RISK FACTORS FOR CARDIOVASCULAR MULTIMORBIDITY IN UK WOMEN



Jae Won Suh

St Catherine's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2022

Abstract

Background: Multimorbidity is the presence of two or more chronic diseases in one person, and in this thesis, cardiovascular multimorbidity (CVM) refers to the co-occurrence of two or more cardiovascular disease subtypes. Multimorbidity is a major public health concern in the UK and CVM is a key component. However, there is currently limited large-scale, prospective evidence on risk factors for CVM.

Aim: This thesis aims to provide a broad overview of CVM, its component conditions and potential risk factors in UK women.

Methods: This was a population-based prospective cohort study of 1.2 million women aged 50-64 years at recruitment in 1996-2001. Using linked hospital admission and death records up to 31 December 2016, each participant was followed-up for nineteen subtypes of incident cardiovascular disease. CVM was defined either by counting the number of cardiovascular diseases in one person or as specific pairs of diseases. Established cardiovascular risk factors, socioeconomic, and reproductive factors were investigated.

Main results: More than 40% of UK women with any incident cardiovascular disease developed CVM (≥2 cardiovascular diseases). The most common components of CVM were ischaemic heart disease (in 55% of CVM cases), atrial fibrillation (46%), heart failure (30%), and other arrhythmias (28%). Smoking, obesity, diabetes, hypertension, and hyperlipidaemia was each associated with 1.7-2.5 times the risk of CVM. Taller height was associated with higher CVM risk, which appeared to be driven by the association between height and atrial fibrillation. There was evidence of somewhat higher CVM risks associated with lower educational attainment and higher area deprivation. However, 80% or more of each of the observed associations was accounted for by four health behaviours measured at recruitment: smoking, alcohol consumption, physical activity, and BMI. Similarly, over 75% of each of the associations between reproductive factors (age at menarche, parity, and breastfeeding) and CVM was accounted for by those four factors.

Conclusion: Roughly 1 in every 2.5 UK women with incident cardiovascular disease had CVM. Established cardiovascular risk factors were also independent risk factors for CVM. Much of the difference in CVM risk by levels of socioeconomic and reproductive factors could be attributed to known modifiable cardiovascular risk factors.

Thesis word count (excluding references, appendices, figures, and tables): 49,236

Acknowledgements

Dominus illuminatio mea – the Lord is my light (Psalm 27:1). I thank God for giving me life in Christ and guiding me to complete every good work during my time at Oxford. There will never be enough words to express my thanks to all those who have made this thesis possible. I would first like to thank the participants of the Million Women Study who have generously provided their data and to the Nuffield Department of Population Health for funding my research. My sincerest thanks go to my supervisors, Dr. Ben Cairns and Dr. Lucy Wright, who have been unwavering in their support for me, both as a researcher and as a person. Thank you to my colleagues at the Cancer Epidemiology Unit and the wider Department who have been willing and eager to share their skills and insights with me and for providing moral support. I would also like to thank all my wonderful friends (too many to list here!) who cheered me on and were there for me through the highs and lows of my DPhil. My deepest thanks to Mom, Dad, Jae Hee, and my grandparents for loving me unconditionally and helping me to be who I am today. Finally, I am eternally grateful to my thoroughly amazing husband, Dr. Jinwoo Leem, who has been my rock and strength through the years. This thesis would not have been possible without you. Thank you, and I love you.

Contents

Abstract ii
Acknowledgements iii
List of Figures x
List of Tables xiii
List of Abbreviations xvi
1 Introduction 1
1.1 Introduction 2
1.2 Cardiovascular disease 2
1.2.1 Definition 2
1.2.2 Prevalence and mortality 4
1.2.3 Pathophysiology and risk factors 4
1.2.4 Cardiovascular disease in women 6
1.3 Multimorbidity 8
1.3.1 Definition 8
1.3.2 Prevalence 10
1.3.3 Burden on patients and healthcare 11
1.3.4 Risk factors 12
1.4 Cardiovascular multimorbidity 13
1.4.1 Definition and prevalence 13
1.4.2 Significance in wider multimorbidity 13
1.4.3 Burden on patients and health care 14
1.4.4 Research gaps 15
1.5 Aim and research questions 16
1.6 Thesis outline 16
1.7 Role of the author in the preparation of this thesis 18
1.8 References 19
2 Risk factors for cardiovascular multimorbidity: a systematic review of prospective studies 30
2.1 Introduction 31
2.1.1 Background 31
2.1.2 Objectives 31
2.2 Methods 33
2.2.1 Registration of protocol 33
2.2.2 Eligibility criteria 33
2.2.3 Information sources 36

2.2.4 Search strategy 36
2.2.5 Selection process 37
2.2.6 Data items and collection process 37
2.2.7 Risk of bias assessment 38
2.2.8 Effect measures and synthesis methods 38
2.3 Results 39
2.3.1 Study selection 39
2.3.2 Study characteristics 41
2.3.3 Risk of bias in studies 43
2.3.4 Methods used to investigate risk factors for multimorbidity 45
2.3.5 Risk factors for overall multimorbidity 48
2.3.6 Risk factors for cardiometabolic multimorbidity 51
2.4 Discussion 54
2.4.1 Key findings in context 54
2.4.2 Definition of multimorbidity 56
2.4.3 Strengths and limitations 57
2.4.4 Implications for future research 58
2.4.5 Implications for this thesis 59
2.4.6 Conclusions 60
2.5 References 61
3 The Million Women Study and Methods 64
3.1 Introduction 65
3.2 The Million Women Study 65
3.2.1 Overall study design 65
3.2.2 Ethics 65
3.2.3 Funding 66
3.2.4 Recruitment 66
3.2.5 Follow-up 73
3.2.6 Cardiovascular research in the Million Women Study 78
3.3 Methods 79
3.3.1 Outcomes 79
3.3.2 Reliability of individual vascular outcomes in the Million Women Study 84
3.3.3 Risk factors 85
3.3.4 Statistical methods 85
3.4 Conclusions 89

3.5 References 90
4 Descriptive study of cardiovascular multimorbidity in UK women 96
4.1 Introduction 97
4.1.1 Background 97
4.1.2 Objective and research questions 97
4.2 Methods 98
4.2.1 Study design and participants 98
4.2.2 Outcomes and definitions 98
4.2.3 Statistical analysis 99
4.3 Results 104
4.3.1 Incidence of cardiovascular diseases and cardiovascular multimorbidity 104
4.3.2 Incidence of pairs of cardiovascular diseases 106
4.3.3 Co-occurrence of pairs of cardiovascular diseases 108
4.3.4 Average age at first diagnosis 110
4.3.5 The number of comorbidities of individual cardiovascular diseases 113
4.3.6 The five most common cardiovascular comorbidities of individual cardiovascular diseases 116
4.3.7 Cardiovascular multimorbidity over time 116
4.3.8 Coding of rheumatic heart disease 121
4.4 Discussion 126
4.4.1 Key findings and interpretation of results 126
4.4.2 Strengths and limitations 131
4.4.3 Conclusions 132
4.5 References 134
5 Established cardiovascular risk factors and risk of incident cardiovascular multimorbidity 140
5.1 Introduction 141
5.1.1 Background 141
5.1.2 Research questions 141
5.2 Methods 142
5.2.1 Outcomes 142
5.2.2 Risk factors 142
5.2.3 Statistical analyses 143
5.3 Results 147
5.3.1 Baseline characteristics 147
5.3.2 Risk factors for cardiovascular multimorbidity 149

5.3.3 Risk factors for complex cardiovascular multimorbidity 152
5.3.4 Risk factors for selected cardiovascular multimorbidity pairs 154
5.3.5 Sensitivity analyses 158
5.4 Discussion 166
5.4.1 Key findings and interpretation of results 166
5.4.2 Potential effect of changes in the prevalence of risk factors over time 171
5.4.3 Potential role of reverse causation 173
5.4.4 Potential role of residual confounding 174
5.4.5 Strengths and limitations 176
5.4.6 Conclusions 177
5.5 References 178
6 Socioeconomic status and risk of incident cardiovascular multimorbidity 183
6.1 Introduction 184
6.1.1 Background 184
6.1.2 Research questions 184
6.2 Methods 186
6.2.1 Outcomes 186
6.2.2 Risk factors 186
6.2.3 Statistical analyses 188
6.3 Results 191
6.3.1 Baseline characteristics by socioeconomic status 191
6.3.2 Risk of cardiovascular multimorbidity and complex cardiovascular multimorbidity in relation to socioeconomic status 195
6.3.3 Risk of selected cardiovascular multimorbidity pairs in relation to socioeconomic status 197
6.3.4 The role of covariates in accounting for the social gradient in the risk of cardiovascular multimorbidity outcomes 200
6.3.5 Sensitivity analyses 205
6.4 Discussion 207
6.4.1 Key findings and interpretation of results 207
6.4.2 Strengths and limitations 215
6.4.3 Conclusions 216
6.5 References 217
7 Reproductive factors and risk of incident cardiovascular multimorbidity 221
7.1 Introduction 222
7.1.1 Background 222

7.1.2 Research questions....222
7.2 Methods....224
7.2.1 Outcomes....224
7.2.2 Risk factors....224
7.2.3 Statistical analyses....225
7.3 Results....228
7.3.1 Baseline characteristics by reproductive factors....228
7.3.2 Risk of cardiovascular multimorbidity and complex cardiovascular multimorbidity in relation to reproductive factors....234
7.3.3 Risk of selected cardiovascular multimorbidity pairs in relation to reproductive factors....241
7.3.4 The role of covariates in accounting for reproductive factor associations with cardiovascular multimorbidity....244
7.3.5 Sensitivity analyses....251
7.4 Discussion....252
7.4.1 Key findings and interpretation of results....252
7.4.2 Strengths and limitations....258
7.4.3 Conclusions....259
7.5 References....260
8 Summary and implications....265
8.1 Introduction....266
8.2 Summary of main results....267
8.2.1 Descriptive study of cardiovascular multimorbidity (Chapter 4)....267
8.2.2 Established cardiovascular risk factors and cardiovascular multimorbidity (Chapter 5)....268
8.2.3 Socioeconomic status and cardiovascular multimorbidity (Chapter 6)....269
8.2.4 Reproductive factors and cardiovascular multimorbidity (Chapter 7)....270
8.3 Summary of thesis strengths and limitations....271
8.3.1 Key strengths....271
8.3.2 Limitations....272
8.4 Implications for clinical practice....275
8.5 Future research....276
8.5.1 Additional analyses in the Million Women Study....276
8.5.2 Differences by sex....277
8.5.3 Different definitions of multimorbidity....278
8.5.4 Analysis of causal association via Mendelian randomisation....278

8.6 References 280
Appendix A: Systematic literature review 285
Appendix B: Million Women Study recruitment questionnaire 298
Appendix C: Descriptive study of cardiovascular multimorbidity 303
Appendix D: Established cardiovascular risk factors and cardiovascular multimorbidity 322
Appendix E: Socioeconomic status and cardiovascular multimorbidity 333
Appendix F: Reproductive factors and cardiovascular multimorbidity.....354

List of Figures

Figure 1.1 The cardiovascular system and examples of cardiovascular diseases 4
Figure 1.2 Stages of atherosclerosis 5
Figure 2.1 PRISMA flow diagram of database search results as of 5 January 2021 40
Figure 3.1 Location of UK NHS Breast Screening Centres which participated in recruitment for the Million Women Study 68
Figure 3.2 Flowchart showing standard exclusions for the Million Women Study and further exclusions leading to the study sample contributing data for the statistical analyses presented in this thesis 71
Figure 4.1 Incidence of each pair of cardiovascular diseases per 1000 women during the follow-up period 107
Figure 4.2 Ratio of observed to expected number of cases for each pair of cardiovascular diseases 109
Figure 4.3 Box and whiskers plot displaying the distribution of age at first record of each individual cardiovascular disease and cardiovascular multimorbidity 112
Figure 4.4 Crude proportion of the number of cardiovascular comorbidities by each incident individual cardiovascular disease 114
Figure 4.5 Age-standardised proportion of the number of cardiovascular comorbidities by each incident individual cardiovascular disease 115
Figure 4.6 UpSet plot of the top five comorbidities of rheumatic heart disease (rhd) 122
Figure 4.7 Distribution of 4-character ICD-10 codes for incident rheumatic heart disease in hospital admission data in the Million Women Study 123
Figure 4.8 UpSet plot of the top five comorbidities of mitral valve stenosis (I05.0) 124
Figure 5.1 Relative risks of incident cardiovascular multimorbidity in relation to cardiovascular risk factors 151
Figure 5.2 Relative risks of incident complex cardiovascular multimorbidity in relation to cardiovascular risk factors 153
Figure 5.3 Relative risks of CVM pairs including ischaemic heart disease in relation to cardiovascular risk factors 156
Figure 5.4 Relative risks of other CVM pairs including cardiac arrhythmias in relation to cardiovascular risk factors 157
Figure 5.5 Risk factor associations of CVM leaving out ischaemic heart disease, atrial fibrillation, or heart failure as sensitivity analyses 159
Figure 5.6 Risk factor associations of CVM leaving out other arrhythmias, stroke, or venous thromboembolism as sensitivity analyses 160
Figure 6.1 Relative risks of incident CVM and complex CVM in relation to education and area deprivation 196
Figure 6.2 Relative risks of CVM pairs including ischaemic heart disease in relation to education and area deprivation 198
Figure 6.3 Relative risks of other CVM pairs including cardiac arrhythmias in relation to education and area deprivation 199
Figure 7.1 Relative risks of incident cardiovascular multimorbidity and complex cardiovascular multimorbidity in relation to age at menarche 235
Figure 7.2 Relative risks of incident cardiovascular multimorbidity and complex cardiovascular multimorbidity in relation to parity 237
Figure 7.3 Relative risk of incident cardiovascular multimorbidity and complex cardiovascular multimorbidity 239

Figure 7.4 Relative risks of CVM pairs including ischaemic heart disease in relation to reproductive factors 242
Figure 7.5 Relative risks of other CVM pairs including cardiac arrhythmias in relation to reproductive factors 243
Figure A1 Revised Newcastle-Ottawa Quality Assessment Scale for cohort studies 291
Figure C1 UpSet plot of the top five comorbidities of hypertensive heart and renal disease (n = 11,344) 304
Figure C2 UpSet plot of the top five comorbidities of ischaemic heart disease (n = 119,847) 305
Figure C3 UpSet plot of the top five comorbidities of mitral valve disorders (n = 12,518) 306
FigureC4 UpSet plot of the top five comorbidities of aortic stenosis (n = 11,993) 307
Figure C5 UpSet plot of the top five comorbidities of other aortic valve disorders (n = 5,208) 308
Figure C6 UpSet plot of the top five comorbidities of cardiomyopathy (n = 4,074) 309
Figure C7 UpSet plot of the top five comorbidities of other arrhythmias (n = 45,241) 310
Figure C8 UpSet plot of the top five comorbidities of atrial fibrillation (n = 76,058) 311
Figure C9 UpSet plot of the top five comorbidities of heart failure (n = 38,172) 312
Figure C10 UpSet plot of the top five comorbidities of stroke (n = 34,408) 313
Figure C11 UpSet plot of the top five comorbidities of transient cerebral ischaemic attacks (n = 12,550) 314
Figure C12 UpSet plot of the top five comorbidities of other cerebrovascular disease (n = 32,940) 315
Figure C13 UpSet plot of the top five comorbidities of vascular dementia (n = 6,602) 316
Figure C14 UpSet plot of the top five comorbidities of venous thromboembolism (n = 31,309) 317
Figure C15 UpSet plot of the top five comorbidities of aortic aneurysm (n = 6,097) 318
Figure C16 UpSet plot of the top five comorbidities of peripheral vascular disease (n = 13,193) 319
Figure C17 UpSet plot of the top five comorbidities of arterial embolism and thrombosis (n = 4,116) 320
Figure C18 UpSet plot of the top five comorbidities of vascular disorders of the intestine (n = 6,568) 321
Figure D1 Relative risks of ischaemic heart disease, atrial fibrillation, and heart failure in relation to established cardiovascular risk factors 325
Figure D2 Relative risks of other arrhythmias, stroke, and venous thromboembolism in relation to established cardiovascular risk factors 326
Figure E1 Relative risks of CVM minus ischaemic heart disease, atrial fibrillation, or heart failure in relation to socioeconomic factors 346
Figure E2 Relative risks of CVM minus other arrhythmias, stroke, or venous thromboembolism in relation to socioeconomic factors 347
Figure E3 Relative risks of ischaemic heart disease, atrial fibrillation, and heart failure in relation to socioeconomic factors 348
Figure E4 Relative risks of other arrhythmias, stroke, and venous thromboembolism in relation to socioeconomic factors 349
Figure F1 Relative risks of CVM minus ischaemic heart disease, atrial fibrillation, or heart failure in relation to reproductive factors 358

Figure F2 Relative risks of CVM minus other arrhythmias, stroke, or venous thromboembolism in relation to reproductive factors 359
Figure F3 Relative risks of ischaemic heart disease, atrial fibrillation, and heart failure in relation to reproductive factors 360
Figure F4 Relative risks of other arrhythmias, stroke, and venous thromboembolism in relation to reproductive factors 361

List of Tables

Table 2.1 Eligibility criteria 34
Table 2.2 Details and examples of exposures (i.e. established cardiovascular risk factors) of interest* 35
Table 2.3 Characteristics of studies included in the systematic review.....42
Table 2.4 Quality assessment for each study using a modified version of the Newcastle-Ottawa Scale 44
Table 2.5 Definition of multimorbidity in each study 46
Table 2.6 Direction of association between cardiovascular risk factors and multimorbidity 50
Table 2.7 Main results from studies specifically investigating cardiometabolic multimorbidity.....53
Table 3.1 Selected characteristics of all study participants at recruitment, after standard exclusions 72
Table 3.2 Overview of all electronically linked data available for follow-up in the Million Women Study at the time of data extraction for this thesis.....74
Table 3.3 List of selected individual cardiovascular diseases, corresponding diagnostic codes, and number of events during follow-up in the study cohort, after standard exclusions only.....81
Table 4.1 A 2x2 contingency table illustrating the relationship between the incidence of diseases A and B at the end of follow-up. 101
Table 4.2 Number and proportion of incident cardiovascular disease and cardiovascular multimorbidity at the end of follow-up.....105
Table 4.3 Average age of onset of cardiovascular multimorbidity and individual cardiovascular diseases (sorted by age in descending order).....111
Table 4.4 Kaplan-Meier failure estimates for cardiovascular multimorbidity and individual cardiovascular diseases at ages 60, 65, 70, 75, and 80 years in all women.....117
Table 4.5 Kaplan-Meier failure estimates for individual cardiovascular diseases in women who had a record of cardiovascular multimorbidity during follow-up 119
Table 4.6 Kaplan-Meier failure estimates for individual cardiovascular diseases in women who had a record of complex cardiovascular multimorbidity during follow-up 120
Table 4.7 Number and proportion of cardiovascular diseases that comprise CVM at the end of follow-up, after leaving out rheumatic heart disease from the definition of CVM 125
Table 5.1 Baseline characteristics in all women and by their incident CVM status at the end of follow-up 148
Table 5.2 Sensitivity analyses controlling for potential reverse causation bias in comparison to the main results.....162
Table 5.3 Relative risks of incident cardiovascular multimorbidity outcomes in relation to alcohol consumption, including non-drinkers as a sensitivity analysis.....165
Table 6.1 Baseline characteristics of study participants by educational attainment 192
Table 6.2 Baseline characteristics of participant by area deprivation level.....194
Table 6.3 Relationship of education with risk of CVM and complex CVM after adjustment for various covariates 201
Table 6.4 Relationship of area deprivation with risk of CVM and complex CVM after adjustment for various covariates.....203
Table 7.1 Baseline characteristics of study participants by age at menarche 229

Table 7.2 Baseline characteristics of study participants by parity 231
Table 7.3 Baseline characteristics of parous women, overall and by total breastfeeding duration 233
Table 7.4 The combined effect of parity and breastfeeding on incident CVM and complex CVM 240
Table 7.5 Association of age at menarche with incident cardiovascular multimorbidity after adjustment for various covariates 246
Table 7.6 Association of parity with incident cardiovascular multimorbidity after adjustment for various covariates 248
Table 7.7 Association of total breastfeeding duration with incident cardiovascular multimorbidity in parous women after adjustment for various covariates 250
Table A1 CINAHL search terms 286
Table A2 EMBASE search terms 287
Table A3 MEDLINE search terms 288
Table A4 Web of Science search terms 289
Table A5 Full list of data extraction items 290
Table A6 Statistical methods and data sources used in each study 292
Table A7 Main findings of each study 293
Table A8 List of diseases to define multimorbidity and exposures investigated in each study 296
Table D1 Number and proportion of missing values for each risk factor and adjustment variable 323
Table D2 Comparison of multivariable hazard ratios for incident IHD-AF to the averaged effects calculated from HRs for component individual CVDs 327
Table D3 Comparison of multivariable hazard ratios for incident IHD-HF to the averaged effects calculated from HRs for component individual CVDs 328
Table D4 Comparison of multivariable hazard ratios for incident IHD-ARR to the averaged effects calculated from HRs for component individual CVDs 329
Table D5 Comparison of multivariable hazard ratios for incident AF-HF to the averaged effects calculated from HRs for component individual CVDs 330
Table D6 Comparison of multivariable hazard ratios for incident ARR-HF to the averaged effects calculated from HRs for component individual CVDs 331
Table D7 Comparison of multivariable hazard ratios for incident AF-STR to the averaged effects calculated from HRs for component individual CVDs 332
Table E1 IHD-AF incidence by education after adjustment for different covariates 334
Table E2 IHD-AF incidence by area deprivation after adjustment for different covariates 335
Table E3 IHD-HF incidence by education after adjustment for different covariates 336
Table E4 IHD-HF incidence by area deprivation after adjustment for different covariates 337
Table E5 IHD-ARR incidence by education after adjustment for different covariates 338
Table E6 IHD-ARR incidence by area deprivation after adjustment for different covariates 339
Table E7 AF-HF incidence by education after adjustment for different covariates 340
Table E8 AF-HF incidence by area deprivation after adjustment for different covariates 341
Table E9 HF-ARR incidence by education after adjustment for different covariates 342

Table E10 HF-ARR incidence by area deprivation after adjustment for different covariates 343
Table E11 AF-STR incidence by education after adjustment for different covariates 344
Table E12 AF-STR incidence by area deprivation after adjustment for different covariates 345
Table E13 CVM and complex CVM incidence by education, excluding women who were being treated for hypertension or diabetes at baseline 350
Table E14 CVM and complex CVM incidence by area deprivation, excluding women who were being treated for hypertension or diabetes at baseline 351
Table E15 CVM and complex CVM incidence by education, restricted to never smokers 352
Table E16 CVM and complex CVM incidence by area deprivation, restricted to never smokers.....353
Table F1 Association of age at menarche with complex cardiovascular multimorbidity risk after adjustment for various covariates.....355
Table F2 Association of parity with complex cardiovascular multimorbidity risk after adjustment for various covariates.....356
Table F3 Association of total breastfeeding duration with complex cardiovascular multimorbidity risk in parous women after adjustment for various covariates.....357

List of Abbreviations

AA	Aortic aneurysm
AF	Atrial fibrillation
ALSWH	Australian Longitudinal Study on Women's Health
APC	Admitted Patient Care
ARR	Other arrhythmias
AS	Aortic stenosis
ATE	Arterial embolism and thrombosis
BMI	Body Mass Index
CBVD	Other cerebrovascular disease
CI	Confidence Interval
CM	Cardiomyopathy
COPD	Chronic obstructive pulmonary disease
CPRD	Clinical Practice Research Datalink
CVD	Cardiovascular disease
CVM	Cardiovascular multimorbidity
DQMI	Data Quality Maturity Index
ELSA	English longitudinal study on Ageing
FCI	Floated confidence interval
GP	General Practitioner
HDL	High-density lipoprotein
HES	Hospital Episode Statistics
HF	Heart failure
HR	Hazard Ratio
HTND	Hypertensive heart and renal disease
ICD-10	International Classification of Diseases, 10th revision
IHD	Ischaemic heart disease
ISD Scotland	Information Services Division Scotland

IQR	Interquartile range
LDL	Low-density lipoprotein
LR statistic	Likelihood-ratio χ^2 test statistic
MET	Metabolic equivalent of task
MM	Multimorbidity
MR	Mendelian Randomisation
MVD	Mitral valve disorders
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NOS	Newcastle-Ottawa scale
O/E	Observed over expected frequency ratio
ONS	Office for National Statistics
OPCS-4	Office of Population Censuses and Surveys classification of surgical operations and procedures, 4th revision
OR	Odds Ratio
OVD	Other aortic valve disorders
PATH	Atlantic Partnership for Tomorrow's Health
PVD	Peripheral vascular disease
RHD	Rheumatic heart disease
SD	Standard deviation
SMR	Scottish morbidity records
TIA	Transient cerebral ischaemic attacks
UK	United Kingdom
USA	United States of America
VAD	Vascular dementia
VDI	Vascular disorders of the intestine
VTE	Venous thromboembolism
WHO	World Health Organisation

1 Introduction

1.1 Introduction

This chapter presents the conceptual background for this thesis, its aims, and an overview of its structure. An introduction to cardiovascular disease (CVD) in Section 1.2 and to multimorbidity in Section 1.3 is followed by a discussion of cardiovascular multimorbidity (CVM) in Section 1.4. The aims of this thesis are outlined in Section 1.5 and the topics addressed in each of the subsequent chapters in Section 1.6.

1.2 Cardiovascular disease

1.2.1 Definition

The cardiovascular, or circulatory, system comprises of the heart and all blood vessels in the human body, including arteries, veins, and capillaries (Figure 1.1). The main functions of the cardiovascular system are delivering oxygen and nutrients to and removing carbon dioxide and waste products from the tissues and organs in the body.

Cardiovascular disease (CVD) is an overarching term for a range of disorders that affect the cardiovascular system (i.e. heart and blood vessels). Individual CVD subtypes generally derive their name from the anatomical location where disease manifests. For instance, cerebrovascular diseases affect blood vessels in the brain, and peripheral arterial disease are disorders of the peripheral circulation (i.e. arms and legs). Without a healthy cardiovascular system, the human body is impaired from performing vital functions such as meeting the demands of physical activity, regulating body temperature, and generally maintaining homeostasis. [1] Depending on the location and severity of disease, some CVD subtypes can cause long-term damage to vital organs and even death if not treated immediately.

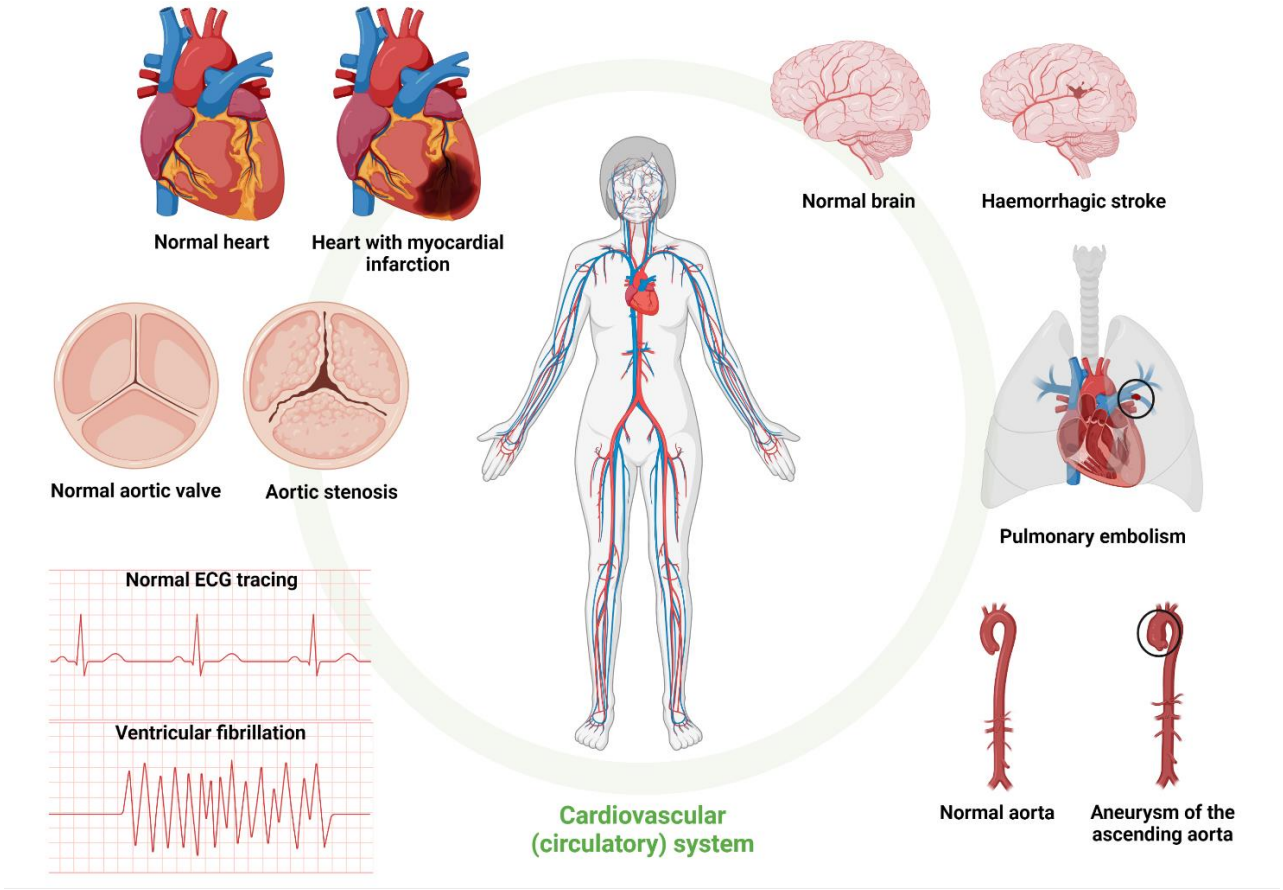


Figure 1.1 The cardiovascular system and examples of cardiovascular diseases

The cardiovascular system consists of the heart and blood vessels. Disorders that affect this system are called cardiovascular diseases (CVDs). CVD can occur in any part of the body, potentially affecting blood flow to vital organs. This figure illustrates some examples of subtypes of CVDs, compared with the healthy state. Clockwise from top left: **Myocardial infarction** occurs when the heart muscle is damaged due to a blockage in the coronary artery, usually due to a blood clot. **Haemorrhagic stroke** is caused by bleeding in or around the brain due to the rupture of a blood vessel, which may cause brain damage. **Pulmonary embolism** is a blockage of blood flow to the lungs, usually by a blood clot that travels to the pulmonary artery from deep veins in the legs. **Aneurysms** are weak spots in a blood vessel wall which bulge abnormally; aneurysms can rupture and cause severe internal bleeding. **Ventricular fibrillation** is a type of abnormal heart rhythm (arrhythmia) that causes the heart to quiver erratically and hence unable to pump blood. **Aortic stenosis** is the obstruction of blood flow across the aortic valve due to the scarring and calcification of the aortic valve. Created with BioRender.com

1.2.2 Prevalence and mortality

CVD is one of the leading causes of death and disability both globally and in the UK. [2, 3] Globally, 523 million people were estimated to be living with CVD in 2019. [4] Approximately 17.9 million people died from CVD in 2019, representing a third of all global deaths; 85% of CVD deaths were due to ischaemic heart disease or stroke.

The British Heart Foundation estimated that 7.6 million individuals in the UK (including 3.6 million women) were living with CVD in 2021. [5] This figure is likely to increase with the ageing population and better survival from CVD. Although the age-standardised absolute CVD mortality per annum declined by 70% between 1979 and 2013 in the UK, CVD still accounts for a quarter of all deaths each year. [6] In 2020, the age-standardised death rate from all subtypes of CVD was 259 per 100,000 people in the UK. [7]

1.2.3 Pathophysiology and risk factors

The primary underlying disease process of many subtypes of CVD is atherosclerosis, a thickening and stiffening of the arterial wall due to the accumulation of plaque (Figure 1.2). [8, 9] Plaque is a deposit of substances such as cholesterol, calcium, fibrin, and cellular debris within the arterial wall, the build-up of which may restrict, and sometimes obstruct, blood flow to organs and tissues. In the coronary arteries, this can eventually lead to several different CVD subtypes: myocardial infarction [9], cardiac arrhythmias [10], and heart failure [11], while in the iliac and femoral arteries, it can lead to peripheral arterial disease [8]. Atherosclerosis, along with other factors that weaken the arterial walls, can also lead to an aneurysm (an abnormal bulging of the vessel wall) which may rupture and cause potentially fatal internal bleeding. [12]

Unstable plaque can sometimes rupture into the blood stream and a blood clot (called a thrombus) can form at the site of the rupture (in a process called thrombosis). [9] It is

especially dangerous if the blood clot completely occludes the artery or if it breaks off and travels downstream to block smaller vessels (thromboembolism). For example, a blood clot that suddenly and completely blocks a coronary artery can damage the heart muscle and cause myocardial infarction. [13] Blood clots lodged in cerebral arteries cause ischaemic stroke. [14] A blood clot can also form in the veins (deep vein thrombosis), and if it travels to the lungs and blocks an artery, this can cause pulmonary embolism. [15]

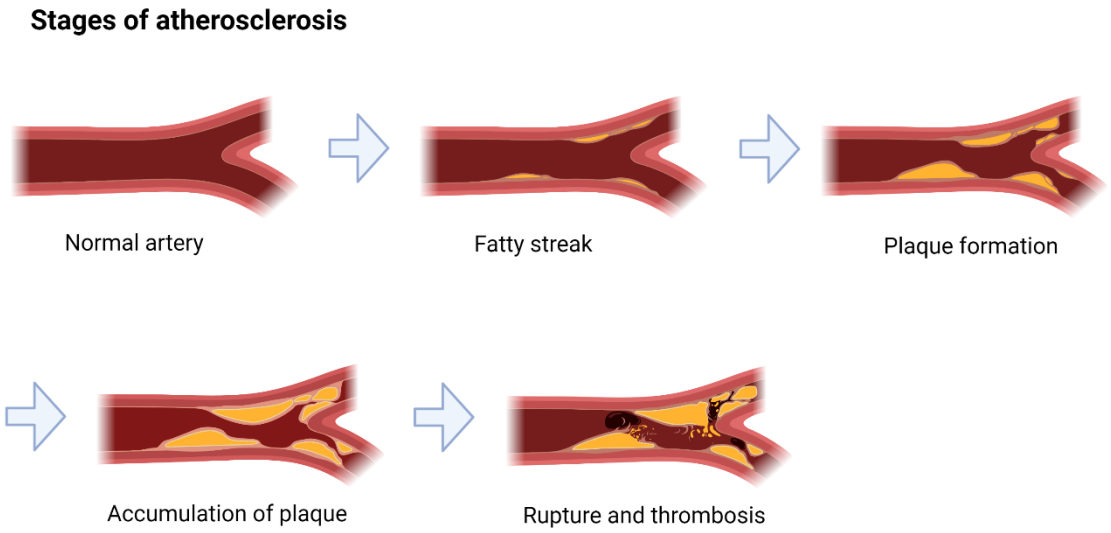


Figure 1.2 Stages of atherosclerosis

Arteries are blood vessels that carry blood from the heart to the rest of the body. The inside of an artery is lined by a single layer of cells called the endothelium, which can be damaged due to physical trauma (e.g. high blood pressure) or chemical stressors in the blood (e.g. high cholesterol, high blood sugar). Macrophages move across the endothelial lining and into the arterial wall at the site of the injury – here they transform into foam cells and collect fatty materials such as cholesterol. Aggregates of foam cells are first visible as fatty streaks, which are present in almost all people by adolescence. Over decades, foam cells, cell debris, calcium, fibrin, and other substances can aggregate to form plaque. As plaque accumulates, it can cause the vessel wall to bulge out, restricting blood flow to the tissues and organs the artery supplies. If the plaque ruptures, a thrombus (blood clot) forms at the site of the rupture. A thrombus can further obstruct the artery, or sometimes suddenly and completely block the artery. Pieces of a thrombus can also break away and block a downstream artery. Created with BioRender.com

Factors that initiate and/or exacerbate atherosclerosis and other disease-causing mechanisms that damage the heart and blood vessels are risk factors for CVD. [16-18] Non-modifiable cardiovascular risk factors include older age, family history of CVD, ethnicity, and male sex. [17] Key modifiable cardiovascular risk factors include long-term high blood pressure, diabetes, high blood cholesterol, obesity, physical inactivity, unhealthy diet, and excessive alcohol consumption. [16-18] Other known CVD risk factors include infection (e.g. in the case of rheumatic heart disease) [19, 20], psychosocial factors (e.g. depression, chronic stress) [16, 21], low socioeconomic position [22-24], and environmental factors (e.g. air pollution) [25]. Furthermore, some CVD subtypes can increase the risk of other CVD subtypes. For example, aortic stenosis reduces the blood flow to the aorta and the rest of the body and can eventually lead to heart failure. [26] Atrial fibrillation also increases the risk of stroke because the irregular heart rhythm may cause a blood clot to form in the left atrium, which can travel to the brain, obstruct blood flow, and cause ischaemic stroke. [27, 28] Finally, CVD subtypes share risk factors with other chronic diseases [29, 30], and some treatments for other chronic diseases are also known risk factors for CVD (e.g. the cardiotoxicity of anthracycline chemotherapy [31]).

1.2.4 Cardiovascular disease in women

More than a third of all deaths in women worldwide are due to CVD; in 2019, 275 million women were diagnosed with CVD and 8.9 million died from it. [32] However, women remain understudied and underrepresented in cardiovascular research. [33] To illustrate, even among relatively recent cardiovascular trials conducted between 2010 to 2017 (n=740), the overall percentage of women was only 38%. [34] The lowest female participation rate was in those aged 61 to 65 years old (26%), despite the largest number of trials conducted in that age group (n=218).

Both women and their physicians tend to underestimate the risk of CVD in women. [35, 36] Perhaps partly because of this lack of awareness, acute coronary heart syndrome tends to be detected later in women than in men. [37] In a UK study, women were 50% more likely to receive an initial incorrect diagnosis following an acute myocardial infarction than men, and those who received a wrong diagnosis were 70% more likely to die after 30 days compared with those who received the correct diagnosis. [38] Women also received less intensive treatment for coronary heart disease, leading to poorer prognosis on average than in men. [39-41] Differences in the detection, treatment, and prognosis of CVD by sex are not limited to coronary heart disease. [42-44] For example, a pooled analysis of international trial data showed that although ischaemic stroke survival was higher in women than in men, among those who survived, women were more likely to be disabled and have a worse quality of life. [44] That study found that although women were more likely to be admitted to an acute stroke unit, they were less likely to be intubated or admitted to an intensive care unit.

Previous studies have shown that cardiovascular risk factor associations are sometimes substantially higher in women than in men. [45] Meta-analyses of prospective cohort studies found that women with diabetes had a 44% higher risk of coronary heart disease than men with diabetes [46], and that women who were current smokers had a 25% higher risk of coronary heart disease than men who were current smokers [47]. Women's CVD risk also appears to be influenced by the reproductive cycle and other factors specific to women. [32] For example, gestational diabetes (a pregnancy complication) was associated with a 25% higher risk of CVD later in life. [48] Lipid profile changes after menopause have been associated with increased CVD risk, although it is unclear if this is due to hormonal changes or a result of ageing. [49] Moreover, breast cancer therapies have been

associated with increased risk of CVD for at least a decade after the initial cancer diagnosis. [50, 51]

In conclusion, existing evidence suggests that women and men have different CVD risk, prognosis, and risk factors. Therefore, investigating sex-specific risk factors for CVD may have important implications for estimating women's CVD risk, targeting treatment for women, and understanding sex differences in CVD.

1.3 Multimorbidity

1.3.1 Definition

Feinstein coined the term comorbidity in 1970 as "any distinct additional entity that has existed or may occur during the clinical course of a patient who has the index disease under study" [52], implying that a comorbidity is relevant only insofar as it affects the treatment and prognosis of the index disease. The terms 'multimorbidity' and 'comorbidity' were initially used somewhat interchangeably, but in 1996 van den Akker et al. [53] proposed a definition for multimorbidity which was distinct from comorbidity: "the co-occurrence of multiple chronic or acute diseases and medical conditions within one person." Although this definition included acute diseases, the current consensus in the scientific literature is that multimorbidity refers to the co-existence of chronic health conditions only. [54-57] Multimorbidity is therefore also known as 'multiple chronic conditions' or 'multiple long-term conditions.' [54] There is some variation in the use of the term 'chronic disease/condition' in the scientific literature [58], but it consistently refers to health conditions of long duration.

In epidemiological research, the types of chronic conditions used to define multimorbidity varied across studies depending on factors such as population characteristics, data availability, and research questions of interest. [59, 60] Studies often included a very broad

range of chronic conditions which were common in the study population or considered important for various reasons. [55, 61] Other studies have been more focussed in scope; for example, some studies have specifically examined cardiometabolic multimorbidity (i.e. the co-occurrence of coronary heart disease, stroke, and/or diabetes) [62-64], and still others have investigated physical-mental multimorbidity (the co-occurrence of at least one physical illness and at least one mental disorder) [65, 66]. In 2018, the Academy of Medical Sciences in the UK defined multimorbidity as "the coexistence of two or more chronic health conditions in an individual, any one of which should be a long-term physical non-communicable disease, a mental health condition of long duration, or a long-term infectious disease" [57], highlighting the different types of chronic conditions which may comprise multimorbidity.

Because no standard list of diseases or method of measuring multimorbidity exists [67, 68], most studies have tailored the definition of multimorbidity to their specific research questions. A 2021 systematic review of 566 studies on multimorbidity found that the number of conditions included to define multimorbidity ranged from 2 to 285 (median 17, IQR 11-23). [55] Furthermore, while 66% of studies used a simple count of diseases to measure multimorbidity, almost 30% of studies used a range of different weighted multimorbidity indices (mostly for predicting health outcomes). The resulting heterogeneity in multimorbidity definitions, which often precludes meaningful comparisons across studies, has been one of the major limitations of multimorbidity research. Several recommendations have been made towards a standard definition of multimorbidity, such as a core list of conditions or a selection of indices to be used for specific research purposes. [55, 60] However, due to discrepancies in data availability and research needs, these recommendations are unlikely to be sufficient for ensuring the comparability of findings across studies.

1.3.2 Prevalence

Multimorbidity has been called "the most common condition experienced by adults" due to the high prevalence of multiple chronic diseases in adults of all ages. [69] A 2021 systematic review and meta-analysis of prevalence studies on multimorbidity found that the pooled prevalence of multimorbidity was 42.4% (95% CI 38.9-46.0), but there was very high heterogeneity between studies (I^2>99%). [70] The mean age of participants and the number of diseases selected to define multimorbidity accounted for about half of the heterogeneity. Multimorbidity prevalence increased with age. The pooled prevalence of multimorbidity in studies with mean participant age of 59-73 years and ≥74 years was 47.6% (42.5-52.8) and 67.0% (60.4-72.9), respectively. Studies which allowed a higher number of conditions to potentially comprise multimorbidity also reported higher prevalence estimates. The pooled prevalence of multimorbidity in studies which used <9 conditions to define multimorbidity was 30.1% (24.9-35.7), whereas the estimate for studies that used 9-19 conditions was 43.7% (39.5-48.0).

In the UK, Barnett et al. [71] investigated 1.75 million primary care patients in Scotland in 2007 (this study was not included in the 2021 meta-analysis). Using a definition of two or more of 40 specified long-term health conditions, the study found that 23% patients were multimorbid overall, and that two out of three people aged 65 years and older were multimorbid. In a more recent UK study based in the Clinical Practice Research Datalink (CPRD), the estimated prevalence of multimorbidity was 27% in a random sample of 403,985 adult patients registered with a general practice on 1 January 2012, while the prevalence of multimorbidity was 54.8 % (95% CI 54.4-55.3) in those aged 65-74 years and 74.0% (73.4-74.4) in those aged 75-84 years. [72] This study defined multimorbidity as two or more of 36 conditions (largely based on the list of conditions used by Barnett et al.) and was included in the 2021 meta-analysis.

It has been estimated that between 2015 and 2035, the prevalence of multimorbidity in those aged 65 years or older will increase from 54% to 68% in the UK, and that the number of people with complex multimorbidity (i.e. four or more chronic conditions) will double. [73] As such, rising multimorbidity – particularly complex multimorbidity – prevalence has been highlighted as a major public health concern in the UK.

1.3.3 Burden on patients and healthcare

Multimorbidity is currently an urgent public health concern in the UK, as highlighted in a 2020 editorial in *The BMJ*. [74] Multimorbidity poses a heavy burden on patients, as it is associated with a range of negative health outcomes such as higher mortality [75], lower quality of life [76, 77], impaired physical function [78], and higher risk of depression [79]. There is some evidence that the negative impact of multimorbidity may worsen with each additional condition. For example, a meta-analysis of 26 studies in older adults found that compared to individuals without multimorbidity (i.e. 0 or 1 condition), those with ≥2 and ≥3 morbidities had 1.73 and 2.72 times the risk of death, respectively. [75] Furthermore, accessing, obtaining, and managing treatment for multimorbidity can be an added burden for patients, their caregivers, and their social networks. [80-82] Patients and caregivers reported frustration due to poor communication and coordination among multiple healthcare providers, and the challenges of adhering to complex treatment recommendations for multiple conditions.

A 2021 systematic review of studies in the UK found that multimorbidity was consistently associated with higher rates of healthcare utilisation and higher costs of care. [83] For example, a study of 180,815 Scottish general practice patients with linked hospital admission data found that patients with complex multimorbidity (≥4 or more diseases) accounted for more than a third of unplanned hospital admissions, despite comprising only 10% of the study sample. [84] The treatment of multimorbidity was often expensive due to

the prescription of multiple drugs and the increased frequency, duration, and complexity of healthcare appointments (e.g. hospital admission). [85-87]

Patients with multimorbidity have often been excluded from randomised controlled trials. [88, 89] Treatment guidelines based on such trials are therefore not necessarily optimal for the care of multimorbid patients. [90] Simultaneously implementing several clinical guidelines developed for individual conditions for a patient with multiple conditions can result in conflicting or burdensome recommendations and even unforeseen drug-drug or drug-disease interactions. [91] Moreover, patients with multimorbidity who see different specialists to address their individual diseases may experience fragmented care and lack of data sharing between the different healthcare providers. [92] In 2016, the National Institute for Health and Care Excellence (NICE) in the UK published guidance on managing patients with multimorbidity to address these issues. [93] However, lack of good quality evidence proved to be one of several limitations. [94]. A 2021 Cochrane review on interventions for multimorbidity management in primary care also highlighted this research gap. [95]

1.3.4 Risk factors

The 2018 report of the Academy of Medical Sciences in the UK highlighted the need for research into the risk factors and patterns of multimorbidity. [57] A systematic review of risk factors for multimorbidity in primary care published in 2014 found that older age, female sex, obesity, and low socioeconomic position were associated with higher multimorbidity across the included studies. [96] All studies included in this review were conducted in high-income countries, mostly in Europe and in the United States of America (USA). However, these studies were predominantly cross-sectional in design and therefore unable to establish the temporal relationship between risk factors and multimorbidity. No systematic review has been published since then to update the available evidence on risk

factors for multimorbidity in the scientific literature, although in recent years, more prospective studies have been published on this topic.

1.4 Cardiovascular multimorbidity

1.4.1 Definition and prevalence

As described in Section 1.2.3 of this chapter, CVD subtypes share risk factors and may sometimes cause one another; this means that CVD subtypes are often found together, particularly in older adults. This thesis defined the co-occurrence of multiple CVD subtypes as cardiovascular multimorbidity (CVM). Data from Medicare beneficiaries in the USA aged ≥65 years in 2012 showed that among 8.7 million patients with ischaemic heart disease, 36% had a diagnosis of heart failure, while 19% had a diagnosis of atrial fibrillation. [97] Moreover, among 1.1 million patients with stroke, 58% had ischaemic heart disease and 37% had heart failure. In the UK, a study in the CPRD found that among 229,205 individuals who had an incident non-fatal CVD between 2000 and 2014, the age and sex-standardised (to the 2013 European Standard Population) prevalence of cardiac arrhythmia was 7.1% (95% CI 6.0-9.6), heart failure was 4.1% (3.8-6.3), and peripheral arterial disease was 3.8% (3.6-6.0). [98] Non-fatal CVD was defined as a diagnosis of ischaemic heart disease, stroke, or transient ischaemic attack in that study.

1.4.2 Significance in wider multimorbidity

Co-occurring CVDs subtypes (e.g. coronary heart disease, stroke) and cardiovascular risk factors (e.g. hypertension, hyperlipidaemia) are the most common components of multimorbidity in older adults across various populations. [96, 97, 99, 100] In a study in the Rochester Epidemiology Project of 138,858 residents of Olmsted county in the USA, all top ten most prevalent pairs and trios of conditions in those aged ≥60 years included at least one CVD subtype (coronary artery disease or arrhythmia) or cardiovascular risk

factor (e.g. hypertension). [100] Similarly, among US Medicare beneficiaries in 2012, the top five most prevalent pairs and trios of chronic conditions in people aged ≥65 years were almost entirely comprised of ischaemic heart disease and cardiovascular risk factors, the only exception being arthritis. [97] Among 19 million individuals with multimorbidity (≥2 chronic conditions), the two most prevalent pairs of conditions were high cholesterol and high blood pressure (57%) and ischaemic heart disease and high blood pressure (37%).

Co-occurring CVDs (i.e. CVM) also featured prominently in multimorbidity patterns identified by studies which investigated evidence of non-random associations among diseases rather than prevalence alone. [101-105] If chronic conditions occur together non-randomly (i.e. more often than expected by chance), that may indicate the presence of shared pathophysiological mechanisms which could potentially be targeted for secondary disease prevention. [101, 106, 107] Regardless of differences in population characteristics and statistical methods used, virtually all studies (conducted without any a priori hypotheses) found that multiple CVD subtypes were found in common multimorbidity patterns more often than expected by chance. [101-105] Furthermore, studies tracing the development of multimorbidity over time found that heart diseases, hypertension, and diabetes were among the most common initial disease in those who later developed multimorbidity [108-113]

1.4.3 Burden on patients and health care

Not only was CVD found frequently and early in the development of multimorbidity, having CVM was also associated with a range of negative health outcomes such as higher mortality [64, 114], poorer quality of life [115], increased risk of cognitive impairment [116], and heavier healthcare utilisation [117]. For example, the Emerging Risk Factors Collaboration found in 689,300 participants (mean age 53 [SD 9] years) pooled from 91 cohorts in Europe or North America that having multiple cardiovascular or metabolic

disorders was associated with a multiplicative increase in mortality risk. [64] Compared to those without a history of myocardial infarction, stroke, or diabetes at baseline, the risk of all-cause mortality was approximately twice as high in individuals with any one of those conditions, nearly four times as high in those with any two, and seven times as high in those with all three. Similar findings were reported in the UK Biobank cohort. [64] In addition, in a random sample of patients over the age of 85 years in the UK CPRD, the trio of coronary heart disease, atrial fibrillation, and heart failure was the multimorbidity pattern associated with the highest rate of mortality, primary care consultations, hospitalisations, and repeat prescriptions. [117] In patients aged 65-84 years, the co-occurrence of coronary heart disease, depression, and pain had the highest rate of death and healthcare utilisation.

1.4.4 Research gaps

CVM is a major component of multimorbidity and is associated with a range of negative health outcomes. [64, 114-117] Despite the importance of CVM, few studies have described the extent to which CVD subtypes co-occur in detail. [97, 98] Although it is known that CVD shares several risk factors with other major chronic conditions (e.g. cancer [30]) and is often the first condition in the trajectory of multimorbidity development [108, 113], there is little prospective evidence on risk factors for CVM (systematically reviewed in Chapter 2). Previous studies have often been limited to studying a few subtypes of CVDs (e.g. coronary heart disease, stroke). [62, 63, 118, 119] Historically, there has been relatively less CVD research in women than in men [33, 34]; with little evidence on CVM in the scientific literature in general, there is currently even less evidence on CVM risk in women or on female-specific risk factors for CVM. Understanding risk factor associations with CVM in a large prospective cohort study of

women will therefore address a significant research gap and have the potential to inform future prevention measures against multimorbidity.

1.5 Aim and research questions

The overall aim of this thesis was to investigate CVM in women in the prospective Million Women Study cohort. The specific CVM outcomes investigated in this thesis were CVM (i.e. ≥2 CVD subtypes), complex CVM (i.e. ≥4 CVD subtypes), and CVM pairs (i.e. pairs of CVD subtypes), which are described in detail in Section 3.3.1 of Chapter 3 (page 79). The research questions addressed in this thesis were:

1. What is already known in the literature about risk factor associations with CVM?
2. When and how does incident CVM present in middle-aged and older UK women?
3. Are established cardiovascular risk factors, socioeconomic factors, and reproductive factors associated with the risk of CVM, complex CVM, and CVM pairs?
4. To what degree are the associations of socioeconomic and reproductive factors with incident CVM accounted for by known cardiovascular risk factors?

1.6 Thesis outline

Chapter 2 is a systematic review of published research studies investigating the prospective associations between known cardiovascular risk factors and multimorbidity that included CVD. It also identifies the main limitations of previous research and briefly discusses the strengths and weaknesses of the Million Women Study (the data source for this thesis) for addressing these limitations and further investigating risk factors for cardiovascular multimorbidity (CVM).

Chapter 3 is a detailed description of the Million Women Study, a large prospective cohort of middle-aged and older UK women. This chapter also describes the linked administrative health datasets, which provided follow-up data for the participants, and the statistical methods that are used across the subsequent results chapters (Chapters 4 to 7).

Chapter 4 is a descriptive study of CVM in the Million Women Study cohort. It explores which individual CVD subtypes most often feature in incident CVM outcomes, which were CVM (i.e. ≥2 CVDs), complex CVM (i.e. ≥4 CVDs), and CVM pairs. This chapter also explores the most common cardiovascular co-morbidities of each individual CVD and at what age individual CVDs occur in women who have CVM during follow-up.

Chapter 5 is an analytical study of the association between known cardiovascular risk factors and incident CVM outcomes. The investigated risk factors were body mass index, height, smoking, physical activity, alcohol consumption, and treatment for diabetes, hypertension, and high blood cholesterol.

Chapter 6 is an analytical study of the association between educational attainment and area deprivation level and incident CVM outcomes. The potential contribution of different lifestyle and reproductive factors to the observed social inequalities in CVM risk is also quantified.

Chapter 7 is an analytical study of the association between reproductive factors (age at menarche, parity, and breastfeeding children) and the risk of CVM outcomes. The potential contribution of different cardiovascular risk factors to the observed associations between reproductive factors and CVM is also quantified.

Chapter 8 summarises the main findings from this thesis. It also discusses the strengths and limitations of this study, and implications for clinical practice, and potential topics for future research.

1.7 Role of the author in the preparation of this thesis

The overarching goal of investigating risk factors for cardiovascular multimorbidity in the Million Women Study was conceived by my supervisors, Dr. Benjamin Cairns and Dr. Lucy Wright. As part of the application for the Doctor of Philosophy programme at the Nuffield Department of Population Health, I devised the initial project proposal for this thesis with advice and feedback from my supervisors. I led the research presented in this thesis by planning and performing all the analyses and preparing the thesis chapters. My supervisors provided advice and feedback for each of these steps. Nia Roberts from the Bodleian Health Care Libraries contributed to the work presented in Chapter 2 by providing guidance for developing comprehensive search strings for the systematic review. Dr. Mahmuda Jahan Chadni and Dr. Sunjuri Sun also contributed to Chapter 2 by acting as second reviewers for the systematic review. Other researchers from the Cancer Epidemiology Unit and assessors for my Transfer of Status and Confirmation of Status assessments also provided comments and discussion of various aspects of the analyses presented in this thesis.

1.8 References

[1] Klabunde RE. Cardiovascular physiology concepts. 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins/Wolters Kluwer; 2012. 243 p.

[2] Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet. 2020 Oct;396(10258):1204–1222. Available from: *https://doi.org/10.1016/S0140-6736(20)30925-9*.

[3] Roth GA, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2018 Nov;392(10159):1736–1788. Available from: *https://doi.org/-10.1016/S0140-6736(18)32203-7*.

[4] Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. J Am Coll Cardiol. 2020 Dec;76(25):2982–3021. Available from: *https://www.sciencedirect.com/science/article/pii/S0735109720377755*.

[5] British Heart Foundation. UK Factsheet; 2022. Available from: *https://-www.bhf.org.uk/-/media/files/research/heart-statistics/bhf-cvd-statistics-uk-factsheet.pdf?la=en&rev=231fa9066a64451e83b107aa63461a37&hash=1CECE96BC260AB985C5BC119F07267E43825D891*.

[6] Bhatnagar P, Wickramasinghe K, Wilkins E, Townsend N. Trends in the epidemiology of cardiovascular disease in the UK. Heart. 2016 Dec;102(24):1945. Available from: *https://doi.org/10.1136/heartjnl-2016-309573*.

[7] British Heart Foundation. Heart and CirculatoryDisease Statistics 2022; 2022. Available from: *https://www.bhf.org.uk/-/media/files/research/heart-statistics/bhf-statistics-compendium-2022.pdf?la=en&rev=975608be0d814948aff0761df7973a3c&hash=C02E7DFD2202AA03D77A43AD65A72A6090206118*.

[8] Herrington W, Lacey B, Sherliker P, Armitage J, Lewington S. Epidemiology of Atherosclerosis and the Potential to Reduce the Global Burden of Atherothrombotic Disease. Circulation Research. 2016 Feb;118(4):535–546. Available from: *https://doi.org/-10.1161/CIRCRESAHA.115.307611*.

[9] Bentzon JF, Otsuka F, Virmani R, Falk E. Mechanisms of Plaque Formation and Rupture. Circulation Research. 2014 Jun;114(12):1852–1866. Available from: *https://-doi.org/10.1161/CIRCRESAHA.114.302721*.

[10] Ghuran AV, Camm AJ. Ischaemic heart disease presenting as arrhythmias. Br Med Bull. 2001 Oct;59(1):193–210. Available from: *https://doi.org/10.1093/bmb/59.1.193*.

[11] Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the

Heart Failure Association (HFA) of the ESC. European heart journal. 2016 Jul;37:2129–2200.

[12] Golledge J, Norman PE. Atherosclerosis and Abdominal Aortic Aneurysm. Arteriosclerosis, Thrombosis, and Vascular Biology. 2010 Jun;30(6):1075–1077. Available from: *https://doi.org/10.1161/ATVBAHA.110.206573*.

[13] Palasubramaniam J, Wang X, Peter K. Myocardial Infarction–From Atherosclerosis to Thrombosis. Arteriosclerosis, Thrombosis, and Vascular Biology. 2019 Aug;39(8):e176–e185. Available from: *https://doi.org/10.1161/ATVBAHA.119.312578*.

[14] Campbell BCV, De Silva DA, Macleod MR, Coutts SB, Schwamm LH, Davis SM, et al. Ischaemic stroke. Nature Reviews Disease Primers. 2019 Oct;5(1):70. Available from: *https://doi.org/10.1038/s41572-019-0118-8*.

[15] Kearon C. Natural History of Venous Thromboembolism. Circulation. 2003 Jun;107(23_suppl_1):I22–I30. Available from: *https://doi.org/10.1161/-01.CIR.0000078464.82671.78*.

[16] Yusuf S, Joseph P, Rangarajan S, Islam S, Mente A, Hystad P, et al. Modifiable risk factors, cardiovascular disease, and mortality in 155722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. Lancet. 2020 Mar;395:795–808. Available from: *https://doi.org/10.1016/s0140-6736(19)32008-2*.

[17] Wood D. Established and emerging cardiovascular risk factors. American Heart Journal. 2001 Feb;141(2, Supplement):S49–S57. Available from: *https://-www.sciencedirect.com/science/article/pii/S0002870301011516*.

[18] Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts)Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). Eur Heart J. 2016 Aug;37(29):2315–2381. Available from: *https://doi.org/10.1093/eurheartj/-ehw106*.

[19] Szwed P, Gasecka A, Zawadka M, Eyileten C, Postua M, Mazurek T, et al. Infections as Novel Risk Factors of Atherosclerotic Cardiovascular Diseases: Pathophysiological Links and Therapeutic Implications. Journal of clinical medicine. 2021 Jun;10:2539. Available from: *https://doi.org/10.3390/jcm10122539*.

[20] Marijon E, Mirabel M, Celermajer DS, Jouven X. Rheumatic heart disease. Lancet. 2012 Mar;379:953–964. Available from: *https://doi.org/10.1016/s0140-6736(11)61171-9*.

[21] Kuper H, Adami HO, Theorell T, Weiderpass E. Psychosocial determinants of coronary heart disease in middle-aged women: a prospective study in Sweden. American journal of epidemiology. 2006 Aug;164:349–57.

[22] Lawlor DA, Ebrahim S, Davey Smith G. Adverse socioeconomic position across the lifecourse increases coronary heart disease risk cumulatively: findings from the British women's heart and health study. Journal of epidemiology and community health. 2005 Sep;59:785–93.

[23] Manrique-Garcia E, Sidorchuk A, Hallqvist J, Moradi T. Socioeconomic position and incidence of acute myocardial infarction: a meta-analysis. J Epidemiol Community Health. 2011 Apr;65(4):301. Available from: *https://doi.org/10.1136/jech.2009.104075*.

[24] Pujades-Rodriguez M, Timmis A, Stogiannis D, Rapsomaniki E, Denaxas S, Shah A, et al. Socioeconomic deprivation and the incidence of 12 cardiovascular diseases in 1.9 million women and men: implications for risk prediction and prevention. PloS one. 2014 Aug;9:e104671. Available from: *https://doi.org/10.1371/journal.pone.0104671*.

[25] Cosselman KE, Navas-Acien A, Kaufman JD. Environmental factors in cardiovascular disease. Nature Reviews Cardiology. 2015 Nov;12(11):627–642. Available from: *https://doi.org/10.1038/nrcardio.2015.152*.

[26] Zile MR, Gaasch WH. Heart Failure in Aortic Stenosis – Improving Diagnosis and Treatment. N Engl J Med. 2003 May;348(18):1735–1736. Available from: *https://doi.org/-10.1056/NEJMp030035*.

[27] Arboix A. Cardiovascular risk factors for acute stroke: Risk profiles in the different subtypes of ischemic stroke. World J Clin Cases. 2015 May;3(5):418–429. Available from: *https://doi.org/10.12998/wjcc.v3.i5.418*.

[28] Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke. 1991 Aug;22(8):983–988. Available from: *https://doi.org/10.1161/01.STR.22.8.983*.

[29] Parikh NI, Hwang SJ, Larson MG, Meigs JB, Levy D, Fox CS. Cardiovascular Disease Risk Factors in Chronic Kidney Disease: Overall Burden and Rates of Treatment and Control. Arch Intern Med. 2006 Sep;166(17):1884–1891. Available from: *https://-doi.org/10.1001/archinte.166.17.1884*.

[30] Koene RJ, Prizment AE, Blaes A, Konety SH. Shared Risk Factors in Cardiovascular Disease and Cancer. Circulation. 2016 Mar;133(11):1104–1114. Available from: *https://doi.org/10.1161/CIRCULATIONAHA.115.020406*.

[31] Moudgil R, Yeh ETH, Gottlieb RA, Mehta PK. Chapter 4 - Molecular Mechanisms of Anthracycline-Induced Cardiotoxicity. In: Cardio-Oncology. Boston: Academic Press; 2017. p. 55–68. Available from: *https://doi.org/10.1016/B978-0-12-803547-4.00004-5*.

[32] Vogel B, Acevedo M, Appelman Y, Bairey Merz CN, Chieffo A, Figtree GA, et al. The Lancet women and cardiovascular disease Commission: reducing the global burden by 2030. The Lancet. 2021 Jun;397(10292):2385–2438. Available from: *https://doi.org/10.1016/S0140-6736(21)00684-X*.

[33] Ravindran TKS, Teerawattananon Y, Tannenbaum C, Vijayasingham L. Making pharmaceutical research and regulation work for women. BMJ. 2020 Oct;371:m3808. Available from: *https://doi.org/10.1136/bmj.m3808*.

[34] Jin X, Chandramouli C, Allocco B, Gong E, Lam CSP, Yan LL. Women's Participation in Cardiovascular Clinical Trials From 2010 to 2017. Circulation. 2020 Feb;141(7):540–548. Available from: *https://doi.org/10.1161/-CIRCULATIONAHA.119.043594*.

[35] Lehto HR, Lehto S, Havulinna AS, Jousilahti P, Salomaa V. Gender differences in the prevalence, causes and treatment of high cardiovascular risk: findings from the

FINRISK Survey. Eur J Prev Cardiol. 2012 Oct;19(5):1153–1160. Available from: *https://doi.org/10.1177/1741826711422454*.

[36] Mosca L, Linfante AH, Benjamin EJ, Berra K, Hayes SN, Walsh BW, et al. National Study of Physician Awareness and Adherence to Cardiovascular Disease Prevention Guidelines. Circulation. 2005 Feb;111(4):499–510. Available from: *https://-doi.org/10.1161/01.CIR.0000154568.43333.82*.

[37] Elsaesser A, Hamm CW. Acute Coronary Syndrome: The Risk of Being Female. Circulation. 2004 Feb;109(5):565–567. Available from: *https://doi.org/10.1161/-01.CIR.0000116022.77781.26*.

[38] Wu J, Gale CP, Hall M, Dondo TB, Metcalfe E, Oliver G, et al. Impact of initial hospital diagnosis on mortality for acute myocardial infarction: A national cohort study. Eur Heart J Acute Cardiovasc Care. 2018 Mar;7(2):139–148. Available from: *https://-doi.org/10.1177/2048872616661693*.

[39] Wilkinson C, Bebb O, Dondo TB, Munyombwe T, Casadei B, Clarke S, et al. Sex differences in quality indicator attainment for myocardial infarction: a nationwide cohort study. Heart. 2019 Apr;105(7):516. Available from: *http://dx.doi.org/10.1136/heartjnl-2018-313959*.

[40] Blomkalns AL, Chen AY, Hochman JS, Peterson ED, Trynosky K, Diercks DB, et al. Gender disparities in the diagnosis and treatment of non-ST-segment elevation acute coronary syndromes: Large-scale observations from the CRUSADE (Can Rapid Risk Stratification of Unstable Angina Patients Suppress Adverse Outcomes With Early Implementation of the American College of Cardiology/American Heart Association Guidelines) National Quality Improvement Initiative. Journal of the American College of Cardiology. 2005;45(6):832–837. Available from: *https://doi.org/10.1016/-j.jacc.2004.11.055*.

[41] Jneid H, Fonarow GC, Cannon CP, Hernandez AF, Palacios IF, Maree AO, et al. Sex Differences in Medical Care and Early Death After Acute Myocardial Infarction. Circulation. 2008 Dec;118(25):2803–2810. Available from: *https://doi.org/10.1161/-CIRCULATIONAHA.108.789800*.

[42] Dahl S, Hjalmarsson C, Andersson B. Sex differences in risk factors, treatment, and prognosis in acute stroke. Womens Health (Lond Engl). 2020 Jan;16:1745506520952039. Available from: *https://doi.org/10.1177/1745506520952039*.

[43] Westerman S, Wenger N. Gender Differences in Atrial Fibrillation: A Review of Epidemiology, Management, and Outcomes; 2019. Available from: *https://doi.org/-10.2174/1573403x15666181205110624*.

[44] Carcel C, Wang X, Sandset EC, Delcourt C, Arima H, Lindley R, et al. Sex differences in treatment and outcome after stroke. Neurology. 2019 Dec;93(24):e2170. Available from: *https://doi.org/10.1212/WNL.0000000000008615*.

[45] Legato MJ, Johnson PA, Manson JE. Consideration of Sex Differences in Medicine to Improve Health Care and Patient Outcomes. JAMA. 2016 Nov;316(18):1865–1866. Available from: *https://doi.org/10.1001/jama.2016.13995*.

[46] Peters SAE, Huxley RR, Woodward M. Diabetes as risk factor for incident coronary heart disease in women compared with men: a systematic review and meta-

analysis of 64 cohorts including 858,507 individuals and 28,203 coronary events. Diabetologia. 2014 Aug;57:1542–51.

[47] Huxley RR, Woodward M. Cigarette smoking as a risk factor for coronary heart disease in women compared with men: a systematic review and meta-analysis of prospective cohort studies. The Lancet. 2011 Oct;378(9799):1297–1305. Available from: *https://doi.org/10.1016/S0140-6736(11)60781-2*.

[48] Goueslard K, Cottenet J, Mariet AS, Giroud M, Cottin Y, Petit JM, et al. Early cardiovascular events in women with a history of gestational diabetes mellitus. Cardiovascular Diabetology. 2016 Jan;15(1):15. Available from: *https://doi.org/10.1186/-s12933-016-0338-0*.

[49] Janssen I, Powell LH, Crawford S, Lasley B, Sutton-Tyrrell K. Menopause and the Metabolic Syndrome: The Study of Women's Health Across the Nation. Arch Intern Med. 2008 Jul;168(14):1568–1575. Available from: *https://doi.org/10.1001/-archinte.168.14.1568*.

[50] Bradshaw PT, Stevens J, Khankari N, Teitelbaum SL, Neugut AI, Gammon MD. Cardiovascular Disease Mortality Among Breast Cancer Survivors. Epidemiology. 2016;27(1). Available from: *https://doi.org/10.1097/ede.0000000000000394*.

[51] Jones LW, Haykowsky MJ, Swartz JJ, Douglas PS, Mackey JR. Early Breast Cancer Therapy and Cardiovascular Injury. Journal of the American College of Cardiology. 2007 Oct;50(15):1435–1441. Available from: *https://doi.org/10.1016/-j.jacc.2007.06.037*.

[52] Feinstein AR. The pre-therapeutic classification of co-morbidity in chronic disease. Journal of Chronic Diseases. 1970 Dec;23(7):455–468. Available from: *https://doi.org/-10.1016/0021-9681(70)90054-8*.

[53] van den Akker M, Buntinx F, Knottnerus JA. Comorbidity or multimorbidity. null. 1996 Jan;2(2):65–70. Available from: *https://doi.org/10.3109/13814789609162146*.

[54] NIHR Collection. Multiple long-term conditions (multimorbidity): making sense of the evidence; 2021. Available from: *https://doi.org/10.3310/collection_45881*.

[55] Ho ISS, Azcoaga-Lorenzo A, Akbari A, Black C, Davies J, Hodgins P, et al. Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies. The Lancet Public Health. 2021 Aug;6(8):e587–e597. Available from: *https://doi.org/10.1016/S2468-2667(21)00107-9*.

[56] World Health Organization. Multimorbidity: Technical Series on Safer Primary Care. Geneva; 2015. Licence: CC BY-NC-SA 3.0 IGO. Available from: *https://-apps.who.int/iris/handle/10665/252275*.

[57] The Academy of Medical Sciences. Multimorbidity: a priority for global health research. London: The Academy of Medical Sciences; 2018.

[58] Bernell S, Howard SW. Use Your Words Carefully: What Is a Chronic Disease? Frontiers in public health. 2016 Aug;4:159. Available from: *https://doi.org/10.3389/-fpubh.2016.00159*.

[59] Huntley AL, Johnson R, Purdy S, Valderas JM, Salisbury C. Measures of Multimorbidity and Morbidity Burden for Use in Primary Care and Community Settings:

A Systematic Review and Guide. Ann Fam Med. 2012 Mar;10(2):134. Available from: *https://doi.org/10.1370/afm.1363*.

[60] Stirland LE, González-Saavedra L, Mullin DS, Ritchie CW, Muniz-Terrera G, Russ TC. Measuring multimorbidity beyond counting diseases: systematic review of community and population studies and guide to index choice. BMJ. 2020 Feb;368:m160. Available from: *http://www.bmj.com/content/368/bmj.m160.abstract*.

[61] Diederichs C, Berger K, Bartels DB. The measurement of multiple chronic diseases–a systematic review on existing multimorbidity indices. The journals of gerontology Series A, Biological sciences and medical sciences. 2011 Mar;66:301–11. Available from: *https://doi.org/10.1093/gerona/glq208*.

[62] Kivimäki M, Kuosma E, Ferrie JE, Luukkonen R, Nyberg ST, Alfredsson L, et al. Overweight, obesity, and risk of cardiometabolic multimorbidity: pooled analysis of individual-level data for 120813 adults from 16 cohort studies from the USA and Europe. Lancet Public Health. 2017 Jun;2:e277–e285. Available from: *https://doi.org/10.1016/-s2468-2667(17)30074-9*.

[63] Singh-Manoux A, Fayosse A, Sabia S, Tabak A, Shipley M, Dugravot A, et al. Clinical, socioeconomic, and behavioural factors at age 50 years and risk of cardiometabolic multimorbidity and mortality: A cohort study. PLoS Med. 2018 May;15(5):e1002571. Available from: *https://doi.org/10.1371/journal.pmed.1002571*.

[64] The Emerging Risk Factors Collaboration. Association of Cardiometabolic Multimorbidity With Mortality. JAMA. 2015 Jul;314(1):52–60. Available from: *https://-doi.org/10.1001/jama.2015.7008*.

[65] Wang YP, Nunes BP, Coêlho BM, Santana GL, do Nascimento CF, Viana MC, et al. Multilevel Analysis of the Patterns of Physical-Mental Multimorbidity in General Population of SÃ£o Paulo Metropolitan Area, Brazil. Scientific Reports. 2019 Feb;9(1):2390. Available from: *https://doi.org/10.1038/s41598-019-39326-8*.

[66] Romano I, Buchan C, Baiocco-Romano L, Ferro MA. Physical-mental multimorbidity in children and youth: a scoping review. BMJ Open. 2021 May;11(5):e043124. Available from: *http://dx.doi.org/10.1136/bmjopen-2020-043124*.

[67] Holzer BM, Siebenhuener K, Bopp M, Minder CE. Evidence-based design recommendations for prevalence studies on multimorbidity: improving comparability of estimates. Population health metrics. 2017 Mar;15:9. Available from: *https://doi.org/-10.1186/s12963-017-0126-4*.

[68] Busija L, Lim K, Szoeke C, Sanders KM, McCabe MP. Do replicable profiles of multimorbidity exist? Systematic review and synthesis. European journal of epidemiology. 2019 Nov;34:1025–1053. Available from: *https://doi.org/10.1007/s10654-019-00568-5*.

[69] Tinetti ME, Fried TR, Boyd CM. Designing health care for the most common chronic condition–multimorbidity. JAMA. 2012 Jun;307:2493–4. Available from: *https://-doi.org/10.1001/jama.2012.5265*.

[70] Ho ISS, Azcoaga-Lorenzo A, Akbari A, Davies J, Hodgins P, Khunti K, et al. Variation in the estimated prevalence of multimorbidity: systematic review and meta-analysis of 193 international studies. BMJ Open. 2022 Apr;12(4):e057017. Available from: *https://doi.org/10.1136/bmjopen-2021-057017*.

[71] Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet. 2012 Jul;380(9836):37–43. Available from: *https://doi.org/-10.1016/S0140-6736(12)60240-2*.

[72] Cassell A, Edwards D, Harshfield A, Rhodes K, Brimicombe J, Payne R, et al. The epidemiology of multimorbidity in primary care: a retrospective cohort study. Br J Gen Pract. 2018 Apr;68(669):e245. Available from: *http://bjgp.org/content/68/669/-e245.abstract*.

[73] Kingston A, Robinson L, Booth H, Knapp M, Jagger C, for the MODEM project. Projections of multi-morbidity in the older population in England to 2035: estimates from the Population Ageing and Care Simulation (PACSim) model. Age Ageing. 2018 May;47(3):374–380. Available from: *https://doi.org/10.1093/ageing/afx201*.

[74] Whitty CJM, MacEwen C, Goddard A, Alderson D, Marshall M, Calderwood C, et al. Rising to the challenge of multimorbidity. BMJ. 2020 Jan;368:l6964. Available from: *https://doi.org/10.1136/bmj.l6964*.

[75] Nunes BP, Flores TR, Mielke GI, Thumé E, Facchini LA. Multimorbidity and mortality in older adults: A systematic review and meta-analysis. Archives of gerontology and geriatrics. 2016 Nov;67:130–8.

[76] Kanesarajah J, Waller M, Whitty JA, Mishra GD. Multimorbidity and quality of life at mid-life: A systematic review of general population studies. Maturitas. 2018 Mar;109:53–62. Available from: *https://doi.org/10.1016/j.maturitas.2017.12.004*.

[77] Makovski TT, Schmitz S, Zeegers MP, Stranges S, van den Akker M. Multimorbidity and quality of life: Systematic literature review and meta-analysis. Ageing Research Reviews. 2019 Aug;53:100903. Available from: *https://www.sciencedirect.com/-science/article/pii/S1568163719300066*.

[78] Steeves JA, Shiroma EJ, Conger SA, Van Domelen D, Harris TB. Physical activity patterns and multimorbidity burden of older adults with different levels of functional status: NHANES 2003-2006. Disability and Health Journal. 2019;12(3):495–502. Available from: *https://doi.org/10.1016/j.dhjo.2019.02.005*.

[79] Read JR, Sharpe L, Modini M, Dear BF. Multimorbidity and depression: A systematic review and meta-analysis. Journal of Affective Disorders. 2017;221:36–46. Available from: *https://doi.org/10.1016/j.jad.2017.06.009*.

[80] Gill A, Kuluski K, Jaakkimainen L, Naganathan G, Upshur R, Wodchis WP. "Where do we go from here?" Health system frustrations expressed by patients with multimorbidity, their caregivers and family physicians. Healthcare policy. 2014 May;9:73–89. Available from: *https://doi.org/10.12927/hcpol.2014.23811*.

[81] May CR, Eton DT, Boehmer K, Gallacher K, Hunt K, MacDonald S, et al. Rethinking the patient: using Burden of Treatment Theory to understand the changing dynamics of illness. BMC Health Services Research. 2014;14(1):281. Available from: *https://doi.org/10.1186/1472-6963-14-281*.

[82] Payne RA. The epidemiology of polypharmacy. Clin Med. 2016 Oct;16(5):465. Available from: *https://doi.org/10.7861/clinmedicine.16-5-465*.

[83] Soley-Bori M, Ashworth M, Bisquera A, Dodhia H, Lynch R, Wang Y, et al. Impact of multimorbidity on healthcare costs and utilisation: a systematic review of the UK literature. Br J Gen Pract. 2021 Jan;71(702):e39. Available from: *http://bjgp.org/-content/71/702/e39.abstract*.

[84] Payne RA, Abel GA, Guthrie B, Mercer SW. The effect of physical multimorbidity, mental health conditions and socioeconomic deprivation on unplanned admissions to hospital: a retrospective cohort study. CMAJ. 2013 Mar;185(5):E221. Available from: *https://doi.org/10.1503/cmaj.121349*.

[85] Lehnert T, Heider D, Leicht H, Heinrich S, Corrieri S, Luppa M, et al. Review: health care utilization and costs of elderly persons with multiple chronic conditions. Medical care research and review : MCRR. 2011 Aug;68:387–420. Available from: *https://doi.org/10.1177/1077558711399580*.

[86] Bähler C, Huber CA, Brüngger B, Reich O. Multimorbidity, health care utilization and costs in an elderly community-dwelling population: a claims data based observational study. BMC Health Services Research. 2015 Jan;15(1):23. Available from: *https://-doi.org/10.1186/s12913-015-0698-2*.

[87] Moffat K, Mercer SW. Challenges of managing people with multimorbidity in today's healthcare systems. BMC Family Practice. 2015 Oct;16(1):129. Available from: *https://doi.org/10.1186/s12875-015-0344-4*.

[88] He J, Morales DR, Guthrie B. Exclusion rates in randomized controlled trials of treatments for physical conditions: a systematic review. Trials. 2020 Feb;21(1):228. Available from: *https://doi.org/10.1186/s13063-020-4139-0*.

[89] Buffel du Vaure C, Dechartres A, Battin C, Ravaud P, Boutron I. Exclusion of patients with concomitant chronic conditions in ongoing randomised controlled trials targeting 10 common chronic conditions and registered at ClinicalTrials.gov: a systematic review of registration details. BMJ Open. 2016 Sep;6(9):e012265. Available from: *http://-dx.doi.org/10.1136/bmjopen-2016-012265*.

[90] Guthrie B, Payne K, Alderson P, McMurdo MET, Mercer SW. Adapting clinical guidelines to take account of multimorbidity. BMJ. 2012 Oct;345:e6341. Available from: *https://doi.org/10.1136/bmj.e6341*.

[91] Zulman DM, Asch SM, Martins SB, Kerr EA, Hoffman BB, Goldstein MK. Quality of Care for Patients with Multiple Chronic Conditions: The Role of Comorbidity Interrelatedness. Journal of General Internal Medicine. 2014 Mar;29(3):529–537. Available from: *https://doi.org/10.1007/s11606-013-2616-9*.

[92] Stange KC. In This Issue: Challenges of Managing Multimorbidity. Ann Fam Med. 2012 Jan;10(1):2. Available from: *https://doi.org/10.1370/afm.1354*.

[93] Farmer C, Fenu E, O'Flynn N, Guthrie B. Clinical assessment and management of multimorbidity: summary of NICE guidance. BMJ. 2016 Sep;354:i4843. Available from: *https://doi.org/10.1136/bmj.i4843*.

[94] Hughes LD, McMurdo MET, Guthrie B. Guidelines for people not for diseases: the challenges of applying UK clinical guidelines to people with multimorbidity. Age Ageing. 2013 Jan;42(1):62–69. Available from: *https://doi.org/10.1093/ageing/afs100*.

[95] Smith SM, Wallace E, O'Dowd T, Fortin M. Interventions for improving outcomes in patients with multimorbidity in primary care and community settings. Cochrane Database of Systematic Reviews. 2021;(1). Available from: *https://doi.org//10.1002/-14651858.CD006560.pub4*.

[96] Violan C, Foguet-Boreu Q, Flores-Mateo G, Salisbury C, Blom J, Freitag M, et al. Prevalence, Determinants and Patterns of Multimorbidity in Primary Care: A Systematic Review of Observational Studies. PLOS ONE. 2014 Jul;9(7):e102149. Available from: *https://doi.org/10.1371/journal.pone.0102149*.

[97] Arnett DK, Goodman RA, Halperin JL, Anderson JL, Parekh AK, Zoghbi WA. AHA/ACC/HHS Strategies to Enhance Application of Clinical Practice Guidelines in Patients With Cardiovascular Disease and Comorbid Conditions: From the American Heart Association, American College of Cardiology, and U.S. Department of Health and Human Services. Journal of the American College of Cardiology. 2014 Oct;64(17):1851–1856. Available from: *https://doi.org/10.1016/j.jacc.2014.07.012*.

[98] Tran J, Norton R, Conrad N, Rahimian F, Canoy D, Nazarzadeh M, et al. Patterns and temporal trends of comorbidity among adult patients with incident cardiovascular disease in the UK between 2000 and 2014: A population-based cohort study. PLoS Med. 2018 Mar;15(3):e1002513. Available from: *https://doi.org/10.1371/-journal.pmed.1002513*.

[99] Weiss CO, Boyd CM, Yu Q, Wolff JL, Leff B. Patterns of Prevalent Major Chronic Disease Among Older Adults in the United States. JAMA. 2007 Sep;298(10):1158–1162. Available from: *https://doi.org/10.1001/jama.298.10.1160-b*.

[100] Rocca WA, Boyd CM, Grossardt BR, Bobo WV, Finney Rutten LJ, Roger VL, et al. Prevalence of multimorbidity in a geographically defined American population: patterns by age, sex, and race/ethnicity. Mayo Clinic proceedings. 2014 Oct;89:1336–49. Available from: *https://doi.org/10.1016/j.mayocp.2014.07.010*.

[101] Prados-Torres A, Calderón-Larrañaga A, Hancco-Saavedra J, Poblador-Plou B, van den Akker M. Multimorbidity patterns: a systematic review. Journal of clinical epidemiology. 2014 Mar;67:254–66.

[102] Zemedikun DT, Gray LJ, Khunti K, Davies MJ, Dhalwani NN. Patterns of Multimorbidity in Middle-Aged and Older Adults: An Analysis of the UK Biobank Data. Mayo Clinic Proceedings. 2018 Jul;93(7):857–866. Available from: *https://doi.org/-10.1016/j.mayocp.2018.02.012*.

[103] Prados-Torres A, Poblador-Plou B, Calderón-Larrañaga A, Gimeno-Feliu LA, González-Rubio F, Poncel-Falcó A, et al. Multimorbidity Patterns in Primary Care: Interactions among Chronic Diseases Using Factor Analysis. PLOS ONE. 2012 Feb;7(2):e32190. Available from: *https://doi.org/10.1371/journal.pone.0032190*.

[104] Abad-Díez JM, Calderón-Larraáaga A, Poncel-Falcó A, Poblador-Plou B, Calderón-Meza JM, Sicras-Mainar A, et al. Age and gender differences in the prevalence and patterns of multimorbidity in the older population. BMC Geriatrics. 2014 Jun;14(1):75. Available from: *https://doi.org/10.1186/1471-2318-14-75*.

[105] Guisado-Clavero M, Roso-Llorach A, López-Jimenez T, Pons-Vigués M, Foguet-Boreu Q, Muñoz MA, et al. Multimorbidity patterns in the elderly: a prospective cohort

study with cluster analysis. BMC Geriatrics. 2018 Jan;18(1):16. Available from: *https://-doi.org/10.1186/s12877-018-0705-7*.

[106] van den Akker M, Buntinx F, Roos S, Knottnerus JA. Problems in determining occurrence rates of multimorbidity. Journal of Clinical Epidemiology. 2001 Jul;54(7):675–679. Available from: *https://doi.org/10.1016/S0895-4356(00)00358-9*.

[107] Wong A, Boshuizen HC, Schellevis FG, Kommer GJ, Polder JJ. Longitudinal administrative data can be used to examine multimorbidity, provided false discoveries are controlled for. Journal of Clinical Epidemiology. 2011 Oct;64(10):1109–1117. Available from: *https://doi.org/10.1016/j.jclinepi.2010.12.011*.

[108] Vetrano DL, Roso-Llorach A, Fernández S, Guisado-Clavero M, Violán C, Onder G, et al. Twelve-year clinical trajectories of multimorbidity in a population of older adults. Nature Communications. 2020 Jun;11(1):3223. Available from: *https://doi.org/10.1038/-s41467-020-16780-x*.

[109] Freisling H, Viallon V, Lennon H, Bagnardi V, Ricci C, Butterworth AS, et al. Lifestyle factors and risk of multimorbidity of cancer and cardiometabolic diseases: a multinational cohort study. BMC Medicine. 2020 Jan;18(1):5. Available from: *https://-doi.org/10.1186/s12916-019-1474-7*.

[110] Ashworth M, Durbaba S, Whitney D, Crompton J, Wright M, Dodhia H. Journey to multimorbidity: longitudinal analysis exploring cardiovascular risk factors and sociodemographic determinants in an urban setting. BMJ Open. 2019 Dec;9:e031649. Available from: *http://dx.doi.org/10.1136/bmjopen-2019-031649*.

[111] Haug N, Deischinger C, Gyimesi M, Kautzky-Willer A, Thurner S, Klimek P. High-risk multimorbidity patterns on the road to cardiovascular mortality. BMC Medicine. 2020 Mar;18(1):44. Available from: *https://doi.org/10.1186/s12916-020-1508-1*.

[112] Strauss VY, Jones PW, Kadam UT, Jordan KP. Distinct trajectories of multimorbidity in primary care were identified using latent class growth analysis. Journal of clinical epidemiology. 2014 Oct;67:1163–71. Available from: *https://doi.org/10.1016/-j.jclinepi.2014.06.003*.

[113] Kudesia P, Salimarouny B, Stanley M, Fortin M, Stewart M, Terry A, et al. The incidence of multimorbidity and patterns in accumulation of chronic conditions: A systematic review. Journal of Multimorbidity and Comorbidity. 2021 Jan;11:26335565211032880. Available from: *https://doi.org/10.1177/-26335565211032880*.

[114] Marian AJ, Bhatnagar A, Bolli R, Izpisua Belmonte JC. Introduction to Cardiovascular Aging Compendium. Circulation Research. 2018 Sep;123(7):737–739. Available from: *https://doi.org/10.1161/CIRCRESAHA.118.313940*.

[115] Djärv T, Wikman A, Lagergren P. Number and burden of cardiovascular diseases in relation to health-related quality of life in a cross-sectional population-based cohort study. BMJ Open. 2012 Jan;2(5):e001554. Available from: *http://dx.doi.org/10.1136/-bmjopen-2012-001554*.

[116] Qiu C, Fratiglioni L. A major role for cardiovascular burden in age-related cognitive decline. Nature Reviews Cardiology. 2015 Sep;12(5):267–277. Available from: *https://doi.org/10.1038/nrcardio.2014.223*.

[117] Zhu Y, Edwards D, Mant J, Payne RA, Kiddle S. Characteristics, service use and mortality of clusters of multimorbid patients in England: a population-based study. BMC Medicine. 2020 Apr;18(1):78. Available from: *https://doi.org/10.1186/s12916-020-01543-8*.

[118] Xu X, Mishra GD, Dobson AJ, Jones M. Progression of diabetes, heart disease, and stroke multimorbidity in middle-aged women: A 20-year cohort study. PLOS Medicine. 2018 Mar;15(3):e1002516. Available from: *https://doi.org/10.1371/-journal.pmed.1002516*.

[119] Calderón-Larrañaga A, Saadeh M, Hooshmand B, Refsum H, Smith AD, Marengoni A, et al. Association of Homocysteine, Methionine, and MTHFR 677C>T Polymorphism With Rate of Cardiovascular Multimorbidity Development in Older Adults in Sweden. JAMA Netw Open. 2020 May;3(5):e205316–e205316. Available from: *https://doi.org/10.1001/jamanetworkopen.2020.5316*.

2 Risk factors for cardiovascular multimorbidity: a systematic review of prospective studies

2.1 Introduction

2.1.1 Background

As outlined in Chapter 1, multimorbidity is increasing in global prevalence. [1] In the United Kingdom (UK), more than 60% of people aged 65 years and above were estimated to be classified as multimorbid in 2012. [2] Multimorbidity has been associated with higher mortality, lower quality of life, and greater use of health services and is a major public health concern in countries with ageing populations like the UK. [3]

A 2014 systematic review on the prevalence, patterns, and determinants of multimorbidity found that cardiovascular diseases (CVDs) were the most common components of multimorbidity. [4] Preventing the co-occurrence of two or more CVDs – or cardiovascular multimorbidity (CVM) as defined in this thesis – will likely reduce the overall multimorbidity burden, but the relationships between known cardiovascular risk factors and CVM are not well understood.

Most studies identified in the 2014 review were cross-sectional and had a small number of events, making them inadequate for studying risk factors for multimorbidity. [4] At the time of writing, no published systematic review has summarised the prospective evidence on risk factors for multimorbidity or has focussed on CVM.

2.1.2 Objectives

The objective of this chapter was to conduct a systematic review of all available prospective evidence on the associations of known cardiovascular risk factors with multimorbidity, with a focus on CVM. The following research questions were addressed:

- What is known about the direction and magnitude of the relationships between known cardiovascular risk factors and incident multimorbidity (particularly CVM) in any population or setting?

- What methods have been used in the literature to investigate the associations of known cardiovascular risk factors with incident multimorbidity, and particularly with CVM?

A secondary objective was to identify the limitations in the existing evidence on risk factors for CVM and their implications for the subsequent analyses in this thesis.

2.2 Methods

2.2.1 Registration of protocol

The protocol for this systematic review was registered on PROSPERO, the International Prospective Register of Systematic Reviews. It is available from:

https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021240491

2.2.2 Eligibility criteria

The full eligibility criteria for this systematic review are presented in Table 2.1. The outcome of interest was incident multimorbidity. Each study defined multimorbidity using a list of potential component diseases, and studies that did not include any CVDs were excluded from this review. Prospective observational studies conducted in adult human populations in any country or setting with at least 500 participants or more were included in this review. Studies that restricted the entire cohort to those with pre-specified medical conditions were excluded. The exposures of interest were established cardiovascular risk factors including socio-demographic, lifestyle, and clinical characteristics. Table 2.2 describes the exposures of interest in more detail. Studies that did not conduct formal statistical analyses on the relationships between risk factors and multimorbidity were also excluded.

Table 2.1 Eligibility criteria

	Inclusion criteria	**Exclusion criteria**
Language	English	Non-English
Sample size	500 or more	Under 500
Study type	Primary prospective observational studies (cohort, nested case-control, Mendelian randomisation, etc.) of human populations	Case report, case series, ecological study, cross-sectional study, validation study, qualitative study, editorial, letter, review, randomised controlled trial, lab-based experimental study, etc.
Demographic characteristics	Predominantly adults (aged 18 years or older), any sex	Predominantly children to adolescents
Other population characteristics		Entire study population is restricted to participants with pre-specified medical conditions
Outcome	Investigated outcome(s) include incident multimorbidity, which is defined using at least one cardiovascular disease	Study does not address multimorbidity as an outcome, or multimorbidity definition does not include any cardiovascular diseases
Exposure	Investigated exposure(s) are established cardiovascular risk factors including, but not limited to, socio-demographic characteristics, biochemical or physiological characteristics, reproductive factors, early life factors, lifestyle, psychosocial factors, genetic factors, and clinical factors in mental and physical health. (see Table 2.2 for examples)	Acute CVDs (such as coronary heart disease, stroke, etc.). Other exposures (e.g. surgical and medical interventions including pharmaceuticals and alternative or complementary medicines for self-care and disease management. Health services use and costs, general health status, quality of life measures, disease symptoms, general disability, other psychological factors, non-specific occupation exposures, etc.)
Methods		No formal statistical test conducted on the relationship between the exposure(s) and multimorbidity

*Table 2.2 Details and examples of exposures (i.e. established cardiovascular risk factors) of interest**

Category	Examples
Socio-demographic	age, sex, ethnicity, education, economic conditions, marital status
Biochemical or physiological	blood pressure, HDL cholesterol, LDL cholesterol, triglycerides, glycaemia, adiposity, thrombogenic factors, markers of chronic inflammation
Reproductive	use of oral contraceptives, menopausal status, parity
Early life	birth weight, having been breastfed, childhood social circumstances
Lifestyle	smoking/tobacco, diet, physical activity, alcohol
Psychosocial	chronic stress, social isolation
Genetic	family history of cardiovascular disease, genetic markers
Clinical	diabetes, obesity, cognitive impairment, anxiety, depression

*This table does not completely describe all possible exposures that might be included, but rather sets out the type of exposure that is considered to fall within the remit of the review; HDL=high-density lipoprotein; LDL=low-density lipoprotein

2.2.3 Information sources

A first search was conducted on 17 November 2017 (with no language or date restrictions) of the following databases: CINAHL (via EBSCOhost, 1981 to 16/11/17), EMBASE (via Ovid, 1974 to 16/11/17), MEDLINE (via Ovid, 1946 to 16/11/17), and Web of Science (via Thomson Reuters, 1945 to 16/11/17). The database search results included published studies as well as 'grey literature' including conference proceedings and dissertations. Reference lists of identified systematic reviews on multimorbidity [3, 4] and of the studies included in this systematic review were examined to find any additional studies. The screening process was independently conducted by me and Dr. Mahmuda Jahan Chadni, the second reviewer. This only yielded two studies that fully met the inclusion criteria. As a result, a second search of the same databases using the same search terms was conducted on 5 January 2021 to identify any relevant studies published since the initial search in 2017. For the second search, the screening procedure was independently conducted by me and Dr. Sunjuri Sun, who acted as the second reviewer for the search updates. The results presented in this chapter are therefore based on the search results of the four databases up to 5 January 2021.

2.2.4 Search strategy

The full search strategies used in each database are reported in Appendix Table A1-A4 (page 286-9). Nia Roberts – Outreach Librarian and Information Specialist at the Bodleian Health Care Libraries of the University of Oxford – provided expert guidance and feedback on the development of the search strategies. Search strings were formulated based on MeSH headings and keywords of previously identified studies on CVM, two of which were included in the final data synthesis. [5, 6] Relevant indexing terms (e.g. MeSH), text search terms, and their synonyms and permutations were included in the search strings to capture the concept of risk factors for multimorbidity of CVDs or chronic

diseases. The searches were limited to observational studies based on the Observational Studies Filter developed by the Scottish Intercollegiate Guidelines Network. [7] The filter was modified to include specific study types (e.g. population surveys) and articles indexed as epidemiological studies of CVDs or chronic diseases. The search strings were not restricted to prospective studies, because based on test searches, doing so decreased the sensitivity of the results. Therefore, prospective studies were not filtered at this stage, but in the full-text screening step instead.

2.2.5 Selection process

Duplicates of records identified through database searches were removed using Endnote X9. [8] Then, the eligibility criteria were applied to the titles and abstracts to exclude clearly irrelevant reports. The remaining reports were read in full to determine whether they should be included in the final synthesis according to the eligibility criteria. These procedures were independently carried out by two reviewers for each report. Any disagreements were adjudicated initially between the two reviewers (as described in Section 2.2.3, above) and if an agreement could not be reached, through further discussion with my supervisors. Covidence software was used for full-text screening and adjudication. [9]

2.2.6 Data items and collection process

Data extraction sheets were prepared in Covidence and Microsoft Excel to obtain information on study characteristics, methods, participant characteristics, data sources, and multimorbidity definition, incidence, and association with cardiovascular risk factors. Data collection was independently carried out by the two reviewers and discrepancies were adjudicated together in meetings. The full list of data extraction items is available in Appendix Table A5 (page 290).

2.2.7 Risk of bias assessment

Each paper included in the final data synthesis was independently assessed by the two reviewers for risk of bias and other quality measures using the Newcastle-Ottawa Quality Assessment Tool [10]. This tool was used to assess a range of study quality characteristics including the representativeness of the study sample, method of ascertainment of the exposure and outcome variables, adjustment for confounders, and adequacy of follow-up. The assessment tool is in Appendix Figure A1 (page 291).

2.2.8 Effect measures and synthesis methods

Due to the heterogeneity of the definitions of multimorbidity used in the identified studies/ reports, it was not possible to quantitatively synthesise the findings across studies (e.g. by formal meta-analysis). Instead, the qualitative direction of association was compared across studies when at least two included studies investigated the exposure. Adjusted measures of association (e.g. odds ratio [OR], hazard ratio [HR]) between cardiovascular risk factors and incident multimorbidity were reported for each study in Appendix Table A7 (page 293).

Three studies investigated cardiometabolic multimorbidity, defined as having at least two conditions from coronary heart disease, stroke, and diabetes. [5, 11, 12] Among the definitions of multimorbidity identified in this review, cardiometabolic multimorbidity was best aligned with CVM as defined in this thesis. Therefore, the main findings of these studies were reported separately and in more detail. Nevertheless, since cardiometabolic multimorbidity included diabetes – a known cardiovascular risk factor rather than a CVD – it was not considered to be a type of CVM.

2.3 Results

2.3.1 Study selection

Figure 2.1 displays the PRISMA flow diagram for this systematic review. CINAHL, Embase, Medline, and Web of Science returned a total of 11,353 records, of which 4,208 duplicate records were removed. No additional records were identified through manual searching of reference lists. A total of 7,145 titles and abstracts were screened, of which 6,902 were excluded for not meeting the eligibility criteria. In total, 243 reports were downloaded and screened in full text. At this stage, 128 non-prospective studies were excluded. A further 19 studies were excluded because multimorbidity was not studied as an outcome, and four studies were excluded since the definition of multimorbidity did not include any CVDs. In total, 10 studies were included and described in this review.

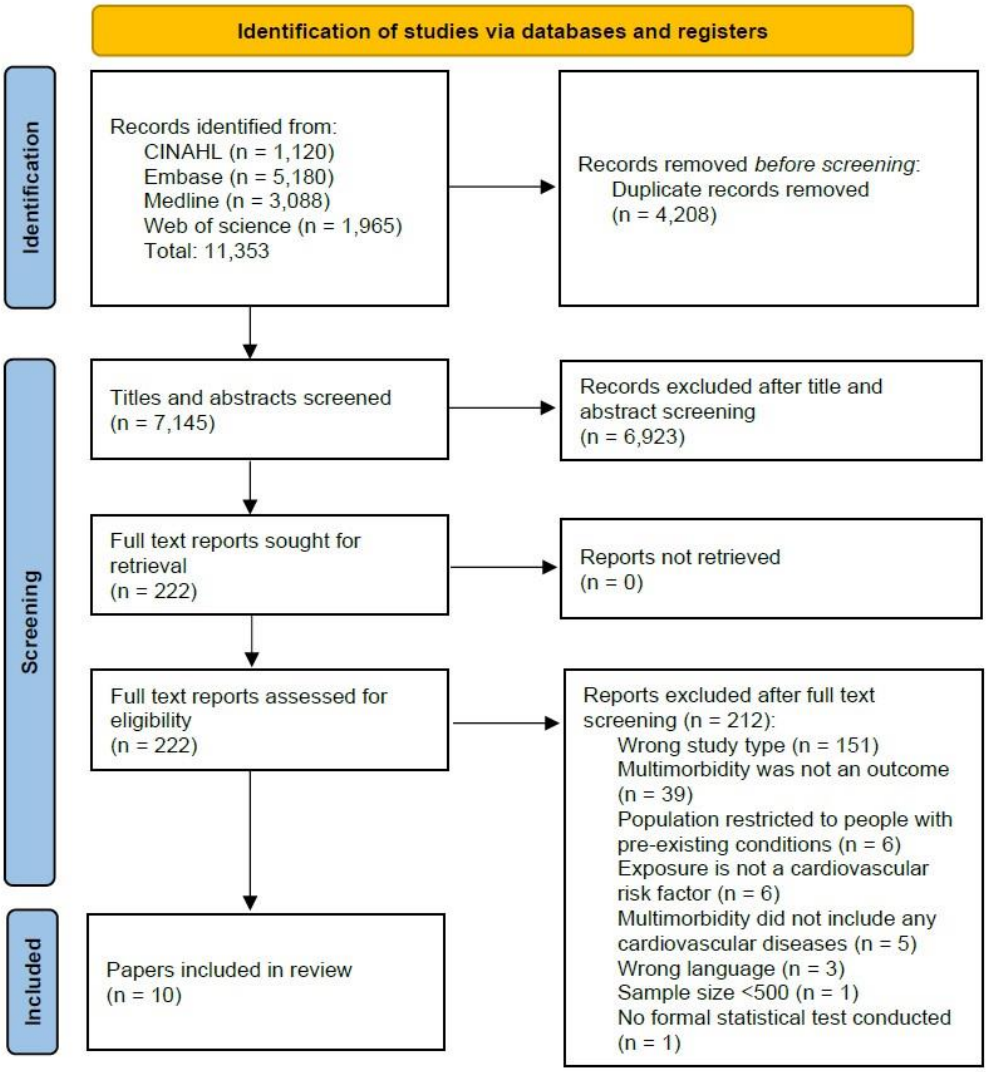


Figure 2.1 PRISMA flow diagram of database search results as of 5 January 2021

2.3.2 Study characteristics

Characteristics of the 10 included studies are presented in Table 2.3. All included studies were published in or after 2015 and were based in high-income countries. There was some overlap of study cohorts: two studies used data from the Whitehall II Cohort Study, two studies from the Australian Longitudinal Study on Women's Health (ALSWH), and one paper was a pooled analysis of 16 different prospective cohort studies, of which two (Whitehall II and the English longitudinal study on Ageing [ELSA]) overlapped with other studies in this review [5]. The median sample size was 10,105.5 (interquartile range [IQR] 6,912- 31,207). The maximum sample size was 120,813. Most studies reported a higher percentage of women than men, and the median proportion of females was 54.4% (IQR 50-70.6). The two studies using ALSWH data only studied women. [11, 13] In most studies, the mean age of participants was around 50 years, but a notable exception was Li et al. [14], in which 80% of participants were under the age of 50 years. Recruitment of participants often occurred in the 1990s and 2000s, and recruitment years across the studies ranged from 1973-2015.

Table 2.3 Characteristics of studies included in the systematic review

First author, Publication year	Country	Study name	Sample size	Proportion female, %	Age, years Range (mean)	Recruitment years	Median duration of follow-up, years	Incident MM (≥2 diseases), n	Selected diseases [b], n
Dugravot, 2020 [15]	UK	Whitehall II	6,425	28.8	50	1985-1988	23.6	1,694	9 (2)
Keats, 2020 [16]	Canada	PATH	6,912	70.6	35-69 (53.5[a])	2009-2015	n/a	2,638	5 (1)
Kivimäki, 2017 [5]	USA & Europe	16 prospective cohort studies	120,813	59.1	35-103 (51.4)	1973-2012	Mean 10.7	1,627	3 (2)
Li, 2021 [14]	Canada	Pooled national health surveys	29,838	51	≥18	1996-2003	10	3,151 (women), 2,598 (men)	15 (5)
Mounce, 2018 [17]	UK	ELSA	1,477	50	≥50	2002-2003	Mean 10.1	377	15 (4)
Shang, 2020 [18]	Australia	45 and Up	53,867	56.5	45-64	2006-2009	Mean 8.9	7,098 (women), 5,955 (men)	11 (2)
Singh-Manoux, 2018 [12]	UK	Whitehall II	8,270	33[a]	50	1985-1988	Mean 23.7	511	3 (2)
Wikstrom, 2015 [6]	Finland	FINRISK	31,207	52.2[a]	25-64	1982-2002	10	166 (women), 311 (men)	5 (1)
Xu, 2018 [11]	Australia	ALSWH	11,941	100	45-50	1996	20	423	3 (2)
Xu, 2019b [13]	Australia	ALSWH	7,357	100	45-50 (47.6[a])	1996	20	4,442	9 (1)

[a]Figure not reported but derived from relevant information in the study [b]Number of cardiovascular diseases is indicated in parentheses; n/a=not available; MM=multimorbidity; UK=United Kingdom; USA=United States of America; PATH=Atlantic Partnership for Tomorrow's Health; ELSA=The English Longitudinal Study of Ageing; ALSWH=Australian Longitudinal Study on Women's Health

2.3.3 Risk of bias in studies

The Newcastle-Ottawa scale (NOS) score of each study – quantifying study quality out of a maximum score of 8 – are presented in Table 2.4. The NOS scores ranged between 4 to 7 across the 10 included studies, with two scoring 4, three scoring 5, three scoring 6, and two with a score of 7. Most studies scored high on selection and comparability, but low on outcome. Only two studies [12, 15] were judged to have adequate completeness of follow-up (i.e. follow-up >90%, or follow-up <90% and description provided of those lost), and both were based in the Whitehall II cohort study. However, these studies were restricted to a sample of British civil servants and did not adjust for smoking, and overall were assessed to have a study quality NOS score of 5. Keats et al. [16] and Shang et al. [18] scored the lowest among the included studies (NOS score of 4) mainly due to poor reporting of methods. Keats et al. did not report the duration of follow-up and relied on non-validated self-report for both exposure and outcome ascertainment. Shang et al. did not report what statistical method was used to obtain the effect estimates or what the adjustment variables were.

Table 2.4 Quality assessment for each study using a modified version of the Newcastle-Ottawa Scale

AUTHOR, YEAR	SELECTION			COMPARABILITY		OUTCOME			TOTAL
	Representa-tiveness	Ascertainm-ent of exposure	Outcome not present at study baseline	Controls for age, sex, and smoking	Additional adjustment	Assessment of outcome	Adequacy of duration of follow-up	Adequacy of completeness of follow-up	/ 8
DUGRAVOT, 2020 [15]		*	*			*	*	*	5
KEATS, 2020 [16]	*		*	*	*				4
KIVIMÄKI, 2017 [5]	*	*	*	*		*	*		6
LI, 2021 [14]	*	*	*	*	*	*	*		7
MOUNCE, 2018 [17]	*	*	*	*	*		*		6
SHANG, 2020 [18]	*		*			*	*		4
SINGH-MANOUX, 2018 [12]		*	*			*	*	*	5
WIKSTROM, 2015 [6]	*	*	*	*	*	*	*		7
XU, 2018 [11]	*		*	*	*	*	*		6
XU, 2019B [13]	*		*	*	*		*		5

Study quality was assessed based on a modified version of the Newcastle-Ottawa scale (Appendix Figure A1). Each study could be awarded a maximum of one star (*) per each column in the selection, comparability, and outcome categories. A higher number of total stars indicates higher study quality. A maximum of eight stars could be awarded in total. **Representativeness**: study was given a star if the study cohort was either truly representative of the target population or comprised a large, diverse sample somewhat representative of the average in the target population. **Ascertainment of exposure:** star given if exposure data was collected via secure records, structured interview, direct measurement, or validated self-report. **Additional adjustment:** star given if additionally controls for one or more of body mass index, hypertension, hypercholesterolemia, diabetes, or socioeconomic status. **Assessment of outcome:** star given if outcome was assessed via independent blind assessment, record linkage, or validated self-report. **Adequacy of duration of follow-up:** star given if follow-up period was at least five years. **Adequacy of follow-up:** star given if follow-up is complete or if subjects lost to follow-up are unlikely to introduce bias (follow-up rate over 90% or description provided of those lost).

2.3.4 Methods used to investigate risk factors for multimorbidity

2.3.4.1 Definition of multimorbidity

Table 2.5 summarises the definition of multimorbidity used in each study. Each study provided a list of diseases that could be counted towards multimorbidity. The median number of selected diseases was 7, ranging from 3 to 15. All studies but Shang et al. [18] explained the reasons for selecting the set of diseases. The majority of studies selected the diseases according to definitions of multimorbidity that other studies had used [5, 11, 12, 14, 17] or based on official guidelines [13]. Other studies selected the diseases due to high prevalence [6, 16] or as leading causes of death in the target population [15].

No study included in this review defined multimorbidity only using CVDs: all studies included diabetes, a known cardiovascular risk factor. Three studies defined multimorbidity as two or more of coronary heart disease, stroke, or diabetes (i.e. cardiometabolic multimorbidity). [5, 11, 12] Four other studies also included coronary heart disease, stroke, and diabetes, but in addition to other non-cardiometabolic diseases. [14, 15, 17, 18] Two studies included heart failure and cardiac arrhythmias. [14, 17] No other individual CVDs were included in any of the ten studies. Two studies included a composite CVD outcome comprising coronary heart disease and/or stroke. [6, 16] One study included "cardiovascular disease" with no further explanation. [13]

Most included studies defined multimorbidity as the co-occurrence of two or more chronic diseases. Shang et al. [18] was the exception, which investigated having ≥2, ≥3, or ≥4 diseases as three separate definitions of multimorbidity. Kivimäki et al. [5] and Xu et al. (2018) [11] additionally reported risk factor associations with cardiometabolic multimorbidity subdivided according to the temporal order of occurrence of the component diseases (i.e. vascular disease followed by diabetes or diabetes followed by vascular disease).

Table 2.5 Definition of multimorbidity in each study

Author, year	Disease selection criteria	Number of selected diseases*, n	List of selected cardiovascular diseases for each study
Dugravot, 2020 [15]	Leading causes of death in high-income countries	9 (2)	coronary heart disease, stroke
Keats, 2020 [16]	Based on high prevalence in Canada and the Atlantic provinces	5 (1)	cardiovascular disease (myocardial infarction and/or stroke)
Kivimäki, 2017 [5]	Based on previous work of the Emerging Risk Factors Collaboration	3 (2)	coronary heart disease, stroke
Li, 2021 [14]	Based on previous Canadian studies and recommended by a systematic review of multimorbidity studies	15 (5)	acute myocardial infarction, cardiac arrhythmia, chronic coronary syndrome, congestive heart failure, stroke
Mounce, 2018 [17]	Measured in ELSA and typically included by previous studies	15 (4)	ischemic heart disease (angina or myocardial infarction), congestive heart failure, heart arrhythmia, stroke
Shang, 2020 [18]	Not specified	11 (2)	heart disease, stroke
Singh-Manoux, 2018 [12]	Based on previous work of the Emerging Risk Factors Collaboration	3 (2)	coronary heart disease, stroke
Wikstrom, 2015 [6]	Most common chronic diseases within the study population	5 (1)	cardiovascular disease (coronary heart disease and/or stroke)
Xu, 2018 [11]	Based on previous evidence	3 (2)	heart disease (angina, heart attack), stroke
Xu, 2019b [13]	Based on a previous systematic review [19] and identified as the National Health Priority Areas in Australia	9 (1)	cardiovascular disease (unspecified)

*Number of cardiovascular diseases analysed in the study is indicated in parentheses

2.3.4.2 Ascertainment of exposure and outcome

Appendix Table A6 (page 292) summarises the methods used to ascertain exposure and outcome data, and the statistical methods used to investigate the association between them. For outcome measurement, four studies [11, 13, 16, 17] used self-report only and four [6, 12, 14, 18] used linked health records only, while two [5, 15] used a combination of both. For exposure measurement, although different data sources were often used for different exposures in the same paper, four studies [11, 13, 14, 16] used self-report only, two [15, 18] used self-report and data linkage, and four [5, 6, 12, 17] used self-report and direct measure.

2.3.4.3 Statistical methods

All studies used either proportional hazards regression or logistic regression to study the association between risk factors and multimorbidity.

Two studies went beyond only investigating the exposure-outcome association to additionally assessing the role of risk factors in the trajectory from healthy state to multimorbidity and then to mortality [12, 15], by using multistate models. This method uses time-to-event data to estimate the effect of risk factors on the movement of individuals through multiple possible states (events are the transitions between states). [20, 21] Dugravot et al. [15] considered three possible states (healthy, multimorbidity, and mortality) and investigated three types of transitions (healthy to multimorbidity, healthy to mortality, and multimorbidity to mortality). They found that lower education was associated with a higher risk of transition from healthy state to multimorbidity, but that there was no association between educational attainment and the transitions from healthy to mortality or from multimorbidity to mortality. In the study by Singh-Manoux et al. [12] hypertension, obesity, and high blood cholesterol were associated with the transition from

healthy to first disease but not with the transition from first disease to cardiometabolic multimorbidity.

2.3.5 Risk factors for overall multimorbidity

Table 2.6 displays the direction of association between selected cardiovascular risk factors and multimorbidity reported by the 10 included studies. The selected risk factors were BMI, physical activity, smoking, alcohol consumption, hypertension, high cholesterol, education, and depression, which were chosen because their relationship with multimorbidity was investigated by at least two studies. The fully adjusted effect estimates and adjustment variables in each study are presented in Appendix Table A7 (page 293). The complete list of diseases included in the multimorbidity definition and exposures investigated in each study can be found in Appendix Table A8 (page 296).

BMI was the most frequently studied risk factor. [5, 6, 11-14, 17, 18] All eight studies that investigated BMI found a statistically significant association between higher BMI and higher risk of multimorbidity, despite the differences in cohort characteristics and the selection of diseases to define multimorbidity.

Physical activity was investigated in six studies [6, 11, 12, 14, 16, 17]: three found that higher levels of physical activity were associated with a lower risk of multimorbidity (i.e. inverse association) [6, 16, 17], while the other three found no statistically significant association [11, 12, 14]. Two of the studies reported separate effect estimates for men and women – one [14] found no association in both sexes while the other [6] found an inverse association in both sexes. The largest effect estimate was found in Xu et al. (2018) [11], where low levels of physical activity (0-39 MET min/week) was associated with a higher risk of multimorbidity compared to high levels of physical activity (≥1200 MET min/week): odds ratio 1.38, 95% CI 1.02-1.86.

Five studies [6, 11, 12, 14, 17] examined smoking as a risk factor for multimorbidity: three found smoking to be associated with a higher risk of multimorbidity (i.e. positive association) [6, 11, 12], one reported no statistically significant association [17], and one reported no statistically significant association in women but found a positive association in men [14]. The effect estimates of the largest magnitude were found in Wikstrom et al. [6], which reported separate HRs for men (HR 2.68, 2.10-3.41) and women (HR 2.55, 1.76-3.71) in current smokers compared to never smokers.

Four studies [6, 11, 15, 17] investigated the relationship between education and incident multimorbidity. Lower education was associated with a higher risk of multimorbidity in one study [15] and no association in two studies [11, 17]. One study found that lower education was associated with a higher risk of multimorbidity in men but found no association in women. [6]

All other risk factors were investigated by only two or three studies per risk factor. Alcohol consumption was investigated as a risk factor in two studies [12, 14], but only one found a statistically significant association with multimorbidity (HR for abstention or heavy drinking [vs. moderate drinking]: 1.30, 1.09-1.56) [12]. Hypertension was associated with a higher risk of multimorbidity in two studies [11, 12], while one study reported evidence that hypertension increased the risk of multimorbidity in women but no such evidence for men [6]. High cholesterol did not have a statistically significant association with multimorbidity in both studies that investigated it. [6, 12] Depression was not associated with multimorbidity in one study [11], while in one other study, depression was associated with a higher risk of multimorbidity in women but not in men [14].

Table 2.6 Direction of association between cardiovascular risk factors and multimorbidity

Author, year	Adjustment	High BMI	Low physical activity	Current smoking	Alcohol consumption	Hyper-tension	High choles-terol	Low education	Depression
Dugravot, 2020 [15]								↑	
Keats, 2020 [16]			↑						
Kivimäki, 2017 [5]	✓	↑							
Li, 2021 [14] (women)	✓	↑	-	-	-				↑
Li, 2021 [14] (men)	✓	↑	-	↑	-				-
Mounce, 2018 [17]	✓	↑	↑	-				-	
Shang, 2020 [18]		↑							
Singh-Manoux, 2018 [12]		↑	-	↑	↑	↑	-		
Wikstrom, 2015 [6] (women)	✓	↑	↑	↑		-	-	-	
Wikstrom, 2015 [6] (men)	✓	↑	↑	↑		↑	-	↑	
Xu, 2018 [11]	✓	↑	-	↑		↑		-	-
Xu, 2019b [13]	✓	↑							

Based on fully adjusted effect estimates; Adjustment = at minimum adjusted for age, sex, BMI, and smoking); BMI=body mass index; ↑=associated with a higher risk of multimorbidity; - =no association with multimorbidity, no marker indicates that the risk factor was not investigated by the study

2.3.6 Risk factors for cardiometabolic multimorbidity

Cardiometabolic multimorbidity – defined as having at least two conditions from coronary heart disease, stroke, and diabetes – was the type of multimorbidity that was most closely aligned with the CVM definition for this review, and as such is explored in more detail in this section. The main findings of the studies on cardiometabolic multimorbidity are summarised in Table 2.7.

BMI was studied as a risk factor in all three studies. [5, 11, 12] Kivimäki et al. [5] found that the risk of cardiometabolic multimorbidity was twice as high in overweight individuals (BMI 25.0-29.0 kg/m^2), five times as high in those with class I obesity (30.0-34.9 kg/m^2), and 15 times as high in those with class II obesity or above (≥35 kg/m^2) compared with those of healthy weight (20.0-24.9 kg/m^2), after adjusting for socio-demographic and lifestyle behavioural factors. As a pooled analysis of 16 cohort studies, this was by far the largest study among all included studies (cohort size of 120,813) and had the highest number of cases of cardiometabolic multimorbidity (1,627 incident cases).

Singh-Manoux et al. [12] found in a cohort of 8,270 British civil servants with 511 incident cases of cardiometabolic multimorbidity that being overweight or obese (BMI ≥25 kg/m^2) was associated with a 79% higher risk of cardiometabolic multimorbidity (HR 1.79, 95% CI 1.50-2.14) compared to those with a lower BMI (<25 kg/m^2). They did not, however, adjust for any lifestyle behavioural factors such as smoking or physical activity. Xu et al. (2018) [11] found in a cohort of 8,224 Australian women with 423 cases of incident cardiometabolic multimorbidity that compared to normal weight (18.5-24.9 kg/m^2), the OR associated with overweight (BMI 25-29.9 kg/m^2) was 1.75 (1.32-2.32) and with obesity (≥30 kg/m^2) was 3.66 (2.76-4.86), after adjusting for sociodemographic and lifestyle behavioural factors and other prior chronic conditions.

Singh-Manoux et al. and Xu et al. both also investigated hypertension and smoking, among other risk factors. Singh-Manoux et al. found that hypertension (defined as measured blood pressure of ≥140/90 mmHg or use of antihypertensive medication) was associated with a higher risk of cardiometabolic multimorbidity (HR 1.49, 95% CI 1.23-1.80). Xu et al. found that self-reported diagnosis of hypertension was associated with almost twice the risk of cardiometabolic multimorbidity (OR 1.94, 95% CI 1.43-2.64). Current smoking compared to never smoking was associated with about twice the risk of cardiometabolic multimorbidity in Singh-Manoux et al. (HR 2.01, 95% CI 1.64-2.47) and in Xu et al. (OR 2.00, 95% CI 1.51-2.65).

Table 2.7 Main results from studies specifically investigating cardiometabolic multimorbidity

Author, year	Cases / cohort size (location)	List of selected diseases	Exposure variables; adjustments	Main findings (fully adjusted)
Kivimäki et al., 2017 [5]	1,627 / 120,813 (USA & Europe)	diabetes, coronary heart disease, stroke	body mass index (BMI); age, sex, ethnic origin, lifestyle factors (smoking, physical activity, alcohol consumption)	Odds ratios (ORs) in participants with BMI of <20, 20-24.9, 25-29.9, 30-34.9, and ≥35 kg/m^2 were 0.9 (0.6-1.5), 1.0 (reference), 2.0 (1.7-2.4), 4.5 (3.5-5.8), and 14.5 (10.1-21.0), respectively
Singh-Manoux et al., 2018 [12]	511 / 8,270 (UK)	diabetes, coronary heart disease, stroke	clinical profile (hypertension, high cholesterol, overweight/obesity, family history of cardiometabolic disease) and behavioural factors (smoking, alcohol consumption, diet, physical activity); age, sex, race, marital status, and birth cohort	Hazard ratios for hypertension 1.49 (1.23-1.80), high cholesterol 1.32 (0.98-1.77), overweight 1.79 (1.50-2.14), family history of cardiometabolic disease 1.07 (0.86-1.34), current smoking 2.01 (1.64-2.47), alcohol consumption 1.34 (1.13-1.61), poor diet 1.20 (1.01-1.43), physical inactivity 1.21 (1.01-1.45)
Xu et al., 2018 [11]	423 / 11,941 (Australia)	diabetes, coronary heart disease, stroke	sociodemographic factors, lifestyle factors, and other prior chronic conditions; age, marital status, area of residence, education, country of birth, ability to manage on income, BMI, hypertension, physical activity, smoking status, other chronic conditions (depression/anxiety, COPD, asthma, cancer, arthritis, osteoporosis)	ORs for no educational qualifications 1.48 (0.99-2.20) [vs. university/higher degree], obesity 3.66 (2.76-4.86) [vs. normal weight], hypertension 1.94 (1.43-2.64) [vs. no hypertension], physical inactivity 1.31 (0.91-1.88) [vs. high activity], current smoking 2.00 (1.51-2.65) [vs. never smoking], depression/anxiety 1.18 (0.86-1.62) [vs. no other condition]

2.4 Discussion

2.4.1 Key findings in context

2.4.1.1 Key findings

This was the first systematic review to summarise the prospective evidence on risk factors for multimorbidity involving CVDs specifically. Although none of the included studies examined co-occurring CVDs only (i.e. CVM), three studies investigated cardiometabolic multimorbidity, which was the most closely related to CVM. Higher BMI was statistically significantly associated with higher cardiometabolic multimorbidity risk in all three studies. Hypertension and smoking were also associated with a higher risk, which were investigated in two of these studies.

Coronary heart disease and stroke were included in the definition of multimorbidity in all ten included studies, either as individual diseases or as part of a composite CVD outcome. Similar to cardiometabolic multimorbidity, higher BMI was associated with a higher risk of overall multimorbidity, while the associations with physical activity and education were inconsistent across studies. There was very limited information on any other risk factors: most were studied in fewer than four studies, and the evidence was conflicting.

Study quality was generally high, however the number of incident cases of multimorbidity was fewer than 5,000 in all but two studies and fewer than 1,000 in four studies. In eight out of ten studies, more than 10% of study participants were lost to follow-up and no description of those lost was provided. The only two studies with adequate completeness of follow-up were both based in the Whitehall II Cohort Study of British civil servants.

2.4.1.2 Comparison with previous evidence

The only previous systematic review on risk factors for multimorbidity was Violan et al. [4], which was published in 2014. The studies identified by Violan et al. were restricted to

the primary care setting, predominantly cross-sectional in design, and the only known cardiovascular risk factors investigated were age, sex, and socioeconomic status. Violan et al. found older age and female sex to be largely consistently associated with a higher prevalence of multimorbidity. Age is a well-established risk factor in chronic disease epidemiology in general. [22, 23] Only one study (Mounce et al. [17]) included in the review in this chapter investigated age and sex, which found that older age was associated with a higher risk of multimorbidity. Additionally, they reported that female sex was not statistically significantly associated with multimorbidity after multivariable adjustment. However, this may have been partly due to study size; Mounce et al. [17] had the smallest number of incident cases among all included studies in this chapter (n=377), and therefore did not have adequate statistical power to reliably detect moderate differences in risk.

In contrast to the consistent inverse association between socioeconomic status and multimorbidity reported by Violan et al., the review in this chapter found inconsistent relationships between education and multimorbidity across four studies. Two studies [11, 17] found no association between education and multimorbidity, one study [15] found an inverse association, and one study [6] found different results by sex. This was possibly because the studies in Violan et al. did not adjust for potential mediators/confounders such as smoking behaviour and BMI, whereas three [6, 11, 17] of the four prospective studies included in this chapter did. These three studies [6, 11, 17] also each had fewer than 500 cases of incident multimorbidity. The study in this chapter that found an inverse association between education and multimorbidity [15] (consistent with the findings of Violan et al.) did not adjust for covariates other than age, sex, ethnicity, marital status, and birth cohort, and had the highest number of cases of incident multimorbidity among the four studies that examined education as a risk factor (n=1,694).

2.4.2 Definition of multimorbidity

The lack of a standard definition of multimorbidity to facilitate comparisons across studies has been highlighted as a challenge in multimorbidity research by previous systematic reviews. [24, 25] Heterogeneity in the list of diseases to define multimorbidity precluded quantitative comparisons or syntheses of effect estimates in this chapter. Moreover, details of the reasoning and processes of selecting diseases to define multimorbidity were not reported in most of the included studies.

Although previous systematic reviews have recommended existing indices [26] or a list of core conditions [27] to define multimorbidity for some high-level research objectives, the needs of and information available for each study were unique. It has also been suggested that including as many diseases as possible rather than restricting to an arbitrary list could enhance the comparability of estimates. [4] However, this is likely to lead to a large majority of study participants being classified as multimorbid, making it difficult to derive meaningful insights on risk factor associations with multimorbidity. Including a large number of heterogeneous diseases in the definition of multimorbidity can also result in a high number of multimorbidity cases that are not necessarily similar to each other in their association with risk factors, complicating the interpretation of findings. For example, Xu et al. (2018) and Xu et al. (2019b) were studies included in this systematic review using data from the same cohort study, but the number of incident cases of multimorbidity was 10.5 times higher in Xu et al. (2019b), which defined multimorbidity using 9 diseases (n=4,442), compared to in Xu et al. (2018), which defined multimorbidity using 3 diseases (n=423). Xu et al. (2019b) included many common conditions such as hypertension and osteoporosis as well as more serious diseases like cancer. Within this study, someone with asthma and osteoporosis was considered equally multimorbid as someone with cancer and

CVD, although the risk factor associations of the two types of multimorbidity are likely to be dissimilar.

Therefore, when investigating risk factor associations with multimorbidity, defining multimorbidity using a limited number of well-defined conditions with shared aetiology is likely to make the results interpretable and comparable across different studies. Indeed, the only comparable studies in this review were those on cardiometabolic multimorbidity, defined using the same three conditions (coronary heart disease, stroke, and diabetes). [5, 11, 12] Alternatively, including an extensive list of chronic diseases in the definition of multimorbidity while also investigating more defined and comparable forms of multimorbidity (such as disease pairs or selected clusters) could also enhance the interpretability and comparability of results across studies. Some previous descriptive studies characterising multimorbidity have taken this approach. [28-31]

Regarding the choice of CVDs, individual CVDs included in multimorbidity definitions were at most limited to coronary heart disease, stroke, heart failure, and cardiac arrhythmia in the studies identified in this review. None of the studies included other common CVDs such as heart valve disorders or venous thromboembolism, or examined CVD subtypes, apart from myocardial infarction and angina as subtypes of coronary heart disease. Future studies on multimorbidity – and particularly CVM – could include a wider selection of CVDs and CVD subtypes (e.g. atrial fibrillation as a subtype of cardiac arrhythmia) for a more comprehensive picture of co-occurring CVDs.

2.4.3 Strengths and limitations

This was the first systematic review to focus specifically on and summarise the relationship between cardiovascular risk factors and multimorbidity involving CVDs. A strength of this review was the extensive database searches: a broad range of databases

were searched, and the search strings were tested against a selection of target studies before being formally implemented. A manual search of the reference lists of other systematic reviews and the included studies did not return any studies which had not already been captured by the primary searches.

However, there was large variation in the methods for defining multimorbidity (e.g. eligible conditions, data sources) in the included studies, and due to methodological heterogeneity and variable population characteristics, quantitative syntheses of effect estimates (e.g. by meta-analysis) would have been inappropriate. Publication bias may also be an issue if results of published and unpublished studies are systematically different from each other, for example due to studies with statistically significant results being more likely to be published. Although the potential for publication bias was not formally tested, searching grey literature did not yield any unpublished studies which were relevant to this systematic review.

2.4.4 Implications for future research

This systematic review found that there is a lack of prospective evidence on risk factors for multimorbidity with CVDs in the scientific literature, and almost none on CVM. The most relevant evidence was from three studies that specifically investigated cardiometabolic multimorbidity, which included two CVDs, but also diabetes in the definition. [5, 11, 12] The only prospective study which included a wide range of CVDs (and did not include any cardiovascular risk factors like diabetes) did not investigate any established cardiovascular risk factors and was therefore not eligible for inclusion in the review – the only risk factors investigated were total serum homocysteine and total serum methionine. [32] There is a need for good-quality, large-scale prospective evidence on risk factor associations with CVM defined using a comprehensive selection of CVDs.

It is important for future studies to report their methods clearly and in detail. Most studies included in this review had substantial loss to follow-up but did not report whether study participants who were lost to follow-up were different from those who remained under observation until the end of follow-up. Many studies also did not provide sufficient description of why or how conditions included in the definition of multimorbidity were selected. Moreover, there was no prospective evidence based in populations in low and middle-income countries, and future studies must address these gaps and limitations in the scientific literature.

2.4.5 Implications for this thesis

There is a need for prospective studies with adequate follow-up to examine various cardiovascular risk factors for CVM. The Million Women Study has a uniquely large cohort with a high number of individual CVDs, which makes it likely to have sufficient cases to allow for the exploration of various forms of CVM. The Million Women Study also has information on a variety of cardiovascular risk factors, including BMI, physical activity, smoking behaviour, and alcohol consumption, among others, which can be assessed as risk factors as well as adjusted for as potential confounders.

In this review, three articles (based on two cohort studies) were based in the UK [12, 15, 17], and of those articles, one in particular on cardiometabolic multimorbidity. [12] Compared to the Million Women Study, these UK studies had a smaller sample size and did not have as wide a range of information relevant to women's cardiovascular health. The sample sizes of the three UK studies were under 10,000, with the numbers of incident cases of multimorbidity ranging from 377 to 1,694; therefore, they were unlikely to have adequate statistical power to reliably assess risk factor associations with more detailed components of multimorbidity such as pairs of CVDs. The study with the largest number of cases of incident multimorbidity (n=1,694) did not have information on variables such

as BMI, smoking, or physical activity. [15] Moreover, two [12, 15] of the three study cohorts were approximately 70% male, while one was 50% female [17]. Analyses in the Million Women Study will be able to address these limitations of previous UK studies on risk factors for multimorbidity.

2.4.6 Conclusions

Prospective studies on risk factors for multimorbidity were heterogeneous in many aspects, particularly regarding the multimorbidity definition. High BMI and current smoking were relatively consistently associated with a higher risk of multimorbidity across the included studies. Overall, there was a lack of large-scale, high-quality prospective studies that assess cardiovascular risk factors for multimorbidity, and no studies specific to CVM. None of the studies had cohort sizes comparable to that of the Million Women Study (approximately 1.2 million participants). Analyses of risk factors for CVM in the Million Women Study in the subsequent chapters of this thesis will offer a unique contribution to the scientific literature on multimorbidity due to the very large number of events, information on many potential risk factors, substantially complete and documented follow-up, and the broad range of CVDs selected to define CVM. The findings of this thesis will be particularly important for women's cardiovascular health in the UK.

2.5 References

[1] Uijen AA, van de Lisdonk EH. Multimorbidity in primary care: prevalence and trend over the last 20 years. The European journal of general practice. 2008;14 Suppl 1:28–32. Available from: *https://doi.org/10.1080/13814780802436093*.

[2] Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet. 2012 Jul;380(9836):37–43. Available from: *https://doi.org/-10.1016/S0140-6736(12)60240-2*.

[3] Marengoni A, Angleman S, Melis R, Mangialasche F, Karp A, Garmen A, et al. Aging with multimorbidity: a systematic review of the literature. Ageing research reviews. 2011 Sep;10:430–9. Available from: *https://doi.org/10.1016/j.arr.2011.03.003*.

[4] Violan C, Foguet-Boreu Q, Flores-Mateo G, Salisbury C, Blom J, Freitag M, et al. Prevalence, Determinants and Patterns of Multimorbidity in Primary Care: A Systematic Review of Observational Studies. PLOS ONE. 2014 Jul;9(7):e102149. Available from: *https://doi.org/10.1371/journal.pone.0102149*.

[5] Kivimäki M, Kuosma E, Ferrie JE, Luukkonen R, Nyberg ST, Alfredsson L, et al. Overweight, obesity, and risk of cardiometabolic multimorbidity: pooled analysis of individual-level data for 120813 adults from 16 cohort studies from the USA and Europe. Lancet Public Health. 2017 Jun;2:e277–e285. Available from: *https://doi.org/10.1016/-s2468-2667(17)30074-9*.

[6] Wikström K, Lindström J, Harald K, Peltonen M, Laatikainen T. Clinical and lifestyle-related risk factors for incident multimorbidity: 10-year follow-up of Finnish population-based cohorts 1982-2012. European journal of internal medicine. 2015 Apr;26:211–6. Available from: *https://doi.org/10.1016/j.ejim.2015.02.012*.

[7] Scottish Intercollegiate Guidelines Network. Search filters [Internet]. Edinburgh, UK: Healthcare Improvement Scotland;. [cited 7 February 2022]. Available from: *https://-www.sign.ac.uk/what-we-do/methodology/search-filters/*.

[8] The EndNote Team, Clarivate Analytics. EndNote X9 64 bit. Philadelphia, PA; 2013.

[9] Veritas Health Innovation. Covidence systematic review software. Melbourne, Austrailia;. Available at www.covidence.org.

[10] Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al.. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses [Internet];. [cited 2022 Jan 07]. Available from: *http://www.ohri.ca/-programs/clinical_epidemiology/oxford.asp*.

[11] Xu X, Mishra GD, Dobson AJ, Jones M. Progression of diabetes, heart disease, and stroke multimorbidity in middle-aged women: A 20-year cohort study. PLOS Medicine. 2018 Mar;15(3):e1002516. Available from: *https://doi.org/10.1371/-journal.pmed.1002516*.

[12] Singh-Manoux A, Fayosse A, Sabia S, Tabak A, Shipley M, Dugravot A, et al. Clinical, socioeconomic, and behavioural factors at age 50 years and risk of

cardiometabolic multimorbidity and mortality: A cohort study. PLoS Med. 2018 May;15(5):e1002571. Available from: *https://doi.org/10.1371/journal.pmed.1002571*.

[13] Xu X, Mishra GD, Dobson AJ, Jones M. Short-term weight gain is associated with accumulation of multimorbidity in mid-aged women: a 20-year cohort study. International journal of obesity (2005). 2019 Sep;43:1811–1821. Available from: *https://doi.org/-10.1038/s41366-018-0250-7*.

[14] Li A, Rosella LC, Kurdyak P, Wodchis WP. Depression as a Risk Factor for Physical Illness and Multimorbidity in a Cohort with No Prior Comorbidity. Canadian journal of psychiatry Revue canadienne de psychiatrie. 2021 Aug;66:726–736. Available from: *https://doi.org/10.1177/0706743720974832*.

[15] Dugravot A, Fayosse A, Dumurgier J, Bouillon K, Rayana TB, Schnitzler A, et al. Social inequalities in multimorbidity, frailty, disability, and transitions to mortality: a 24-year follow-up of the Whitehall II cohort study. The Lancet Public Health. 2020 Jan;5(1):e42–e50. Available from: *https://doi.org/10.1016/S2468-2667(19)30226-9*.

[16] Keats MR, Cui Y, DeClercq V, Grandy SA, Sweeney E, Dummer TJB. Associations between Neighborhood Walkability, Physical Activity, and Chronic Disease in Nova Scotian Adults: An Atlantic PATH Cohort Study. International journal of environmental research and public health. 2020 Nov;17(33233809):8643. Available from: *https://doi.org/10.3390/ijerph17228643*.

[17] Mounce LTA, Campbell JL, Henley WE, Tejerina Arreal MC, Porter I, Valderas JM. Predicting Incident Multimorbidity. Annals of family medicine. 2018 Jul;16:322–329. Available from: *https://doi.org/10.1370/afm.2271*.

[18] Shang X, Peng W, Wu J, He M, Zhang L. Leading determinants for multimorbidity in middle-aged Australian men and women: A nine-year follow-up cohort study. Preventive medicine. 2020 Dec;141:106260. Available from: *https://doi.org/10.1016/-j.ypmed.2020.106260*.

[19] Xu X, Mishra GD, Jones M. Evidence on multimorbidity from definition to intervention: An overview of systematic reviews. Ageing research reviews. 2017 Aug;37:53–68. Available from: *https://doi.org/10.1016/j.arr.2017.05.003*.

[20] Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Statist Med. 2007 May;26(11):2389–2430. Available from: *https://-doi.org/10.1002/sim.2712*.

[21] Gillaizeau F, Sénage T, Le Borgne F, Le Tourneau T, Roussel JC, Leffondrè K, et al. Inverse probability weighting to control confounding in an illness-death model for interval-censored data. Statistics in Medicine. 2018 Apr;37(8):1245–1258. Available from: *https://doi.org/10.1002/sim.7550*.

[22] Christensen K, Doblhammer G, Rau R, Vaupel JW. Ageing populations: the challenges ahead. Lancet (London, England). 2009 Oct;374:1196–208. Available from: *https://doi.org/10.1016/S0140-6736(09)61460-4*.

[23] Fontana L, Kennedy BK, Longo VD, Seals D, Melov S. Medical research: Treat ageing. Nature. 2014 Jul;511(7510):405–407. Available from: *https://doi.org/10.1038/-511405a*.

[24] Holzer BM, Siebenhuener K, Bopp M, Minder CE. Evidence-based design recommendations for prevalence studies on multimorbidity: improving comparability of estimates. Population health metrics. 2017 Mar;15:9. Available from: *https://doi.org/-10.1186/s12963-017-0126-4*.

[25] Busija L, Lim K, Szoeke C, Sanders KM, McCabe MP. Do replicable profiles of multimorbidity exist? Systematic review and synthesis. European journal of epidemiology. 2019 Nov;34:1025–1053. Available from: *https://doi.org/10.1007/s10654-019-00568-5*.

[26] Stirland LE, González-Saavedra L, Mullin DS, Ritchie CW, Muniz-Terrera G, Russ TC. Measuring multimorbidity beyond counting diseases: systematic review of community and population studies and guide to index choice. BMJ. 2020 Feb;368:m160. Available from: *http://www.bmj.com/content/368/bmj.m160.abstract*.

[27] Ho ISS, Azcoaga-Lorenzo A, Akbari A, Black C, Davies J, Hodgins P, et al. Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies. The Lancet Public Health. 2021 Aug;6(8):e587–e597. Available from: *https://doi.org/10.1016/S2468-2667(21)00107-9*.

[28] St Sauver JL, Boyd CM, Grossardt BR, Bobo WV, Finney Rutten LJ, Roger VL, et al. Risk of developing multimorbidity across all ages in an historical cohort study: differences by sex and ethnicity. BMJ Open. 2015 Feb;5(2):e006413. Available from: *http://bmjopen.bmj.com/content/5/2/e006413.abstract*.

[29] Coste J, Valderas JM, Carcaillon-Bentata L. Estimating and characterizing the burden of multimorbidity in the community: A comprehensive multistep analysis of two large nationwide representative surveys in France. PLOS Medicine. 2021 Apr;18(4):e1003584. Available from: *https://doi.org/10.1371/journal.pmed.1003584*.

[30] Ward BW, Schiller JS. Prevalence of Multiple Chronic Conditions Among US Adults: Estimates From the National Health Interview Survey, 2010. Prev Chronic Dis. 2013 Apr;10:E65. Available from: *http://dx.doi.org/10.5888/pcd10.120203*.

[31] Goodman RA, Ling SM, Briss PA, Parrish RG, Salive ME, Finke BS. Multimorbidity Patterns in the United States: Implications for Research and Clinical Practice. J Gerontol A Biol Sci Med Sci. 2016 Feb;71(2):215–220. Available from: *https://doi.org/10.1093/gerona/glv199*.

[32] Calderón-Larrañaga A, Saadeh M, Hooshmand B, Refsum H, Smith AD, Marengoni A, et al. Association of Homocysteine, Methionine, and MTHFR 677C>T Polymorphism With Rate of Cardiovascular Multimorbidity Development in Older Adults in Sweden. JAMA Netw Open. 2020 May;3(5):e205316–e205316. Available from: *https://doi.org/10.1001/jamanetworkopen.2020.5316*.

3 The Million Women Study and Methods

3.1 Introduction

The two aims of this chapter are to describe (1) the Million Women Study, which is the data source of the statistical analyses in this thesis, and (2) the methods that are common to the results chapters that follow.

3.2 The Million Women Study

3.2.1 Overall study design

The Million Women Study is an ongoing population-based prospective cohort study. It was set up in 1996, initially to study the relationship between the use of menopausal hormone therapy and the risk of breast cancer in women in the United Kingdom (UK). [1, 2] The Million Women Study is suitable for studying the risk of different health outcomes (e.g. disease incidence, mortality) associated with various risk factors because of the diverse range of prospectively collected risk factors and linkage with routinely collected health data.

Women were invited to participate in the Million Women Study through the National Health Service (NHS) Breast Screening Centres. In 1996-2001, approximately 1.3 million UK women aged 50 to 64 years joined the study. These women completed a recruitment questionnaire on a range of demographic, lifestyle, reproductive, and anthropometric factors. The women were followed up for hospital admission, cancer registration, and death through linkage to NHS databases. Resurvey questionnaires were posted to the participants every 3 to 5 years. All questionnaires are accessible on the study website. [3]

3.2.2 Ethics

Ethics approval for primary data collection was obtained separately for individual regions in 1996 and as a multi-centre cohort study in 1997 (REC reference 97/5/001). Since then, several amendments were sought and granted for further data collection (e.g. resurvey

questionnaires, new data linkage). Approval was granted by the Oxford and Anglia Multi-Centre Research Ethics Committee, which is now known as the East of England - Cambridge South Research Ethics Committee.

At recruitment, participants consented in writing to re-contact and follow-up via medical records, with the understanding that their data will be kept safe and will only be used for research purposes. There were no apparent health risks from participating in the study: pilot studies in 1994-1996 confirmed that adding a self-administered questionnaire to the usual invitation did not affect uptake of breast screening. [4] Participation in the Million Women Study has always been entirely voluntary, and information on options for withdrawal at any time are provided on the study website. A freephone line and postal address are available for any enquires for participants with no internet access.

3.2.3 Funding

The Million Women Study was set up in collaboration with the NHS Breast Screening Programme and is funded by the UK Medical Research Council (grants MR/K02700X/1, U137686853) and Cancer Research UK (grants C570/A1649, A29186).

3.2.4 Recruitment

3.2.4.1 Data collection

Recruitment for the Million Women Study was conducted via collaborating breast screening centres in 1996-2001. At the time, the NHS Breast Screening Programme invited women aged 50 to 64 years, who were registered with a general practitioner, to mammographic screening every three years. Sixty-one screening centres in England and five centres in Scotland participated in the study, covering about half of all screening centres in the UK. Figure 3.1 shows the location of screening centres which took part in recruitment for the Million Women Study.

The Million Women Study recruitment questionnaire (Appendix B, page 298) was posted in the same envelope as the screening invitation sent out from each participating centre. Almost three million women were invited to join the study. Women were asked to complete the questionnaire (including the consent form) and bring it to their screening appointment if they wished to participate. Around 1.25 million women in England and 120,000 women in Scotland completed the questionnaire and joined the study, which was approximately one in four of all UK women aged 50 to 64 years in 1996-2001.



Figure 3.1 Location of UK NHS Breast Screening Centres which participated in recruitment for the Million Women Study

Reproduced from Figure 1 of the original article 'The Million Women Study Collaborative Group: The Million Women Study: design and characteristics of the study.' [1] The original article is an open access article distributed under the terms of the Creative Commons Attribution License [http://creativecommons.org/ licenses/by/2.0], which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

3.2.4.2 Data entry and storage

Each participating breast screening centre collected the completed study questionnaires and sent them to the study coordinating centre in the Cancer Epidemiology Unit of the University of Oxford. Trained staff shredded all questionnaires which did not have signed consent and scanned the remainder using an optical character recognition software [Eyes and Hands, Readsoft Inc, Slough, UK]. Random samples of the scanned data were validated against the original questionnaires. Paper questionnaires were shredded as soon as possible once they were saved in electronic format. Any personally-identifiable information was removed before the data were entered into a central database.

All study data are stored and maintained on a secure site in Oxford. Only authorised staff have access to the high security areas that store the data. Researchers can access subsets of the database upon approval of an application to the principal investigators.

3.2.4.3 Participant response rate and characteristics

Participants of the Million Women Study were largely representative of women who attended breast cancer screening. The first 227,000 recruitment questionnaires sent out by the Million Women Study coordinating centre were analysed to estimate the typical response rate. [1] Among them, 121,000 (53%) were completed and returned to the study coordinating centre, representing 71% of women who attended breast screening at the participating centres. About 7% of the completed questionnaires did not include written consent or failed to provide enough identifying information for follow-up.

According to a study of two general practices in England in 1998, women who attended breast screening were from less deprived areas and were more likely to have had a prescription for menopausal hormone therapy than women who did not attend screening.

[5] The two groups of women showed no differences in age and recent prescriptions of medication for hypertension, heart disease, hypercholesterolaemia, and diabetes mellitus.

The initial cohort available for analysis in the Million Women Study was 1.37 million women. Standard exclusions excluded participants with unsigned forms, those with missing information on age or aged under 48 years at recruitment, those without a matching ONS record, those with missing information on follow-up date, or those whose date of censoring at ONS was prior to recruitment (Figure 3.2). Table 3.1 presents selected characteristics of women at recruitment after standard exclusions.

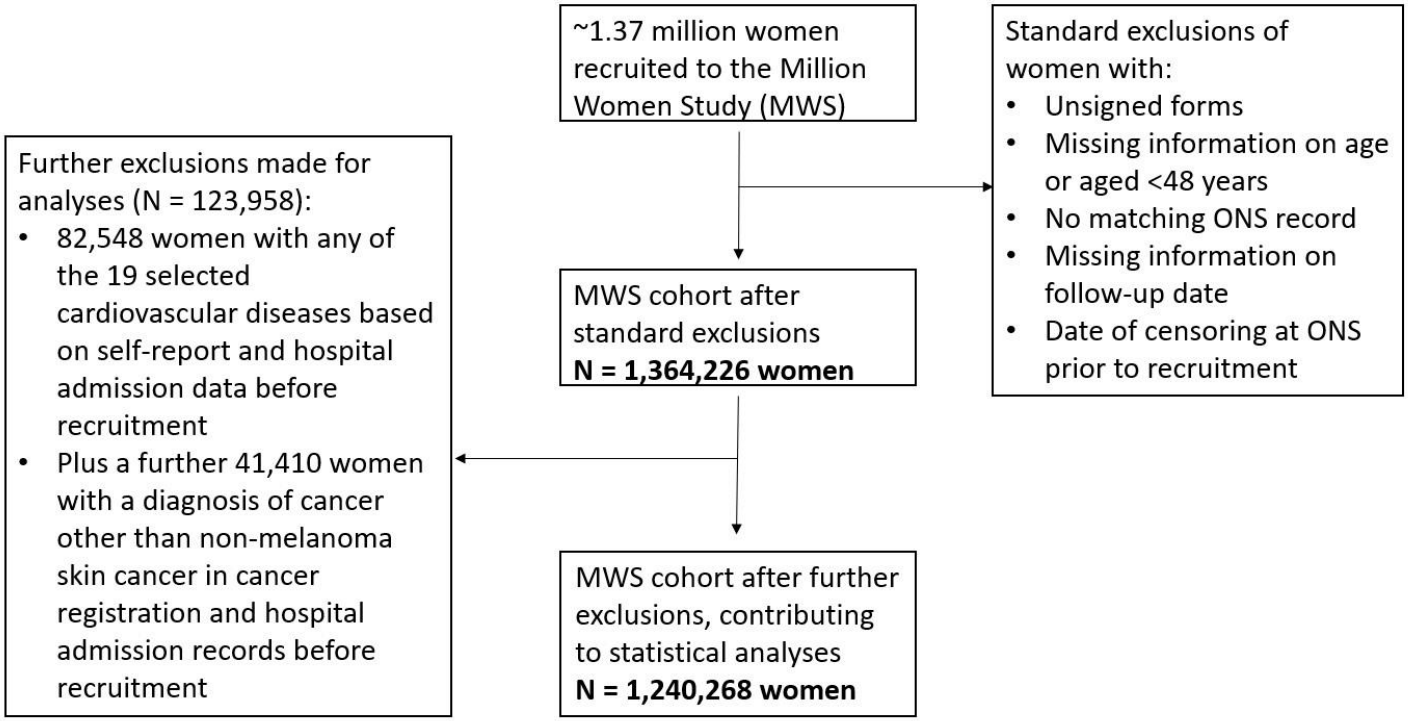


Figure 3.2 Flowchart showing standard exclusions for the Million Women Study and further exclusions leading to the study sample contributing data for the statistical analyses presented in this thesis

Table 3.1 Selected characteristics of all study participants at recruitment, after standard exclusions

Women, n	1,364,226
Age at recruitment, mean (SD), years	56.7 (4.9)
Lifestyle and anthropometric factors	
Current smoker, % (n)	20.5 (262,825)
Drinker of alcoholic beverages, % (n)	64.1 (867,990)
No strenuous physical activity, % (n)	61.3 (804,431)
Body mass index, mean (SD), kg/m^2	26.2 (4.7)
Height, mean (SD), cm	162.0 (6.6)
Socio-economic factors	
Most deprived quintile, % (n)	19.9 (269,650)
No educational qualifications, % (n)	43.8 (581,868)
Reproductive factors	
Age at menarche, mean (SD), years	13.0 (1.6)
Age at first birth, mean (SD), years	23.8 (4.3)
Age at natural menopause, mean (SD), years	49.1 (4.3)
Pre/perimenopausal, % (n)	12.3 (167,164)
Parity, mean (SD)	2.2 (1.3)
Ever breastfed children in parous women, % (n)	67.9 (652,842)
Ever use of contraceptive pills, % (n)	59.1 (797,065)
Ever use of hormone replacement therapy, % (n)	50.0 (674,220)
Morbidities	
Being treated for hypertension, % (n)	16.1 (219,328)
Being treated for diabetes mellitus, % (n)	2.5 (33,750)
Being treated for high blood cholesterol, % (n)	3.9 (53,829)
Being treated for/Ever had heart disease, % (n)	4.7 (64,431)
Ever had stroke, % (n)	1.2 (16,306)

3.2.5 Follow-up

Participants were mainly followed-up via electronic linkage to routinely-collected NHS databases using their unique NHS number, gender, age, and postcode. The databases contained nearly complete information on emigrations, hospital admissions (day case and inpatient), and deaths. Additional linked data (e.g. primary care data) were available for subsets of participants. Table 3.2 is an overview of linked data which was available for follow-up in the Million Women Study at the time of data extraction for this thesis. Only linked data for deaths, cancer registration, and hospital admissions were used in the analyses presented in this thesis. Nearly 20 years on average after recruitment, only about 1.5% of the Million Women Study cohort was lost to follow-up. During this period, 13% of participants died, 15% had a cancer registration, and 85% had one or more hospital admissions.

Table 3.2 Overview of all electronically linked data available for follow-up in the Million Women Study at the time of data extraction for this thesis

Data	Data provider	Available dates
All participants in England		
Deaths (e.g. causes of death)	NHS Digital	1953-2017
Loss to follow-up (e.g. reason, date)	NHS Digital	1953-2017
Cancer registration (e.g. cancer site, tumour morphology)	NHS Digital	1953-2017
Hospital inpatient/day patient admissions (e.g. diagnoses, procedures)	NHS Digital	1997-2017
Breast, cervix, and bowel screening invitation date and attendance	PHE	Up to 2013
Tumour characteristics (e.g. stage, grade)	PHE	Up to 2017
All participants in Scotland		
Deaths (e.g. causes of death)	ISD	1948-2017
Loss to follow-up (e.g. reason, date)	ISD	1948-2017
Cancer registration (e.g. cancer site, tumour morphology)	ISD	1948-2017
Hospital inpatient/day patient admissions (e.g. diagnoses, procedures)	ISD	1981-2017
100,000 women in England		
Primary care diagnosis and prescribing	CPRD	Up to 2013
100,000 women in Scotland		
Dispensing of prescription drugs in primary care	ISD	2009-2015

ISD=Information Services Division Scotland. PHE=Public Health England. CPRD=Clinical Practice Research Datalink.

3.2.5.1 Follow-up for deaths and cancer registration

Information on deaths and cancer registrations came from the Office for National Statistics (ONS), via NHS Digital in England and the Information Services Division (ISD) in Scotland (which has been part of Public Health Scotland since April 2020).

Information on the date of death, primary and secondary causes of death, and changes in registration with the NHS due to emigration or other loss to follow-up were provided. Causes of death were coded using the International Classification of Diseases, 10th revision (ICD-10) [6] codes. The end date of follow-up for mortality was 31 December 2017 at the time when the analyses in this thesis were conducted; 13% of Million Women Study participants had died by that date. The quality of death certification data is routinely checked by the ONS. [7, 8]

Cancer registration data provided information on the date of cancer registration and cancer site (coded using ICD-9 or ICD-10). [9, 10] The end date of follow-up for cancer registration was 31 December 2017 at the time of analysis; 15% of Million Women Study participants had at least one cancer registration by that date.

3.2.5.2 Follow-up for hospital admissions

Hospital admission data were acquired from the NHS Hospital Episode Statistics (HES) in England [11] through NHS Digital, and from the Scottish Morbidity Records (SMR) in Scotland [12] through ISD Scotland. At the time when data were extracted for this thesis, HES was available from 1 April 1997 to 31 March 2017, and SMR from 1 January 1981 to 31 March 2017. In the Million Women Study, 85% of women had at least one hospital admission by 31 March 2017. Diagnoses recorded in each hospital episode were coded using ICD-10; ICD-9 was used until April 1995. Procedures were coded using the Office

of Population Censuses and Surveys classification of surgical operations and procedures, fourth revision (OPCS-4); OPCS-3 was used until 1988.

In England, HES contains information on all admissions, accident and emergency attendances, and outpatient appointments at NHS hospitals in England. [11] This includes patients who reside outside of England, private patients seen in NHS hospitals, and care in treatment facilities funded by the NHS, including those in the independent sector. The Million Women Study participants were linked to the Admitted Patient Care (APC) data [13] which includes information on hospital admissions but were not linked to accident and emergency or outpatient data. HES APC is routinely collected for planning hospital services and are currently used for reimbursing hospital activity (a system called Payment by Results). HES APC data are generated from clinicians' patient notes, which are coded into ICD-10 and OPCS-4 by clinical coders in each hospital according to national guidelines. Data are compiled, cleaned, and validated centrally by NHS Digital. HES data are audited internally at various stages in the processing cycle [14], and data quality is measured using the Data Quality Maturity Index (DQMI). [15] DQMI reports are published monthly by NHS Digital.

In Scotland, SMR is compiled in a similar way to HES. SMR contains detailed information on deaths, cancer diagnoses, and hospital admissions. Hospitals in Scotland are mostly funded by contracts agreed with local health boards rather than based on hospital activity like in England. This could potentially mean that Scottish hospitals do not have the same incentive to record diagnoses in hospital admission data to the same level of detail or accuracy as English hospitals. There are other differences in health policy and healthcare usage that may result in some differences between HES and SMR. [16] The ISD's Data Quality Assessment team routinely audits SMR data at source hospitals to check for accuracy and adherence to national standards. [17] The national accuracy of SMR for

acute inpatient and day case admissions remained stable between the years 2004 and 2015, at around 89% for main condition (i.e. primary diagnosis) and 94% for main operation at the three-character level. [18]

Routinely collected UK hospital admissions data are often used for clinical and epidemiological studies due to the extensive coverage of the population and the long data collection period. The standardised coding system for clinical diagnoses makes it convenient for international comparisons and secondary uses. However, since the data are not collected for research purposes and coding practices may vary across hospitals or change over time, it is possible that some diagnoses or procedures (such as those less important for financial reimbursement or considered less clinically relevant) may be less accurately coded than others. For example, in 2013 a list of mandatory comorbidities that must always be considered as clinically relevant and hence recorded as secondary diagnoses was issued by the NHS (Coding Clinic Ref 88: Coding of Comorbidities), which may have improved the coding of those conditions since then. For analytical studies in epidemiology, substantial inaccuracies or inconsistencies in coding may bias the effect estimates if there are systematic differences in coding across the levels of the investigated risk factors.

Nevertheless, a 2012 systematic review of studies reported good discharge coding accuracy in the UK: the median accuracy of routinely collected data (when compared with case notes or clinical registries) was 80.3% (IQR 63.3-94.1) for ICD codes and 84.2% (68.7-88.7) for OPCS codes overall. [19] Since the introduction of Payment by Results in England in 2003-2004, the median accuracy of primary diagnoses increased from 73.8% (59.3-92.1) to 96.0% (89.3-96.3). Most studies included in this review compared diagnostic codes at the four-character level, so three-character level accuracy is likely to be

higher than the reported figures. The study reported no apparent differences in accuracy between England and Scotland.

3.2.5.3 Resurvey questionnaires

Resurvey questionnaires were posted directly to the Million Women Study participants from the year 1999, to collect information on new risk factors or risk factors that may change over time (e.g. smoking habits, weight, physical activity). Follow-up questionnaires were sent out to all surviving participants roughly 3, 8, 12, 15, and 18 years since recruitment. There was some loss to follow-up in the resurvey questionnaires compared to the sample at recruitment; for instance, the 3-year resurvey questionnaire was completed by 67% of Million Women Study participants at recruitment. For subsets of the cohort, more data (e.g. diet) has been collected via additional postal and online questionnaires. In this thesis, data on self-rated health from the 3-year resurvey was used to control for potential reverse causation bias in a sensitivity analysis presented in Chapter 5, but data from resurvey questionnaires were not otherwise used.

3.2.6 Cardiovascular research in the Million Women Study

The Million Women Study data have already been used to study a broad range of risk factors in relation to individual cardiovascular outcomes. Risk factors including diabetes [20, 21], depression [22], adiposity [20, 23-26], smoking [20, 27], physical activity [28], menopausal hormone therapy [29], age at menarche [30], hypertension in pregnancy [31], social participation [32], socioeconomic status [33], and marital status [34] have been prospectively studied in relation to outcomes such as coronary heart disease [21, 24, 25, 28, 30-34], cerebrovascular disease [20, 21, 23, 28, 30, 31], and venous thromboembolism [21, 22, 26-29]. Subsequent chapters will discuss these studies and their findings in more detail, as relevant to the content of each chapter.

3.3 Methods

3.3.1 Outcomes

The overarching aim of this thesis was to investigate the prospective associations of various risk factors with incident cardiovascular multimorbidity (CVM). A 2021 systematic review of 566 studies on multimorbidity found that the most widely used definition of multimorbidity was the co-occurrence of two or more chronic conditions in one person, measured by counting the number of diseases each person has from a list of selected individual diseases. [35] This was also the basic definition used in this thesis: CVM was defined as the co-occurrence of two or more CVDs from a list of selected individual CVDs.

As discussed in Section 2.4.2 of Chapter 2, one way to ensure that risk factor associations are interpretable and comparable across studies is to define multimorbidity using an extensive list of diseases with similar aetiology while also investigating more defined and comparable forms of multimorbidity such as disease pairs. With this aim, individual CVDs to comprise CVM in this thesis were systematically selected to comprehensively capture common and serious CVDs. Section 3.3.1.1 describes the selection process, and Section 3.3.1.2 describes the exclusions made to the study cohort given the selection of these individual CVDs. Finally, Section 3.3.1.3 describes CVM outcomes investigated in addition to the basic definition of CVM in this thesis: complex CVM and CVM pairs.

3.3.1.1 Selection of individual cardiovascular diseases

Individual CVDs to comprise CVM in this thesis were selected using the ICD-10 classification system and linked follow-up data in a four-step process to comprehensively capture serious CVDs which are common in the Million Women Study cohort. My supervisors provided guidance throughout the selection process.

In the first step, each three-character code in Chapter IX (Diseases of the Circulatory System) of the ICD-10 was assessed to exclude (1) acute conditions which are not indicative of underlying chronic diseases (e.g. I46: cardiac arrest), (2) common antecedent conditions which are not typically in and of themselves serious disease states (e.g. I10: primary hypertension), and (3) non-specific codes (e.g. I27: other pulmonary heart disease). Non-specific codes that may represent important classes of conditions that could be grouped with a specific condition (e.g. I09: other rheumatic heart diseases) were not excluded.

In the second step, CVDs which are not included in Chapter IX of ICD-10 were identified and selected. Transient cerebral ischaemic attacks and related syndromes (G45), vascular dementia (F01), and vascular disorders of intestine (K55) were included from other chapters of the ICD-10.

In the third step, the three-character ICD-10 codes selected in the first two steps were further aggregated (e.g. I20-I25: ischaemic heart disease) or separated (e.g. I35.0/2 aortic stenosis) to form aetiologically and/or anatomically distinct diseases.

In the fourth and final step, individual CVDs which were relatively common in the Million Women Study cohort during follow-up were included for investigation as components of CVM in this thesis. An event was defined as the first hospital admission after recruitment with a relevant ICD-10 code in any position (i.e. primary or secondary diagnosis fields) or death with the diagnosis as the underlying cause, whichever came first. Only events that were recorded during follow-up (i.e. after the date of recruitment up to 31 December 2016) were counted. Note that these were not strictly incident events, since participants with prior CVD at recruitment had not been excluded yet. Nineteen CVDs with 5,000 or more events were finally included (Table 3.3). The selected CVDs could be broadly grouped

into three disease classes: heart diseases, cerebrovascular diseases, and other vascular diseases.

Table 3.3 List of selected individual cardiovascular diseases, corresponding diagnostic codes, and number of events during follow-up in the study cohort, after standard exclusions only

	ICD-10 code*	Events (N)**
Heart diseases		
Rheumatic heart disease	I05-I09	22,731
Hypertensive heart and renal disease	I12-13	14,329
Ischaemic heart disease	I20-25	154,991
Mitral valve disorders	I34	16,292
Aortic stenosis	135.0, I35.2	14,801
Other aortic valve disorders	I35.1, I35.8, I35.9	6,760
Cardiomyopathy	I42	5,162
Other arrhythmias	I44-45, I47, I49	54,398
Atrial fibrillation	I48	91,980
Heart failure	I11.0, I13.0, I13.2, I50	50,089
Cerebrovascular diseases		
Stroke	I60-61, I63, I64	40,530
Transient cerebral ischaemic attacks and related syndromes	G45	14,918
Other cerebrovascular disease	I62, I65-69	40,995
Vascular dementia	F01	7,810
Arterial, venous, and other vascular diseases		
Venous thromboembolism	I26, I80-82	35,039
Aortic aneurysm	I71	7,487
Peripheral vascular disease	I73.9	17,008
Arterial embolism and thrombosis	I74	5,069
Vascular disorders of the intestine	K55	7,994

*Diagnostic codes used to identify diseases from hospital admission and death records
**An event was defined as the first hospital admission after recruitment with a relevant ICD-10 code in any position or death with the diagnosis as the underlying cause, whichever came first. These were not considered incident events, since women with cardiovascular disease prior to recruitment were not excluded from the cohort yet.

3.3.1.2 Exclusions

Considering the selection of these individual CVDs, further exclusions were made to the study sample (Figure 3.2). The total number of participants after standard exclusions (described in Section 3.2.4.3, page 70) was 1,364,226 women. Of these women, 82,548 (6%) participants who had any of the 19 selected CVDs based on self-report and hospital admission data before recruitment were excluded to reduce reverse causation bias. A

further 41,410 (3%) women were excluded who had a recorded diagnosis of cancer (other than non-melanoma skin cancer, ICD-10 code: C44) in cancer registration and hospital admission records before recruitment. Having cancer can affect cardiovascular risk factors such as BMI or physical activity [36, 37], and cancer survivors have increased medium-term to long-term CVD risks [38], potentially leading to reverse causation bias. Non-melanoma skin cancer was an exception as it is very common and usually curable, and is not reliably recorded in cancer registries. [39] The remaining 1,240,268 women contributed data to the analyses presented in Chapters 4 to 7.

3.3.1.3 Definition of cardiovascular multimorbidity outcomes investigated in this thesis

An incident CVD event was defined as the first hospital admission after recruitment with a relevant ICD-10 code in any diagnosis field or death with the diagnosis as the underlying cause, whichever came first. In this thesis, the most basic CVM outcome investigated was incident CVM. Incident CVM was defined as one person having an incident event of two or more of the 19 selected CVDs (Table 3.3). The date of the second incident CVD event was assigned as the date of CVM. In addition to this basic definition of CVM, other CVM outcomes were explored: complex CVM and pairs of CVDs (i.e. CVM pairs).

Complex (or severe) multimorbidity is a concept that aims to capture relatively serious forms of multimorbidity, which has been defined in several different ways by previous studies (e.g. 3+ diseases from at least three different organ systems [40, 41], or 5+ conditions [42]). A common definition of complex multimorbidity is the co-occurrence of four or more diseases in one person. [43-45] In the UK, complex multimorbidity (4+ diseases) was highlighted as a priority for prevention and future service provision as population modelling estimated that between 2015 and 2035, numbers of older people with 4+ diseases will double. [45] In line with this definition, complex CVM was defined as

having an incident event of four or more of the 19 selected CVDs in this thesis. The date of the fourth incident CVD event was assigned as the date of complex CVM.

Multiple CVDs could be recorded during the same admission, and in those cases, each was counted towards CVM and complex CVM. To illustrate, 1,057 women had ischaemic heart disease and stroke recorded for the first time on the same date, and this was also therefore assigned as the date of onset of CVM. Out of all 114,685 women who developed incident CVM, 44,200 (38.5%) had multiple conditions recorded on the first date of record of any CVD, which was assigned as the date of onset of CVM.

To explore CVM outcomes which are relatively well-defined and potentially comparable with other studies, six common pairs of specific CVDs (i.e. CVM pairs) were also investigated. CVM pairs were selected based on the relatively high number of cases in the cohort (for statistical power) and the distinctiveness of the component CVDs from each other. The date of an incident CVM pair was the date of the second CVD in the pair. The selected CVM pairs were, in descending order of frequency: 'Ischaemic heart disease & Atrial fibrillation' (IHD-AF; n=26,239), 'Ischaemic heart disease & Heart failure' (IHD-HF; n=20,614), 'Ischaemic heart disease & Other Arrhythmias' (IHD-ARR; n=18,157), 'Atrial fibrillation & Heart failure' (AF-HF; n=18,042), 'Other arrhythmias & Heart failure' (ARR-HF; n=9,756), and 'Atrial fibrillation & Stroke' (AF-STR; n=9,562). Although the pairs of 'Atrial fibrillation & Other arrhythmias' (n=15,622) and 'Stroke & Other cerebrovascular disease' (n=13,154) were also common (in fact, more common than ARR-HF or AF-STR), they were not studied because the component conditions were not sufficiently distinct from each other.

3.3.2 Reliability of individual vascular outcomes in the Million Women Study

The reliability of vascular outcomes identified through HES compared with general practice records has been assessed by a previous study in the Million Women Study for coronary heart disease (ICD-10 codes I20-I25), cerebrovascular disease (G54, I60-I69), and venous thromboembolism (I26, I80-I82). [46] Among a random sample of 2,323 women with a HES diagnosis of any of the three vascular outcomes, 93% (n=2,157) also had a record of diagnosis in general practice records. Among 864 women with no HES diagnosis of vascular disease, 97% (n=838) also had no record in general practice records. Agreement between HES and general practice records was highest for serious and specific outcomes such as myocardial infarction and pulmonary embolism. The 3% (n=26) who did not have a hospital admission record of vascular disease but did have a diagnosis in general practice records had less serious forms of disease; none had a diagnosis of myocardial infarction, confirmed stroke, or pulmonary embolism.

Based on these findings, the authors estimated that a substantial number of less severe vascular diseases were missed by hospital admission records. Approximately 35,000 cases of vascular disease in the study cohort were estimated to have been missed by hospital admission data in addition to the 61,000 cases identified through hospital admission data between 1997-2005. Given that there were nearly four times as many vascular disease cases recorded at the time of analysis for this thesis compared to when the validation study was conducted, the number of chronic and less severe cardiovascular conditions was likely to be proportionately underestimated in this thesis.

Nevertheless, a previous study in the Whitehall-II cohort in the UK which assessed the validity of CVD outcomes in linked HES records found that risk factor-disease associations remained unchanged when CVDs were ascertained using the gold standard

method of biomedical examinations. [47] Likewise, it was unlikely that proportions of CVD events which were not recorded in HES and mortality records varied substantially across different levels of the risk factors investigated in this thesis, and therefore the estimated relative risks were likely to be reasonably reliable.

3.3.3 Risk factors

As reported in the systematic review in the previous chapter (Chapter 2), there are currently no published prospective cohort studies which have investigated cardiovascular risk factors for co-occurring CVDs (i.e. CVM as defined in this thesis). Therefore, to provide an initial overview of potential risk factors for CVM, a wide range of self-reported anthropometric, lifestyle, and reproductive factors were investigated in this thesis. Previous studies have reported good reliability of self-reported measures of smoking habits [48-50], physical activity [51], and alcohol intake [52, 53], among other risk factors. Several validation studies of these and other risk factors have also been conducted in the Million Women Study. [54-57] The method of measurement and validity of the risk factors studied will be described in more detail in the following chapters.

3.3.4 Statistical methods

3.3.4.1 Cox proportional hazards regression

Cox proportional hazards regression was used to calculate hazard ratios (HRs) and 95% confidence intervals (CIs) for the CVM outcomes, in relation to the selected risk factors. Attained age was the underlying time variable. Person-years were calculated from the date of recruitment until whichever came first: date of first hospital admission for CVM; date of first record of cancer diagnosis in hospital admission or cancer registration data; date of death; date of emigration or other loss to follow-up; or end of follow-up. Follow-up was censored at the date of first record of cancer to exclude potential indirect and iatrogenic associations in the relationship between risk factors and CVM. The end of follow-up was

on 31 December 2016. Hospital admission data before 1 April 1997 were not available in England, so follow-up was calculated from this date for around 5% of participants who were recruited before this date. Fewer than 1.5% of women were lost to follow-up.

Log-log plots [58, 59] were assessed to check the proportional hazards assumption. Although tests based on Schoenfeld residuals [60, 61] are commonly used to check the proportional hazards assumption, even small deviations from proportionality (which make no material difference to the effect estimates) can be found to be statistically significant in large studies like the Million Women Study, so visual assessment of log-log plots was chosen as the preferred method.

Differences in HRs across the categories of each risk factor were tested using the likelihood ratio test for heterogeneity, with a significance level of $\alpha = 0.05$. The Benjamini-Hochberg procedure was applied in each chapter to control the false discovery rate and account for multiple testing. [62] When multiple hypothesis tests are performed, small p-values are likely to occur by chance and result in incorrect rejections of the null hypotheses – the false discovery rate is the proportion of null hypotheses which are incorrectly rejected (i.e. false positives). The Benjamini-Hochberg procedure arranges the p-values associated with m null hypotheses in ascending order, where P_i is the p-value at the ith position with the associated hypothesis H_i. When k is the largest i for which $P_i \leq \frac{i}{m}\alpha$, the procedure rejects hypotheses $i = 1,2,3,\cdots,k$. This controls the false discovery rate at a level of a. In each chapter, m was determined as the total number of hypothesis tests performed in that chapter.

3.3.4.2 Likelihood-ratio χ^2 test statistic attenuation

The likelihood-ratio χ^2 test (LR) statistic quantifies how well the risk factor predicts the outcome in each given model: the better the risk factor predicts the outcome, the greater

the LR statistic. The LR statistic attenuation method [33, 63] was used to measure the extent to which the relationships between risk factors and CVM outcomes were accounted for by multivariable adjustment, based on the relative change in the LR statistics.

Often, the association between a risk factor and CVM is at least partially accounted for by other covariates, either due to confounding or because the covariates are mediators in the association. In such cases, the LR statistic associated with the risk factor in a model that includes the covariates will be smaller than in an otherwise-identical model that does not include the covariates. This is because the risk factor predicts the outcome less well in the presence of the other covariates. The relative change in the LR statistics for the risk factor in the adjusted and unadjusted models then quantifies the share of the association between the risk factor and CVM which could be accounted for by the covariates.

The relative change in the LR statistics was calculated as $\left(1 - \frac{LR\chi^2_{adj}}{LR\chi^2_{min}}\right) \cdot 100$, where $LR\chi^2_{min}$ is the LR statistic of the minimally adjusted model and $LR\chi^2_{adj}$ is the LR statistic of the additionally adjusted model. This was referred to as the 'attenuation in LR statistics' or '$\chi^2_{attenuation}$' in the text and figures. If the attenuation was great (e.g. >80%) after multivariable adjustment, the remaining observed associations were more likely to be due to residual confounding (since the covariates were measured with error), than when the attenuation was small (e.g. <20%). [64]

3.3.4.3 Statistical software and plotting of forest plots

All analyses were conducted in STATA (version 16.1) [65] and forest plots were generated in R (version 4.0.1) [66] using the Jasper package [67]. Where more than two groups were compared in the forest plots, floated confidence intervals (FCIs) were displayed (although conventional CIs were given in the text where two groups were explicitly compared). Floated confidence intervals allow for the comparison of each category with any other

without specifying a baseline group. [68] The size of the square boxes was inversely proportional to the standard error of the hazard ratio estimate, and the relative change in the LR statistic between the minimally and multivariable adjusted models were indicated as '$\chi^2_{attenuation}$' in the forest plots.

3.4 Conclusions

The Million Women Study is a large prospective cohort study with nearly complete follow-up which was representative of the women in the target age range at the time of recruitment. This cohort is suitable for investigating multimorbidity, since there are large numbers of events of individual CVDs and reasonably high numbers of co-occurring CVDs. Despite the likelihood that linked hospital admission data may have missed sizable numbers of non-hospitalised CVD events, risk factor associations were unlikely to be substantially biased as a result. It was possible in this thesis to comprehensively study co-occurring CVDs, even those that were not reliably investigated in other studies due to the small number of events. Overall, the Million Women Study is generalisable to the target population and has good internal validity. Its design is appropriate for studying the associations between various risk factors and incident CVM in middle-aged UK women.

3.5 References

[1] The Million Women Study Collaborative Group. The Million Women Study: design and characteristics of the study population. Breast Cancer Res. 1999 Dec;1:73. Available from: *https://doi.org/10.1186/bcr16*.

[2] Green J, Reeves GK, Floud S, Barnes I, Cairns BJ, Gathani T, et al. Cohort Profile: the Million Women Study. Int J Epidemiol. 2019 Feb;48(1):28–29e. Available from: *https://doi.org/10.1093/ije/dyy065*.

[3] The Million Women Study. Postal Questionnaires [Internet];. [cited 2020 Feb 11]. Available from: *http://www.millionwomenstudy.org/questionnaires/*.

[4] Banks E, Richardson A, Beral V, Crossley B, Simmonds M, Hilton E, et al. Effect on attendance at breast cancer screening of adding a self administered questionnaire to the usual invitation to breast screening in southern England. J Epidemiol Community Health. 1998 Feb;52:116–119. Available from: *http://dx.doi.org/10.1136/jech.52.2.116*.

[5] Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, et al. Comparison of various characteristics of women who do and do not attend for breast cancer screening. Breast Cancer Res. 2001 Dec;4:R1. Available from: *https://doi.org/10.1186/bcr418*.

[6] World Health Organization. ICD-10: international statistical classification of diseases and related health problems : tenth revision, 2nd ed. World Health Organization; 2004. Available from: *https://apps.who.int/iris/handle/10665/42980*.

[7] NRS. Quality assurance of administrativedata used in population statistics: Births and deaths data. National Records of Scotland; 2020. 9 p. Available from: *https://-www.nrscotland.gov.uk/files//statistics/population/quality-assurance/qaad-births-deaths-data.pdf*.

[8] ONS. Mortality statistics in England and Wales QMI [Internet]; 2020. [cited 2021 Mar 30]. Available from: *https://www.ons.gov.uk/peoplepopulationandcommunity/-birthsdeathsandmarriages/deaths/methodologies/-mortalitystatisticsinenglandandwalesqmi*.

[9] ISD Scotland. Scottish Cancer Registry [Internet]; 2020. [cited 2021 Apr 12]. Available from: *https://www.isdscotland.org/Health-Topics/Cancer/Scottish-Cancer-Registry/*.

[10] Henson KE, Elliss-Brookes L, Coupland VH, Payne E, Vernon S, Rous B, et al. Data Resource Profile: National Cancer Registration Dataset in England. International Journal of Epidemiology. 2019 Apr;49(1):16–16h. Available from: *https://doi.org/-10.1093/ije/dyz076*.

[11] NHS Digital. Hospital Episode Statistics (HES) [Internet]; 2020. [cited 2021 Jan 13]. Available from: *https://digital.nhs.uk/data-and-information/data-tools-and-services/-data-services/hospital-episode-statistics*.

[12] Kendrick SW, Clarke JW. The Scottish Record Linkage System. Health bulletin. 1993 Mar;51(2):72–9.

[13] Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). International

Journal of Epidemiology. 2017 Mar;46(4):1093–1093i. Available from: *https://doi.org/-10.1093/ije/dyx015*.

[14] NHS Digital. The processing cycle and HES data quality [Internet]; 2021. [cited 2021 Mar 30]. Available from: *https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics/the-processing-cycle-and-hes-data-quality*.

[15] NHS Digital. Data Quality [Internet]; 2020. [cited 2021 Jan 13]. Available from: *https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/data-quality*.

[16] Bevan G, Karanikolos M, Exley J, Nolte E, Connolly S, Mays N. The four health systems of the United Kingdom: how do they compare? London, UK: The Health Foundation and Nuffield Trust; 2014. 126 p. Available from: *https://-www.nuffieldtrust.org.uk/files/2017-01/4-countries-report-web-final.pdf*.

[17] ISD Scotland. Data Quality Assurance - products and services [Internet];. [cited 2019 Sep 17]. Available from: *https://www.isdscotland.org/Products-and-Services/Data-Quality/*.

[18] ISD Scotland. Data quality and assurance - Assessment of SMR01 Data Scotland 2014-2015. NHS National Services Scotland; 2016. Available from: *https://-www.isdscotland.org/Products-and-Services/Data-Quality/docs/Assessment-of-SMR01-Data-2014-15-report-181019.pdf*.

[19] Burns EM, Rigby E, Mamidanna R, Bottle A, Aylin P, Ziprin P, et al. Systematic review of discharge coding accuracy. Journal of Public Health. 2011 Jul;34(1):138–148. Available from: *https://doi.org/10.1093/pubmed/fdr054*.

[20] Price AJ, Wright FL, Green J, Balkwill A, Kan SW, Yang TO, et al. Differences in risk factors for 3 types of stroke. Neurology. 2018 Jan;90(4):e298–e306. Available from: *https://doi.org/10.1212/WNL.0000000000004856*.

[21] Spencer EA, Pirie KL, Stevens RJ, Beral V, Brown A, Liu B, et al. Diabetes and modifiable risk factors for cardiovascular disease: the prospective Million Women Study. Eur J Epidemiol. 2008 Nov;23(12):793. Available from: *https://doi.org/10.1007/s10654-008-9298-3*.

[22] Parkin L, Balkwill A, Sweetland S, Reeves GK, Green J, Beral V. Antidepressants, Depression, and Venous Thromboembolism Risk: Large Prospective Study of UK Women. J Am Heart Assoc. 2017 May;6(5):e005316. Available from: *https://doi.org/-10.1161/JAHA.116.005316*.

[23] Kroll ME, Green J, Beral V, Sudlow CLM, Brown A, Kirichek O, et al. Adiposity and ischemic and hemorrhagic stroke. Neurology. 2016 Oct;87(14):1473–1481. Available from: *https://doi.org/10.1212/WNL.0000000000003171*.

[24] Canoy D, Cairns BJ, Balkwill A, Wright FL, Green J, Reeves G, et al. Coronary heart disease incidence in women by waist circumference within categories of body mass index. Eur J Prev Cardiol. 2013 Oct;20(5):759–762. Available from: *https://doi.org/-10.1177/2047487313492631*.

[25] Canoy D, Cairns BJ, Balkwill A, Wright FL, Green J, Reeves G, et al. Body mass index and incident coronary heart disease in women: a population-based prospective study. BMC Med. 2013 Apr;11(1):87. Available from: *https://doi.org/10.1186/1741-7015-11-87*.

[26] Parkin L, Sweetland S, Balkwill A, Green J, Reeves G, Beral V. Body mass index, surgery, and risk of venous thromboembolism in middle-aged women: a cohort study. Circulation. 2012 Apr;125:1897–904. Available from: *https://doi.org/10.1161/-circulationaha.111.063354*.

[27] Sweetland S, Parkin L, Balkwill A, Green J, Reeves G, Beral V. Smoking, surgery, and venous thromboembolism risk in women: United Kingdom cohort study. Circulation. 2013 Mar;127(12):1276–82. Available from: *https://doi.org/10.1161/-circulationaha.113.001428*.

[28] Armstrong MEG, Green J, Reeves GK, Beral V, Cairns BJ. Frequent physical activity may not reduce vascular disease risk as much as moderate activity: large prospective study of women in the United Kingdom. Circulation. 2015 Feb;131(8):721–9. Available from: *https://doi.org/10.1161/circulationaha.114.010296*.

[29] Sweetland S, Beral V, Balkwill A, Liu B, Benson VS, Canonico M, et al. Venous thromboembolism risk in relation to use of different types of postmenopausal hormone therapy in a large prospective study. J Thromb Haemost. 2012;10(11):2277–2286. Available from: *https://doi.org/10.1111/j.1538-7836.2012.04919.x*.

[30] Canoy D, Beral V, Balkwill A, Wright FL, Kroll ME, Reeves GK, et al. Age at menarche and risks of coronary heart and other vascular diseases in a large UK cohort. Circulation. 2015 Jan;131(3):237–44. Available from: *https://doi.org/10.1161/-CIRCULATIONAHA.114.010070*.

[31] Canoy D, Cairns BJ, Balkwill A, Wright FL, Khalil A, Beral V, et al. Hypertension in pregnancy and risk of coronary heart disease and stroke: A prospective study in a large UK cohort. Int J Cardiol. 2016 Nov;222(27529390):1012–1018. Available from: *https://-doi.org/10.1016/j.ijcard.2016.07.170*.

[32] Floud S, Balkwill A, Canoy D, Reeves GK, Green J, Beral V, et al. Social participation and coronary heart disease risk in a large prospective study of UK women. Eur J Prev Cardiol. 2016 Jun;23(9):995–1002. Available from: *https://doi.org/10.1177/-2047487315607056*.

[33] Floud S, Balkwill A, Moser K, Reeves GK, Green J, Beral V, et al. The role of health-related behavioural factors in accounting for inequalities in coronary heart disease risk by education and area deprivation: prospective study of 1.2 million UK women. BMC Med. 2016 Oct;14(1):145. Available from: *https://doi.org/10.1186/s12916-016-0687-2*.

[34] Floud S, Balkwill A, Canoy D, Wright FL, Reeves GK, Green J, et al. Marital status and ischemic heart disease incidence and mortality in women: a large prospective study. BMC Med. 2014 Mar;12(1):42. Available from: *https://doi.org/10.1186/1741-7015-12-42*.

[35] Ho ISS, Azcoaga-Lorenzo A, Akbari A, Black C, Davies J, Hodgins P, et al. Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies. The Lancet Public Health. 2021 Aug;6(8):e587–e597. Available from: *https://doi.org/10.1016/S2468-2667(21)00107-9*.

[36] Nicholson BD, Hamilton W, O'Sullivan J, Aveyard P, Hobbs FDR. Weight loss as a predictor of cancer in primary care: a systematic review and meta-analysis. Br J Gen Pract. 2018 May;68(670):e311. Available from: *http://bjgp.org/content/68/670/-e311.abstract*.

[37] Irwin ML, Crumley D, McTiernan A, Bernstein L, Baumgartner R, Gilliland FD, et al. Physical activity levels before and after a diagnosis of breast carcinoma. Cancer. 2003 Apr;97(7):1746–1757. Available from: *https://doi.org/10.1002/cncr.11227*.

[38] Strongman H, Gadd S, Matthews A, Mansfield KE, Stanway S, Lyon AR, et al. Medium and long-term risks of specific cardiovascular diseases in survivors of 20 adult cancers: a population-based cohort study using multiple linked UK electronic health records databases. The Lancet. 2019 Sep;394(10203):1041–1054. Available from: *https://-doi.org/10.1016/S0140-6736(19)31674-5*.

[39] National Cancer Registration and Analysis Service. Non-melanoma skin cancer in England, Scotland, Northern Ireland and Ireland; 2013. [cited 2022 Apr 16]. Available from: *http://www.ncin.org.uk/publications/data_briefings/-non_melanoma_skin_cancer_in_england_scotland_northern_ireland_and_ireland*.

[40] Storeng SH, Vinjerui KH, Sund ER, Krokstad S. Associations between complex multimorbidity, activities of daily living and mortality among older Norwegians. A prospective cohort study: the HUNT Study, Norway. BMC Geriatrics. 2020 Jan;20(1):21. Available from: *https://doi.org/10.1186/s12877-020-1425-3*.

[41] Singer L, Green M, Rowe F, Ben-Shlomo Y, Kulu H, Morrissey K. Trends in multimorbidity, complex multimorbidity and multiple functional limitations in the ageing population of England, 2002-2015. J Comorb. 2019 Jan;9:2235042X19872030. Available from: *https://doi.org/10.1177/2235042X19872030*.

[42] Chamberlain AM, Finney Rutten LJ, Wilson PM, Fan C, Boyd CM, Jacobson DJ, et al. Neighborhood socioeconomic disadvantage is associated with multimorbidity in a geographically-defined community. BMC Public Health. 2020 Jan;20(1):13. Available from: *https://doi.org/10.1186/s12889-019-8123-0*.

[43] Prior A, Fenger-Grøn M, Larsen KK, Larsen FB, Robinson KM, Nielsen MG, et al. The Association Between Perceived Stress and Mortality Among People With Multimorbidity: A Prospective Population-Based Cohort Study. Am J Epidemiol. 2016 Aug;184(3):199–210. Available from: *https://doi.org/10.1093/aje/kwv324*.

[44] Kivimäki M, Strandberg T, Pentti J, Nyberg ST, Frank P, Jokela M, et al. Body-mass index and risk of obesity-related complex multimorbidity: an observational multicohort study. The Lancet Diabetes & Endocrinology. 2022 Apr;10(4):253–263. Available from: *https://www.sciencedirect.com/science/article/pii/S221385872200033X*.

[45] Kingston A, Robinson L, Booth H, Knapp M, Jagger C, for the MODEM project. Projections of multi-morbidity in the older population in England to 2035: estimates from the Population Ageing and Care Simulation (PACSim) model. Age Ageing. 2018 May;47(3):374–380. Available from: *https://doi.org/10.1093/ageing/afx201*.

[46] Wright FL, Green J, Canoy D, Cairns BJ, Balkwill A, Beral V, et al. Vascular disease in women: comparison of diagnoses in hospital episode statistics and general

practice records in England. BMC Med Res Methodol. 2012 Oct;12:161. Available from: *https://doi.org/10.1186/1471-2288-12-161*.

[47] Kivimäki M, Batty GD, Singh-Manoux A, Britton A, Brunner EJ, Shipley MJ. Validity of Cardiovascular Disease Event Ascertainment Using Linkage to UK Hospital Records. Epidemiology. 2017 Sep;28(5). Available from: *https://journals.lww.com/-epidem/Fulltext/2017/09000/Validity_of_Cardiovascular_Disease_Event.14.aspx*.

[48] Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ. 2004 Jun;328:1519. Available from: *https://doi.org/10.1136/bmj.38142.554479.ae*.

[49] Patrick DL, Cheadle A, Thompson DC, Diehr P, Koepsell T, Kinne S. The validity of self-reported smoking: a review and meta-analysis. Am J Public Health. 1994 Jul;84(7):1086–1093. Available from: *https://doi.org/10.2105/ajph.84.7.1086*.

[50] Brigham J, Lessov-Schlaggar CN, Javitz HS, Krasnow RE, Tildesley E, Andrews J, et al. Validity of recall of tobacco use in two prospective cohorts. Am J Epidemiol. 2010 Oct;172(7):828–835. Available from: *https://dx.doi.org/10.1093/aje/kwq179*.

[51] Wareham NJ, Jakes RW, Rennie KL, Schuit J, Mitchell J, Hennings S, et al. Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. Public Health Nutr. 2003 Jun;6(4):407–413. Available from: *https://-doi.org/10.1079/phn2002439*.

[52] Bohlscheid-Thomas S, Hoting I, Boeing H, Wahrendorf J. Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German part of the EPIC project. European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol. 1997 Jan;26(suppl_1):S59–S70. Available from: *https://-doi.org/10.1093/ije/26.suppl_1.s59*.

[53] Ocke MC, Bueno-de Mesquita HB, Pols MA, Smit HA, van Staveren WA, Kromhout D. The Dutch EPIC food frequency questionnaire. II. Relative validity and reproducibility for nutrients. Int J Epidemiol. 1997 Jan;26(suppl_1):S49–S58. Available from: *https://doi.org/10.1093/ije/26.suppl_1.s49*.

[54] Wright FL, Green J, Reeves G, Beral V, Cairns BJ, on behalf of the Million Women Study collaborators. Validity over time of self-reported anthropometric variables during follow-up of a large cohort of UK women. BMC Med Res Methodol. 2015 Oct;15:81. Available from: *https://doi.org/10.1186/s12874-015-0075-1*.

[55] Armstrong MEG, Cairns BJ, Green J, Reeves GK, Beral V, the Million Women Study Collaborators. Reported frequency of physical activity in a large epidemiological study: relationship to specific activities and repeatability over time. BMC Med Res Methodol. 2011 Jun;11:97. Available from: *https://doi.org/10.1186/1471-2288-11-97*.

[56] Roddam AW, Spencer E, Banks E, Beral V, Reeves G, Appleby P, et al. Reproducibility of a short semi-quantitative food group questionnaire and its performance in estimating nutrient intake compared with a 7-day diet diary in the Million Women Study. Public Health Nutr. 2005 Apr;8(2):201–213. Available from: *https://doi.org/-10.1079/phn2004676*.

[57] Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, et al. Agreement between general practice prescription data and self-reported use of hormone replacement therapy and treatment for various illnesses. J Epidemiol Biostat. 2001 Feb;6:357–363. Available from: *https://doi.org/10.1080/13595220152601837*.

[58] Ranstam J, Kärrholm J, Pulkkinen P, Mäkelä K, Espehaug B, Pedersen AB, et al. Statistical analysis of arthroplasty data. II. Guidelines. Acta orthopaedica. 2011 Jun;82:258–67. Available from: *https://doi.org/10.3109/17453674.2011.588863*.

[59] Hess KR. Graphical methods for assessing violations of the proportional hazards assumption in Cox regression. Statistics in medicine. 1995 Aug;14:1707–23. Available from: *https://doi.org/10.1002/sim.4780141510*.

[60] Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika. 1982 Apr;69(1):239–241. Available from: *https://doi.org/10.1093/biomet/-69.1.239*.

[61] Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika. 1994 Sep;81(3):515–526. Available from: *https://doi.org/-10.1093/biomet/81.3.515*.

[62] Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Series B Stat Methodol. 1995;57(1):289–300. Available from: *https://doi.org/10.1111/j.2517-6161.1995.tb02031.x*.

[63] Parish S, Peto R, Palmer A, Clarke R, Lewington S, Offer A, et al. The joint effects of apolipoprotein B, apolipoprotein A1, LDL cholesterol, and HDL cholesterol on risk: 3510 cases of acute myocardial infarction and 9805 controls. Eur Heart J. 2009 Sep;30(19520708):2137–2146. Available from: *https://doi.org/10.1093/eurheartj/ehp221*.

[64] Collaborative Group on Hormonal Factors in Breast Cancer. Alcohol, tobacco and breast cancer - collaborative reanalysis of individual data from 53 epidemiological studies, including 58 515 women with breast cancer and 95 067 women without the disease. British Journal of Cancer. 2002 Nov;87(11):1234–1245. Available from: *https://doi.org/10.1038/-sj.bjc.6600596*.

[65] StataCorp. Stata Statistical Software: Release 16; 2019. College Station, TX: StataCorp LLC.

[66] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2020. Available from: *http://www.R-project.org/*.

[67] Arnold M. Jasper: Jasper makes plots; 2020. R package version 2-266. Available from: *https://github.com/arnhew99/Jasper*.

[68] Easton DF, Peto J, Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. Stat Med. 1991 Jul;10:1025–35. Available from: *https://doi.org/10.1002/sim.4780100703*.

4 Descriptive study of cardiovascular multimorbidity in UK women

4.1 Introduction

4.1.1 Background

The occurrence of multiple cardiovascular diseases (CVDs) in one person is common [1, 2], but there are only a few large-scale studies that describe the components and patterns of cardiovascular multimorbidity (CVM), particularly in women (see Chapter 1). Most studies to date have limited the definition of CVM to a few CVDs, often focussing on ischaemic heart disease and stroke (see Chapters 1 & 2). It is important to understand the burden, distribution, and relationships of different patterns of co-occurring CVDs to develop effective prevention strategies for CVM. In this chapter, a descriptive study of CVM was conducted in the Million Women Study to understand which CVDs most commonly contribute to incident CVM in middle-aged and older UK women. The findings from this chapter informed the design of the analytical studies presented in the following chapters.

4.1.2 Objective and research questions

The main objective of this chapter is to describe the components of incident CVM in the Million Women Study participants, to investigate if certain CVDs disproportionately contribute to CVM or are found concurrently with other CVDs. The following research questions were addressed:

- What are the most common individual and pairs of CVDs that comprise CVM?
- By how much do CVDs occur together more often than expected by chance?
- Which CVDs occur at an earlier age, potentially driving the development of CVM?
- How does the distribution of CVM vary across individual CVDs? What are the most common cardiovascular comorbidities of each CVD?

4.2 Methods

4.2.1 Study design and participants

The design and participants of the Million Women Study are described in detail in Chapter 3, based on prior publications from the Million Women Study Collaborators. [3, 4] Briefly, the Million Women Study is a population-based cohort study of 1.3 million women enrolled between 1996 to 2001 through the National Breast Screening Programme in England and Scotland. Participants were 50 to 64 years of age at the time of recruitment. All participants were linked by their unique NHS identification number to the NHS Central Registers, which provided information on cancer registrations, deaths, and emigrations. Information on hospital admissions and diagnoses were provided by electronic record linkage to the Hospital Episode Statistics (HES) [5] in England and the Scottish Morbidity Records [6] in Scotland. Disease diagnoses and causes of death were coded using the International Classification of Diseases 10th edition (ICD-10). [7] The end of follow-up for the analyses in this thesis was 31 December 2016.

After excluding 82,548 women (6%) who had any of the 19 selected CVDs (see Table 3.3 in Chapter 3, page 81) recorded before recruitment (based on self-report and hospital admission data) and a further 41,221 women (3%) who had a recorded diagnosis of cancer (other than non-melanoma skin cancer) before recruitment, a total of 1,240,268 women who had completed the recruitment questionnaire were included in these analyses.

4.2.2 Outcomes and definitions

The rationale and process of outcomes selection for this thesis has been described in detail in Section 3.3.1 of Chapter 3 (page 79). Nineteen subtypes of CVD were selected based on clinical importance and number of cases in the Million Women Study cohort, primarily chosen from chapter IX (Diseases of the circulatory system) of the ICD-10 (see Table 3.3

in Chapter 3, page 81). These CVDs were broadly grouped into three disease classes: heart diseases, cerebrovascular diseases, and other vascular diseases.

The main outcome was incident CVM, defined as developing two or more diseases from the list of 19 selected CVDs during the study follow-up period. Complex CVM (having ≥4 CVDs) was also studied as a more serious form of CVM. Individual CVDs and CVM pairs (171 possible pairs of the selected CVDs) were studied as outcomes of secondary interest to understand the components of CVM.

An incident case of a selected CVD was defined as the first event after recruitment (following exclusion of women with prevalent CVD, as described above). An event was any record with a relevant ICD-10 code in the primary or secondary diagnosis fields in hospital admission records, or as the underlying cause of death. The date of CVM was defined as the date of record of the second condition, which could be the same as the date of the first condition if two or more conditions are first recorded on the same day. The date of complex CVM was defined as the date of record of the fourth condition, which similarly could be the same as the date of the first, second, and/or third conditions. The date of an incident CVM pair was defined as the latest of the dates of record of the two diseases in a pair (which again could be the same as the earliest of the dates of record).

4.2.3 Statistical analysis

4.2.3.1 Incidence of cardiovascular diseases and cardiovascular multimorbidity

For each CVD diagnosis, CVM pair, CVM, and complex CVM, the number of incident cases and proportion in the study population were calculated. The incidence per 1000 women of each CVM pair during the follow-up period was displayed in a heatmap generated using the ggplot2 package in R. [8, 9] The incidence and proportion of

individual CVDs were presented separately in the subsets of women who developed CVM and women who developed complex CVM.

4.2.3.2 Observed over expected frequency ratios of pairs of cardiovascular diseases

The observed over expected frequency ratio (O/E) of each CVM pair was calculated by dividing the observed number of each incident CVM pair by the expected number. Table 4.1 is an exemplar contingency table showing the number of participants with or without diseases A and B at the end of follow-up. In this example, the observed number of disease pair A-B is a. The chance of having disease A is $(a + b)/N$, and the chance of having disease B is $(a + c)/N$. Assuming diseases A and B are independent, the chance of having both diseases can be calculated as $[(a + b)(a + c)] \,/\, N^2$. Therefore, $[(a + b)(a + c)]/N$ is the number of disease pair A-B expected by chance alone. The O/E can then be calculated as $(a \times N) \,/\, [(a + b)(a + c)]$, representing the degree to which the diseases occurred together more often than expected by chance. The O/E values of all CVM pairs were presented in a heatmap. For each pair, Fisher's exact test [10-12] was performed to assess the evidence against independence of the diseases. Fisher's exact test was used because it is more robust when the data is sparse or unbalanced compared to a chi-squared test. [13]

Table 4.1 A 2x2 contingency table illustrating the relationship between the incidence of diseases A and B at the end of follow-up.

	With disease B	*Without disease B*	*Total*
With disease A	a	b	$a + b$
Without disease A	c	d	$c + d$
Total	$a + c$	$b + d$	$N = a + b + c + d$

4.2.3.3 Cardiovascular comorbidities of each cardiovascular disease

Crude and age-standardised proportions of the numbers of cardiovascular comorbidities which occurred with each individual CVD were presented in stacked bar charts. The numbers of cardiovascular comorbidities were presented in categories of 0, 1, 2, and 3+ (three or more). Age-standardised proportions were calculated by applying direct age standardisation to the 2013 European Standard Population [14], using 5-year age bands from age 45 up to 75+ years. The 2013 European Standard Population reflects the unweighted average age structure of the populations of all EU-27 and European Free Trade Association countries. [14] The data were derived from population projections for the period of 2011 to 2030, calculated based on 2010 census estimates provided to Eurostat. The use of a standard population allows for the adjustment of population-based rates (e.g. disease incidence, mortality) for different age distributions between groups. Age-standardised proportions of different numbers of comorbidities in patients with each individual CVD were compared to understand how the observed multimorbidity patterns of each CVD may be influenced by the age distribution of disease occurrence. The average age at first record of each CVD was also calculated.

The five most common cardiovascular comorbidities of each CVD were identified. All 31 possible multimorbidity combinations of each selected CVD with its top five comorbidities were counted and presented in UpSet plots. [15] The UpSet plot is used to visualise the frequency of intersections between multiple sets. It identifies all possible combinations of the sets and presents the frequency of each combination as a bar in a bar chart. It makes it easy to see the multiple sets that overlap in each combination, which is difficult in traditional set visualisations like Venn/Euler diagrams when there are more than three sets. There are no overlaps among the bars; for instance, members of set combination A-B-C are not counted among the members of set A-B. In the context of multimorbidity, the UpSet plot is useful for visualising the frequency of multiple co-occurring diseases beyond that of disease pairs.

4.2.3.4 Kaplan-Meier failure estimates

To understand the burden of CVM and CVDs in the cohort at different ages, estimates of the Kaplan-Meier failure function were calculated for CVM and individual CVDs at ages 60, 65, 70, and 75 years. These estimates represent the probability of developing CVM or CVDs in anyone who had lived to each specified age during follow-up. To illustrate the distribution of the component conditions of CVM over time, failure function estimates for each individual CVD were calculated separately in women who developed CVM and in those who developed complex CVM.

Using the Kaplan-Meier product-limit method [16-18], the survivor function estimates the probability of surviving to time t and beyond. Conversely, the Kaplan-Meier failure function estimates the probability of having had an event (i.e. failing) by time t.

Let t_i, $i = 1, \ldots$, be the times of failure events, n_i be the number at risk of failure just before time t_i, and d_i be the number of failures at time t_i. The nonparametric maximum likelihood estimate of the survivor function is:

$$\hat{S}(t) = \prod_{i|t_i \leq t} \left(\frac{n_i - \mathrm{d}_i}{n_i} \right)$$

Then, the Kaplan-Meier failure function is $1 - \hat{S}(t)$.

4.2.3.5 Coding of rheumatic heart disease

Rheumatic heart disease was selected for further analysis to explore whether it may have been overestimated in this cohort due to issues with disease coding, and if so, whether that may have substantially biased the composition of CVDs in CVM and complex CVM. Given that rheumatic heart disease is believed to be virtually eradicated in high-income countries like the UK [19], there was an unexpectedly high number of incident cases in this cohort. It was possible that a diagnosis of rheumatic heart disease in hospital admission records largely represented past medical history rather than incident disease, or that nonrheumatic heart valve disorders were misclassified as rheumatic heart disease.

To investigate, the number of hospital admissions with a record of incident rheumatic heart disease (I05-I09) in the Million Women Study was examined at the four-character ICD-10 code level and by diagnosis type (i.e. primary or secondary diagnosis). The five most common cardiovascular comorbidities of mitral valve stenosis (I05.0) were examined in separate UpSet plot, because I05.0 is most likely to represent true rheumatic heart disease. [20, 21] The number and proportion of CVDs that comprise CVM were compared before and after defining CVM based on the co-occurrence of the 18 individual CVDs other than rheumatic heart disease.

4.3 Results

4.3.1 Incidence of cardiovascular diseases and cardiovascular multimorbidity

After excluding women with prior CVD or cancer, 1,240,268 women were included in the analysis (Table 4.2). The mean age of participants was 56.6 (SD 4.8) years at recruitment. Women were followed up for 15.8 (4.8) years on average, and 114,684 (9.2%) had incident CVM. Ischaemic heart disease (9.7%) was by far the most common incident individual CVD, followed by atrial fibrillation (6.1%) and other arrhythmias (3.6%). Among cerebrovascular diseases, stroke (2.8%) and other cerebrovascular disease (2.7%) were the most common. Venous thromboembolism (2.5%) was the most common among other vascular diseases. There were at least 4000 incident cases of each selected individual CVD.

In the subset of women who developed CVM (≥2 CVDs), ischaemic heart disease and atrial fibrillation were the two most common incident CVDs (recorded in 54.9% and 46.4% of these women, respectively). A recorded diagnosis of heart failure (29.5%), other arrhythmias (28.0%), other cerebrovascular disease (22.3%), and stroke (20.6%) were also common in women with incident CVM.

In the subset of women who developed complex CVM (≥4 CVDs), the five most common incident CVDs were the same as in women with CVM, but with higher proportions, reflecting a higher total number of diagnoses. Ischaemic heart disease (73.1%) and atrial fibrillation (69.5%) were the most common incident CVDs in women who developed complex CVM. They were followed by heart failure (59.3%), other arrhythmias (44.2%), other cerebrovascular disease (34.4%) and rheumatic heart disease (33.8%).

Table 4.2 Number and proportion of incident cardiovascular disease and cardiovascular multimorbidity at the end of follow-up

Disease	ICD-10 code[a]	In all women [n= 1,240,268], % (n)	In women with CVM [n=114,684], % (n)	In women with complex CVM [n=26,896], % (n)
Number of diseases				
Any cardiovascular multimorbidity (CVM)	-	9.2 (114,684)	100 (114,684)	100 (26,896)
Complex CVM (i.e. >3 diseases)	-	2.17 (26,896)	23.5 (26,896)	100 (26,896)
Heart diseases				
Rheumatic heart disease	I05-I09	1.4 (17,348)	14.0 (16,014)	33.8 (9,104)
Hypertensive heart and renal disease	I12-13	0.9 (11,344)	6.9 (7,870)	13.4 (3,592)
Ischaemic heart disease	I20-25	9.7 (119,847)	54.9 (62,977)	73.1 (19,650)
Mitral valve disorders	I34	1.0 (12,518)	9.6 (11,041)	21.8 (5,867)
Aortic stenosis	135.0, I35.2	1.0 (11,993)	8.9 (10,196)	18.5 (4,970)
Other aortic valve disorders	I35.1, I35.8, I35.9	0.4 (5,208)	3.9 (4,497)	9.0 (2,410)
Cardiomyopathy	I42	0.3 (4,074)	3.2 (3,615)	7.5 (2,019)
Other arrhythmias	I44-45, I47, I49	3.6 (45,241)	28.0 (32,087)	44.2 (11,898)
Atrial fibrillation	I48	6.1 (76,058)	46.3 (53,141)	69.5 (18,688)
Heart failure	I11.0, I13.0, I13.2, I50	3.1 (38,172)	29.5 (33,874)	59.3 (15,946)
Cerebrovascular diseases				
Stroke	I60-61, I63, I64	2.8 (34,408)	20.6 (23,569)	29.3 (7,873)
Transient cerebral ischaemic attacks	G45	1.0 (12,550)	7.0 (7,989)	10.6 (2,845)
Other cerebrovascular disease	I62, I65-69	2.7 (32,940)	22.3 (25,522)	34.3 (9,236)
Vascular dementia	F01	0.5 (6,602)	4.2 (4,809)	7.0 (1,883)
Other vascular diseases				
Venous thromboembolism	I26, I80-82	2.5 (31,309)	12.7 (14,575)	16.5 (4,425)
Aortic aneurysm	I71	0.5 (6,097)	3.7 (4,215)	7.3 (1,971)
Peripheral vascular disease	I73.9	1.1 (13,193)	8.6 (9,858)	15.8 (4,256)
Arterial embolism and thrombosis	I74	0.3 (4,116)	3.0 (3,395)	5.8 (1,552)
Vascular disorders of the intestine	K55	0.5 (6,568)	3.1 (3,516)	4.6 (1,233)

[a]Diagnostic codes used to identify diseases from hospital admissions records; CVM=cardiovascular multimorbidity

4.3.2 Incidence of pairs of cardiovascular diseases

The incidence of CVM pairs (i.e. two co-occurring CVDs) greatly varied, ranging from 0.04 per 1000 persons for 'Vascular dementia & Cardiomyopathy' to 21.16 per 1000 persons for 'Atrial fibrillation & Ischaemic heart disease' over the follow-up period (Figure 4.1). Pairs involving common diseases such as ischaemic heart disease, atrial fibrillation, and heart failure had the highest incidence. The most common pair not involving these three common diseases was 'Stroke & Other cerebrovascular disease' (10.61 per 1000 persons). Pairs of CVDs within the same disease class (i.e. heart diseases, cerebrovascular diseases, and other vascular diseases) were generally more common than pairs between different classes. The most common between-class pairs were 'Stroke & Atrial fibrillation' (7.71), and 'Ischaemic heart disease & Other cerebrovascular disease' (7.62).

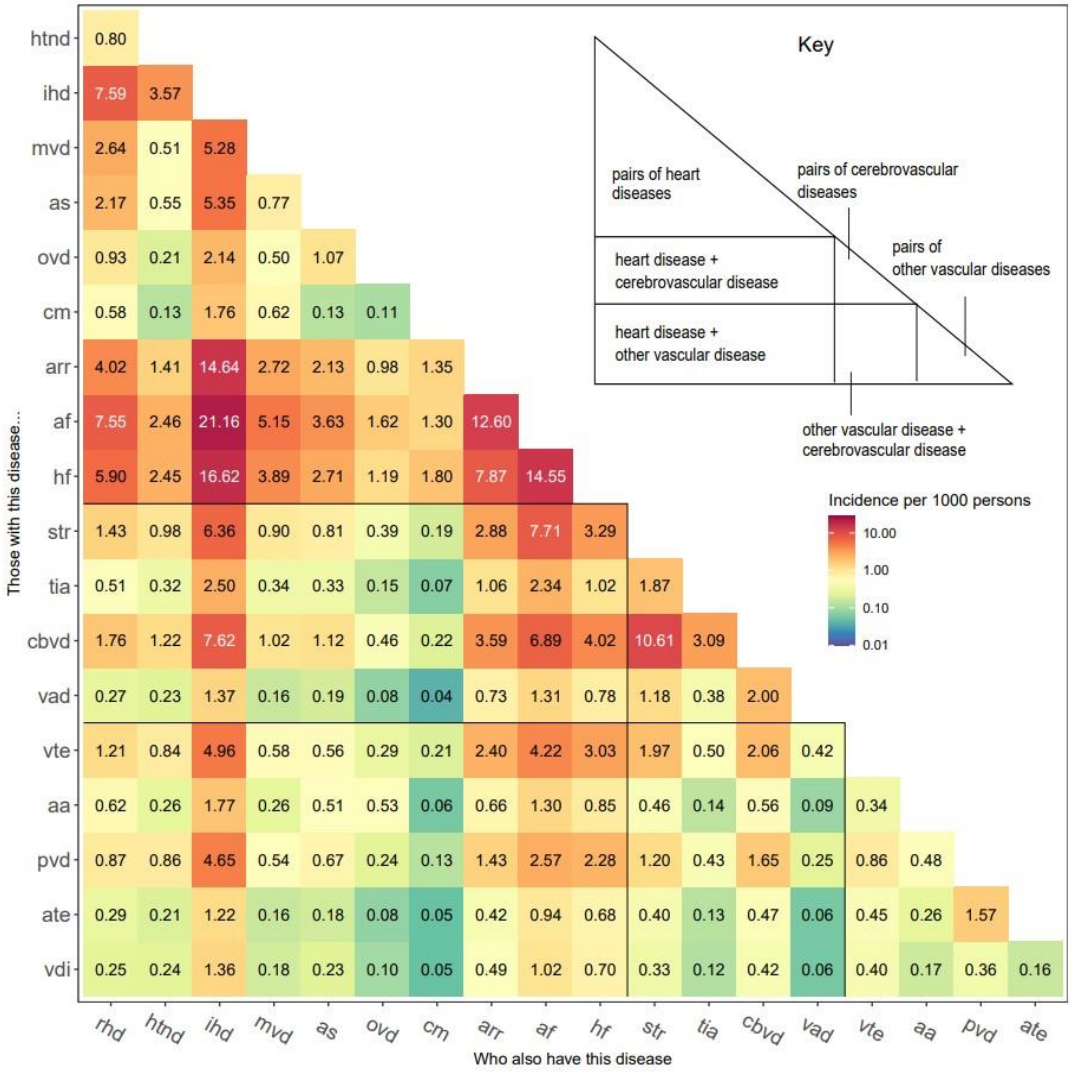


Figure 4.1 Incidence of each pair of cardiovascular diseases per 1000 women during the follow-up period

aa=aortic aneurysm, af=atrial fibrillation, arr=other arrhythmias, as=aortic stenosis, ate=arterial embolism and thrombosis, cbvd=other cerebrovascular disease, cm=cardiomyopathy, hf=heart failure, htnd=hypertensive heart and renal disease, ihd=ischaemic heart disease, mvd=mitral valve disorders, ovd=other aortic valve disorders, pvd=peripheral vascular disease, rhd=rheumatic heart disease, tia=transient cerebral ischaemic attacks, vad= vascular dementia, vdi=vascular disorders of the intestine, vte=venous thromboembolism

4.3.3 Co-occurrence of pairs of cardiovascular diseases

All CVM pairs occurred together at least 2 times more often than expected by chance (Figure 4.2). Fisher's exact test for association suggested strong evidence against independence of diseases in all pairs ($p<0.00001$). The pair of 'Arterial embolism and thrombosis & Peripheral vascular disease' showed by far the strongest association (O/E = 44.4), followed by pairs of 'Aortic aneurysm & Other aortic valve disorders' (25.7), and 'Aortic stenosis & Other aortic valve disorders' (26.5). In general, associations within disease classes and in closely related anatomical sites were much stronger than those between disease classes.

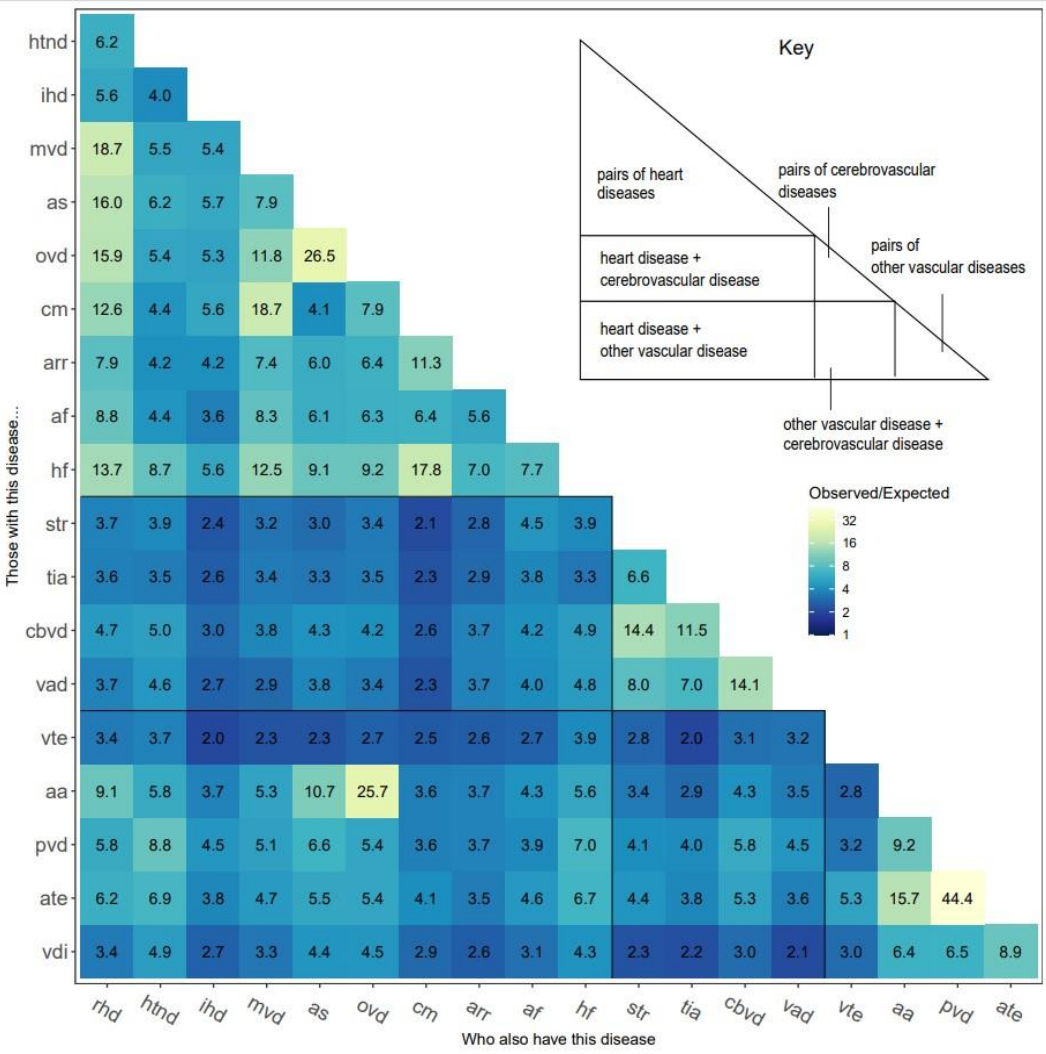


Figure 4.2 Ratio of observed to expected number of cases for each pair of cardiovascular diseases

aa=aortic aneurysm, af=atrial fibrillation, arr=other arrhythmias, as=aortic stenosis, ate=arterial embolism and thrombosis, cbvd=other cerebrovascular disease, cm=cardiomyopathy, hf=heart failure, htnd=hypertensive heart and renal disease, ihd=ischaemic heart disease, mvd=mitral valve disorders, ovd=other aortic valve disorders, pvd=peripheral vascular disease, rhd=rheumatic heart disease, tia=transient cerebral ischaemic attacks, vad= vascular dementia, vdi=vascular disorders of the intestine, vte=venous thromboembolism

4.3.4 Average age at first diagnosis

The average age at first diagnosis (i.e. age of onset) was 71.5 (SD 6.6) years for CVM, and 73.8 (6.2) for complex CVM (Table 4.3). Venous thromboembolism and ischaemic heart disease had the youngest average age of onset (68.4 years [SD 7.3] and 68.6 [6.9], respectively). The disease with the latest onset was vascular dementia (76.4 [5.6]).

Overall, there was no pattern of one class of disease occurring at a younger or older age than another (Figure 4.3). Ischaemic heart disease and cardiomyopathy were the heart diseases with the youngest age of onset, whereas aortic stenosis and rheumatic heart disease occurred latest in the heart disease class. Among cerebrovascular diseases, stroke and transient ischaemic attacks occurred at a younger age than other cerebrovascular diseases on average, and vascular dementia had the latest onset, even later than complex cardiovascular multimorbidity. Among other vascular diseases, venous thromboembolism occurred at the youngest age and aortic aneurysm at the oldest age of onset on average.

Table 4.3 Average age of onset of cardiovascular multimorbidity and individual cardiovascular diseases (sorted by age in descending order)

Disease	ICD-10 code[a]	Average age of onset	Standard deviation
Vascular dementia	F01	76.4	5.6
Complex CVM (≥4 diseases)	**-**	**73.8**	**6.2**
Rheumatic heart disease	I05-I09	72.8	6.3
Aortic stenosis	135.0, I35.2	72.7	6.4
Other cerebrovascular disease	I62, I65-69	72.4	6.9
Aortic aneurysm	I71	72.3	6.1
Heart failure	I11.0, I13.0, I13.2, I50	72.1	6.8
Atrial fibrillation	I48	71.6	6.6
Any CVM (≥2 diseases)	**-**	**71.5**	**6.6**
Other aortic valve disorders	I35.1, I35.8, I35.9	71.5	6.6
Mitral valve disorders	I34	71.2	6.6
Other arrhythmias	I44-45, I47, I49	70.7	7.1
Vascular disorders of the intestine	K55	70.4	6.7
Hypertensive heart and renal disease	I12-13	70.3	6.0
Stroke	I60-61, I63, I64	70.3	7.3
Transient cerebral ischaemic attacks	G45	70.0	6.9
Peripheral vascular disease	I73.9	69.9	7.0
Arterial embolism and thrombosis	I74	69.6	7.0
Cardiomyopathy	I42	69.5	6.5
Ischaemic heart disease	I20-25	68.6	6.9
Venous thromboembolism	I26, I80-82	68.4	7.3

[a]Diagnostic codes used to identify diseases from hospital admissions records; CVM=cardiovascular multimorbidity

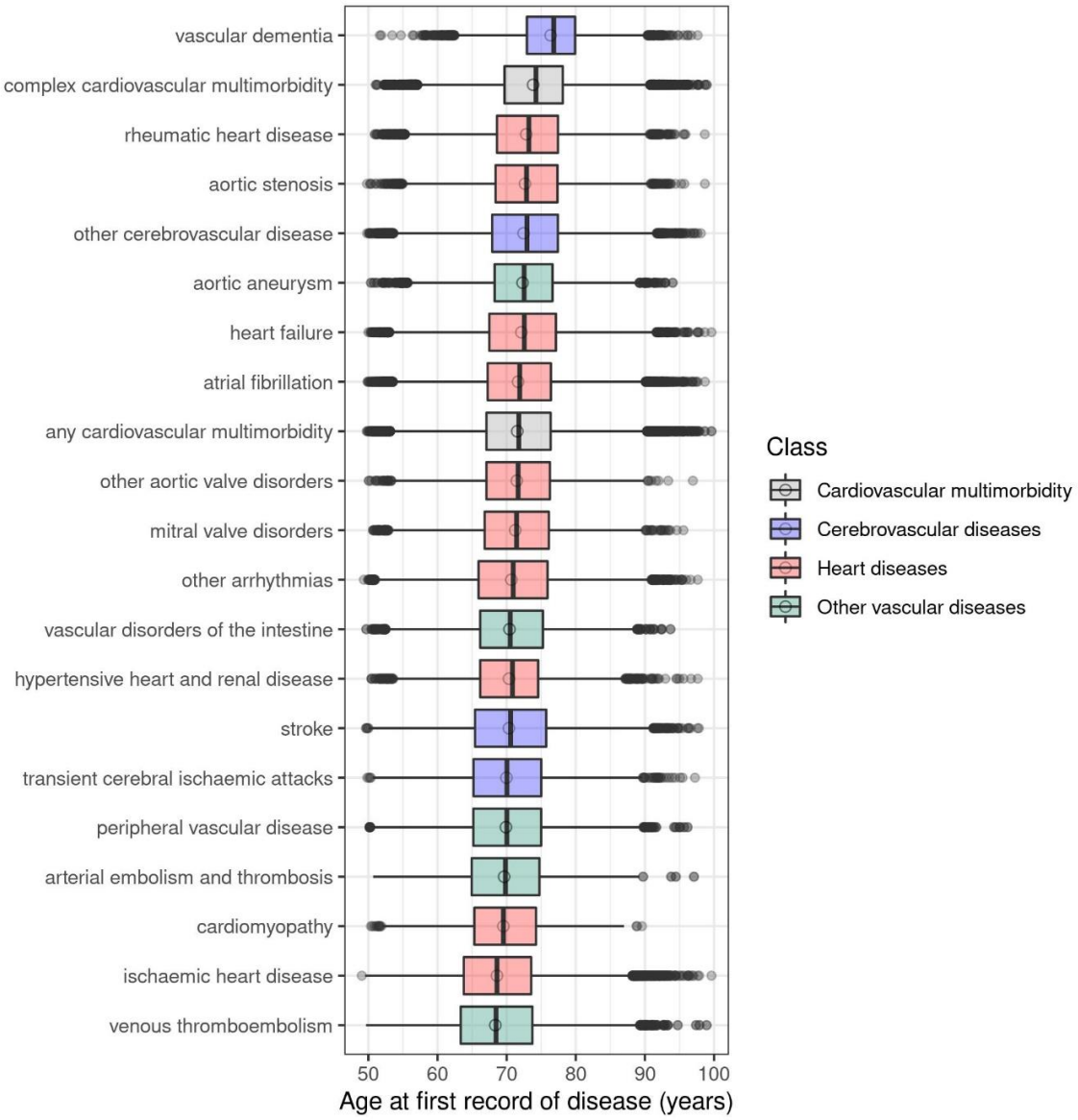


Figure 4.3 Box and whiskers plot displaying the distribution of age at first record of each individual cardiovascular disease and cardiovascular multimorbidity

The solid line in the middle of each box represents the median value. The hollow circle in the middle of each box represents the mean value. The box represents the interquartile range, and the whiskers extend to 1.5 times the interquartile range from each end of the box. The dots beyond the end of the whiskers represent extreme values.

4.3.5 The number of comorbidities of individual cardiovascular diseases

Among 271,510 women with a recorded diagnosis of at least one incident CVD, 114,684 (42%) had at least one other subtype of CVD (i.e. CVM). Figure 4.4 illustrates the crude proportion of the number of cardiovascular comorbidities which occurred with each individual CVD subtype during follow-up. The highest proportions of isolated diagnoses (i.e. incident events of an individual CVD which were not followed by an incident event of another, comorbid CVD before the end of follow-up) were found in the following: venous thromboembolism (53% isolated diagnosis), ischaemic heart disease (47%), and vascular disorders of the intestine (46%). The highest proportions of CVM were found in rheumatic heart disease (92% with at least one comorbidity), cardiomyopathy (89%), and heart failure (89%). Complex CVM (≥3 comorbidities or, ≥4 CVDs in total) was relatively uncommon for many individual CVD subtypes. The exceptions were cardiomyopathy, heart failure, rheumatic heart disease, and other mitral and aortic valve disorders – the proportion of complex CVM was between 40 to 50% for these subtypes.

After age-standardisation to the European Standard Population, the overall pattern of the frequencies of the number of comorbidities remained very similar, although the estimated proportions of women with no comorbidities declined in most individual CVDs (Figure 4.5). Overall, it was estimated that the majority of women who received a diagnosis of any of the 19 CVD subtypes by age 73 years (the mean age at the end of follow-up) were also diagnosed with at least one other of these diseases.

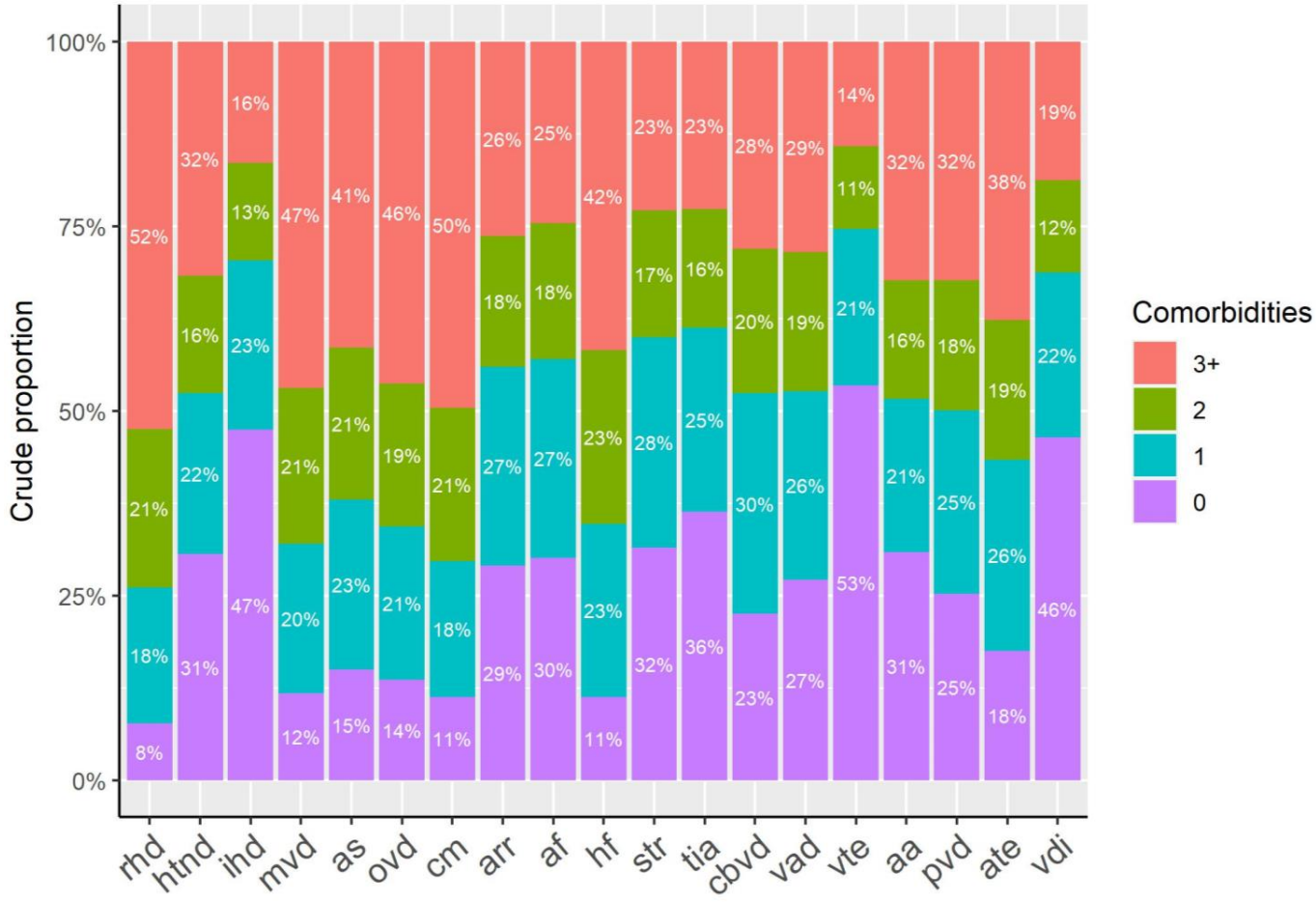


Figure 4.4 Crude proportion of the number of cardiovascular comorbidities by each incident individual cardiovascular disease

rhd=rheumatic heart disease, htnd=hypertensive heart and renal disease, ihd=ischaemic heart disease, mvd=mitral valve disorders, as=aortic stenosis, ovd=other aortic valve disorders, cm=cardiomyopathy, arr=other arrhythmias, af=atrial fibrillation, hf=heart failure, str=stroke, tia=transient cerebral ischaemic attacks, cbvd=other cerebrovascular disease, vad= vascular dementia, vte=venous thromboembolism, aa=aortic aneurysm, pvd=peripheral vascular disease, ate=arterial embolism and thrombosis, vdi=vascular disorders of the intestine

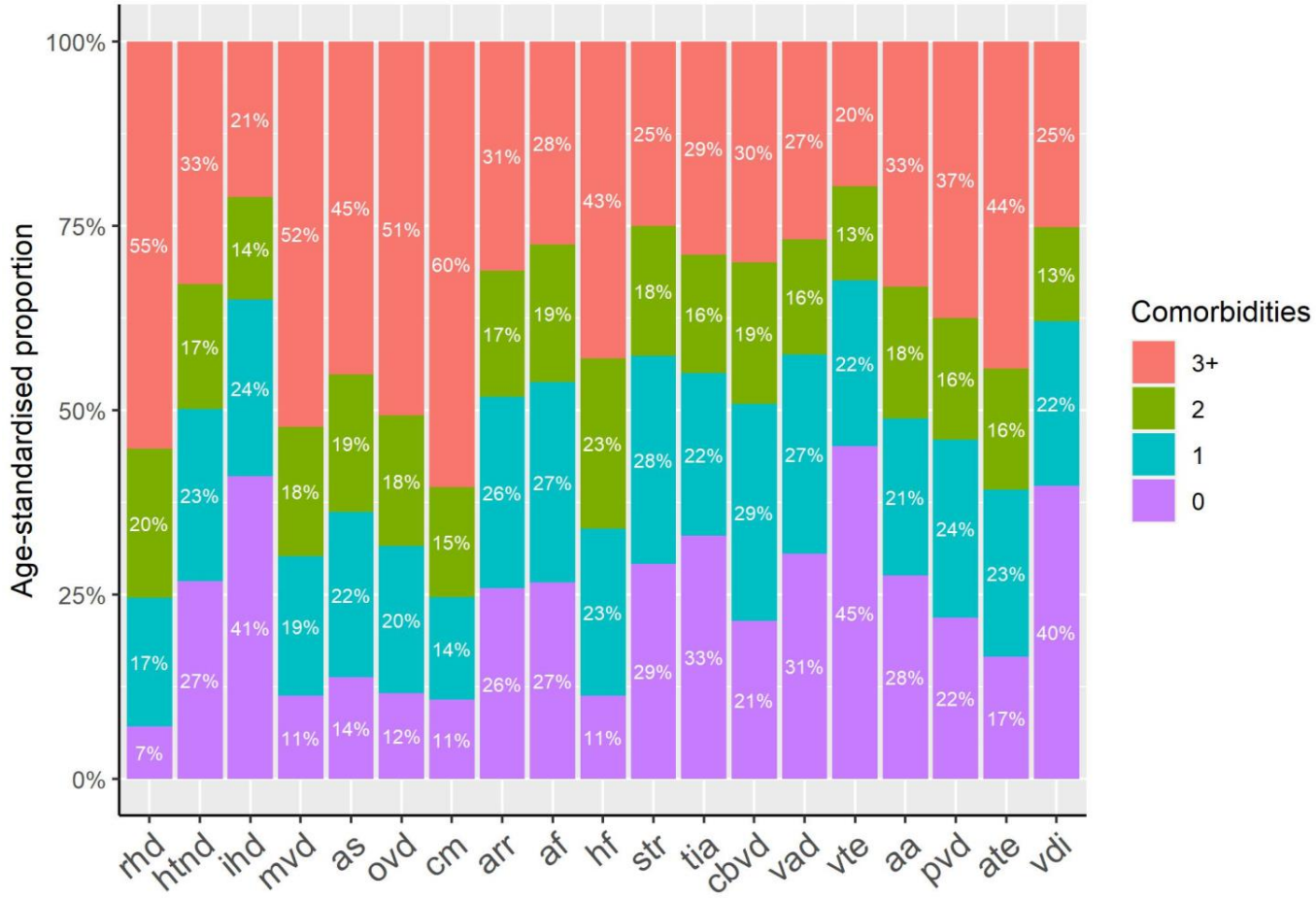


Figure 4.5 Age-standardised proportion of the number of cardiovascular comorbidities by each incident individual cardiovascular disease

rhd=rheumatic heart disease, htnd=hypertensive heart and renal disease, ihd=ischaemic heart disease, mvd=mitral valve disorders, as=aortic stenosis, ovd=other aortic valve disorders, cm=cardiomyopathy, arr=other arrhythmias, af=atrial fibrillation, hf=heart failure, str=stroke, tia=transient cerebral ischaemic attacks, cbvd=other cerebrovascular disease, vad= vascular dementia, vte=venous thromboembolism, aa=aortic aneurysm, pvd=peripheral vascular disease, ate=arterial embolism and thrombosis, vdi=vascular disorders of the intestine

4.3.6 The five most common cardiovascular comorbidities of individual cardiovascular diseases

The most common heart diseases (heart failure, ischaemic heart disease, atrial fibrillation, and other arrhythmias) were also among the top five comorbidities of almost all of the individual CVDs (Appendix C, page 303). Of other CVDs, rheumatic heart disease and other cerebrovascular disease were also relatively common cardiovascular comorbidities.

4.3.7 Cardiovascular multimorbidity over time

4.3.7.1 Cardiovascular multimorbidity and cardiovascular diseases in all women

Women who had not been diagnosed with CVD at age 50 years and lived to at least age 80 years had a 19% probability of developing CVM during follow-up (Table 4.4). The probability of having incident CVM approximately doubled every five years of age from 60 to 80 years. The chance of having incident complex CVM was much lower, at 5% by age 80 years.

Ischaemic heart disease, the most common and one of the earliest onset CVDs, accumulated most quickly in the cohort; the probability of having incident ischaemic heart disease was 3% by age 60 years. For all other diseases, the probability of an incident event was under 1% until at least age 65 years. From age 65 to 80 years, the probability of having incident atrial fibrillation increased by a factor of about 2-3 with every five years. The chance of having other arrhythmias nearly doubled, and stroke increased by a factor of 1.5-2 every five years between age 65 to 80 years.

Table 4.4 Kaplan-Meier failure estimates for cardiovascular multimorbidity and individual cardiovascular diseases at ages 60, 65, 70, 75, and 80 years in all women

Age, years	60 Failure per 100 (95% CI)	65 Failure per 100 (95% CI)	70 Failure per 100 (95% CI)	75 Failure per 100 (95% CI)	80 Failure per 100 (95% CI)
Number of diseases					
Any cardiovascular multimorbidity (CVM)	0.9 (0.9-0.9)	2.2 (2.2-2.2)	4.9 (4.9-5.0)	10.0 (10.0-10.1)	19.1 (19.0-19.2)
Complex CVM	0.1 (0.1-0.1)	0.2 (0.2-0.2)	0.7 (0.7-0.7)	2.0 (2.0-2.0)	5.0 (4.9-5.1)
Heart diseases					
Rheumatic heart disease	0.1 (0.1-0.1)	0.2 (0.2-0.2)	0.6 (0.6-0.6)	1.4 (1.4-1.4)	3.2 (3.1-3.2)
Hypertensive heart and renal disease	0.1 (0.1-0.1)	0.3 (0.2-0.3)	0.5 (0.5-0.6)	1.2 (1.1-1.2)	1.8 (1.8-1.9)
Ischaemic heart disease	2.7 (2.2-3.3)	4.9 (4.4-5.5)	8.1 (7.6-8.6)	12.4 (12.0-12.9)	18.3 (17.9-18.8)
Mitral valve disorders	0.1 (0.1-0.1)	0.3 (0.2-0.3)	0.6 (0.6-0.6)	1.1 (1.1-1.2)	2.2 (2.1-2.2)
Aortic stenosis	0.1 (0.1-0.1)	0.2 (0.2-0.2)	0.4 (0.4-0.4)	1.0 (1.0-1.0)	2.2 (2.1-2.2)
Other aortic valve disorders	0.0 (0.0-0.1)	0.1 (0.1-0.1)	0.2 (0.2-0.2)	0.5 (0.4-0.5)	0.9 (0.9-0.9)
Cardiomyopathy	0.1 (0.1-0.1)	0.1 (0.1-0.1)	0.2 (0.2-0.2)	0.4 (0.4-0.4)	0.7 (0.6-0.7)
Other arrhythmias	0.6 (0.6-0.7)	1.2 (1.2-1.3)	2.3 (2.2-2.4)	4.2 (4.1-4.3)	7.6 (7.5-7.7)
Atrial fibrillation	0.5 (0.5-0.6)	1.4 (1.4-1.4)	3.2 (3.2-3.2)	6.7 (6.6-6.7)	13.0 (12.9-13.1)
Heart failure	0.3 (0.3-0.3)	0.7 (0.7-0.7)	1.5 (1.5-1.5)	3.2 (3.2-3.2)	6.7 (6.6-6.8)
Cerebrovascular diseases					
Stroke	0.6 (0.6-0.6)	1.1 (1.0-1.1)	1.9 (1.8-1.9)	3.3 (3.2-3.3)	5.7 (5.6-5.8)
Transient cerebral ischaemic attacks	0.2 (0.2-0.2)	0.4 (0.4-0.4)	0.7 (0.7-0.7)	1.2 (1.2-1.2)	2.0 (2.0-2.1)
Other cerebrovascular disease	0.3 (0.3-0.3)	0.6 (0.6-0.6)	1.3 (1.2-1.3)	2.7 (2.6-2.7)	5.8 (5.8-5.9)
Vascular dementia	0.0 (0.0-0.0)	0.0 (0.0-0.0)	0.1 (0.1-0.1)	0.4 (0.3-0.4)	1.3 (1.3-1.4)
Other vascular diseases					
Venous thromboembolism	0.8 (0.8-0.9)	1.4 (1.4-1.4)	2.2 (2.1-2.2)	3.3 (3.3-3.4)	5.0 (4.9-5.1)
Aortic aneurysm	0.0 (0.0-0.0)	0.1 (0.1-0.1)	0.2 (0.2-0.2)	0.5 (0.5-0.5)	1.1 (1.1-1.1)
Peripheral vascular disease	0.2 (0.2-0.2)	0.4 (0.4-0.4)	0.7 (0.7-0.8)	1.3 (1.3-1.3)	2.2 (2.1-2.2)
Arterial embolism and thrombosis	0.1 (0.1-0.1)	0.1 (0.1-0.1)	0.2 (0.2-0.2)	0.4 (0.4-0.4)	0.7 (0.6-0.7)
Vascular disorders of the intestine	0.1 (0.1-0.1)	0.2 (0.2-0.2)	0.3 (0.3-0.4)	0.6 (0.6-0.7)	1.1 (1.1-1.1)

4.3.7.2 Cardiovascular diseases in women who developed cardiovascular multimorbidity

In the subset of women who eventually developed CVM by the end of follow-up, the probability of having incident ischaemic heart disease was 23% by age 60 years, rapidly increasing to 71% by age 80 years (Table 4.5). The chance of having incident atrial fibrillation and other arrhythmias was over 10% by age 65 years and increased by a factor of about 1.5 every 5 years until age 80 years. The chance of having incident stroke and venous thromboembolism by age 65 years was 9% and 8%, respectively, and increased by a factor of about 1.3-1.5 every five years until age 80 years.

The subset of women who eventually developed complex CVM had similar patterns of occurrence of cerebrovascular and other vascular diseases but were more likely to have heart diseases – particularly heart failure, heart valve diseases, and cardiomyopathy – at an earlier age than women who developed any CVM (Table 4.6).

Table 4.5 Kaplan-Meier failure estimates for individual cardiovascular diseases in women who had a record of cardiovascular multimorbidity during follow-up

Age, years	60 Failure per 100 (95% CI)	65 Failure per 100 (95% CI)	70 Failure per 100 (95% CI)	75 Failure per 100 (95% CI)	80 Failure per 100 (95% CI)
Heart diseases					
Rheumatic heart disease	1.2 (1.1-1.4)	2.7 (2.6-2.9)	6.0 (5.9-6.2)	11.5 (11.2-11.7)	20.5 (20.2-20.8)
Hypertensive heart and renal disease	1.0 (0.8-1.1)	2.3 (2.1-2.4)	4.1 (4.0-4.3)	7.4 (7.2-7.6)	10.4 (10.1-10.6)
Ischaemic heart disease	22.7 (13.1-37.8)	32.7 (23.4-44.5)	45.0 (36.9-54.1)	58.2 (51.7-64.8)	71.4 (66.8-75.8)
Mitral valve disorders	1.6 (1.4-1.7)	3.0 (2.9-3.2)	5.7 (5.5-5.9)	9.4 (9.2-9.6)	14.6 (14.3-14.9)
Aortic stenosis	1.0 (0.7-1.2)	2.0 (1.8-2.3)	4.2 (3.9-4.5)	7.7 (7.5-8.0)	13.4 (13.1-13.8)
Other aortic valve disorders	0.6 (0.5-0.8)	1.3 (1.1-1.4)	2.3 (2.2-2.5)	3.8 (3.7-4.0)	6.1 (5.9-6.3)
Cardiomyopathy	0.8 (0.7-1.0)	1.5 (1.3-1.6)	2.5 (2.4-2.6)	3.7 (3.5-3.8)	5.0 (4.8-5.2)
Other arrhythmias	6.7 (5.5-8.2)	11.3 (10.1-12.6)	18.1 (17.0-19.3)	27.4 (26.3-28.4)	40.0 (39.0-40.9)
Atrial fibrillation	6.1 (5.8-6.4)	12.6 (12.3-13.0)	23.8 (23.5-24.2)	39.8 (39.4-40.1)	59.6 (59.2-60.0)
Heart failure	4.4 (4.1-4.7)	8.4 (8.1-8.8)	15.2 (14.9-15.6)	25.3 (25.0-25.7)	40.2 (39.8-40.6)
Cerebrovascular diseases					
Stroke	5.3 (4.8-5.8)	8.7 (8.3-9.2)	13.7 (13.2-14.1)	20.7 (20.2-21.1)	30.1 (29.6-30.6)
Transient cerebral ischaemic attacks	1.7 (1.5-1.9)	3.0 (2.8-3.2)	4.8 (4.6-5.1)	7.4 (7.2-7.6)	10.9 (10.6-11.2)
Other cerebrovascular disease	3.6 (3.3-3.9)	6.3 (6.0-6.7)	11.2 (10.9-11.6)	19.0 (18.6-19.3)	31.7 (31.2-32.1)
Vascular dementia	0.0 (0.0-0.1)	0.1 (0.1-0.2)	0.6 (0.6-0.7)	2.0 (1.9-2.1)	6.2 (6.0-6.4)
Other vascular diseases					
Venous thromboembolism	5.0 (4.6-5.5)	7.6 (7.2-8.1)	10.9 (10.5-11.4)	15.3 (14.9-15.8)	20.5 (20.0-21.0)
Aortic aneurysm	0.3 (0.2-0.4)	0.7 (0.6-0.8)	1.6 (1.5-1.8)	3.3 (3.2-3.4)	5.6 (5.4-5.8)
Peripheral vascular disease	2.5 (2.2-2.7)	4.2 (4.0-4.5)	6.6 (6.3-6.8)	9.7 (9.4-10.0)	13.4 (13.1-13.7)
Arterial embolism and thrombosis	0.9 (0.8-1.0)	1.5 (1.4-1.6)	2.3 (2.2-2.5)	3.5 (3.3-3.6)	4.8 (4.6-5.0)
Vascular disorders of the intestine	0.5 (0.4-0.6)	1.0 (0.9-1.1)	1.8 (1.7-1.9)	3.0 (2.9-3.2)	4.8 (4.6-5.0)

Table 4.6 Kaplan-Meier failure estimates for individual cardiovascular diseases in women who had a record of complex cardiovascular multimorbidity during follow-up

Age, years	60 Failure per 100 (95% CI)	65 Failure per 100 (95% CI)	70 Failure per 100 (95% CI)	75 Failure per 100 (95% CI)	80 Failure per 100 (95% CI)
Heart diseases					
Rheumatic heart disease	3.9 (3.3-4.5)	7.4 (6.8-8.0)	14.2 (13.6-14.9)	25.8 (25.1-26.5)	43.0 (42.2-43.8)
Hypertensive heart and renal disease	2.2 (1.7-2.8)	4.6 (4.1-5.2)	8.0 (7.4-8.6)	13.7 (13.1-14.4)	19.2 (18.5-19.9)
Ischaemic heart disease	21.3 (18.9-24.1)	33.8 (31.6-36.1)	48.9 (47.2-50.7)	65.5 (64.3-66.8)	81.4 (80.6-82.1)
Mitral valve disorders	3.9 (3.3-4.5)	7.1 (6.5-7.8)	12.5 (11.9-13.2)	20.1 (19.4-20.8)	29.8 (29.0-30.6)
Aortic stenosis	1.9 (1.5-2.3)	4.0 (3.5-4.5)	8.0 (7.5-8.5)	14.7 (14.1-15.2)	25.3 (24.6-26.0)
Other aortic valve disorders	1.4 (1.0-1.9)	3.0 (2.5-3.5)	5.3 (4.8-5.8)	8.4 (7.9-8.9)	12.8 (12.2-13.5)
Cardiomyopathy	2.4 (1.9-3.0)	3.9 (3.4-4.5)	6.3 (5.7-6.9)	8.7 (8.1-9.3)	11.7 (11.0-12.3)
Other arrhythmias	12.1 (6.6-21.4)	17.8 (12.2-25.7)	26.9 (21.5-33.4)	39.9 (35.2-45.0)	57.2 (53.7-60.7)
Atrial fibrillation	9.9 (9.0-10.8)	19.4 (18.5-20.3)	34.4 (33.6-35.3)	55.0 (54.2-55.7)	77.0 (76.4-77.6)
Heart failure	8.6 (7.8-9.5)	16.1 (15.3-17.0)	27.9 (27.1-28.7)	44.8 (44.0-45.5)	66.2 (65.5-66.9)
Cerebrovascular diseases					
Stroke	6.2 (5.4-7.2)	10.6 (9.8-11.6)	17.1 (16.2-18.0)	26.6 (25.7-27.5)	39.1 (38.2-40.0)
Transient cerebral ischaemic attacks	2.3 (1.9-2.7)	4.1 (3.7-4.6)	6.7 (6.2-7.2)	10.3 (9.8-10.9)	15.7 (15.0-16.3)
Other cerebrovascular disease	4.3 (3.8-5.0)	8.0 (7.4-8.7)	14.8 (14.1-15.4)	25.8 (25.1-26.5)	43.0 (42.3-43.8)
Vascular dementia	0.0 (0.0-0.1)	0.1 (0.1-0.2)	0.8 (0.7-1.0)	2.9 (2.6-3.1)	9.2 (8.7-9.7)
Other vascular diseases					
Venous thromboembolism	6.0 (5.1-7.0)	9.0 (8.1-10.0)	13.0 (12.1-14.0)	18.4 (17.5-19.4)	25.1 (24.2-26.1)
Aortic aneurysm	0.5 (0.3-0.7)	1.3 (1.1-1.6)	3.2 (2.9-3.4)	6.2 (5.8-6.5)	10.3 (9.8-10.8)
Peripheral vascular disease	4.3 (3.7-5.0)	7.1 (6.5-7.8)	11.2 (10.5-11.9)	16.5 (15.8-17.3)	23.0 (22.3-23.8)
Arterial embolism and thrombosis	1.4 (1.1-1.8)	2.5 (2.2-2.9)	4.1 (3.8-4.5)	6.3 (5.9-6.7)	8.7 (8.2-9.2)
Vascular disorders of the intestine	0.5 (0.3-0.8)	1.1 (0.9-1.3)	2.2 (1.9-2.4)	4.0 (3.7-4.4)	6.5 (6.1-6.9)

4.3.8 Coding of rheumatic heart disease

There were 17,348 incident cases of rheumatic heart disease recorded during follow-up; of the 114,684 women with CVM by the end of follow-up, 16,014 (14%) had rheumatic heart disease (Table 4.2, page 105). The five most common cardiovascular comorbidities of rheumatic heart disease were, in descending order, ischaemic heart disease, atrial fibrillation, heart failure, other arrhythmias, and mitral valve disorders (Figure 4.6). More than half of all women with rheumatic heart disease also had a record of incident ischaemic heart disease or atrial fibrillation during the same period of follow-up. Heart failure was recorded in about 40%, and other arrhythmias in about 30% of all women with rheumatic heart disease. Other mitral valve disorders were recorded in about a fifth of these women.

The most common form of CVM with rheumatic heart disease was its pairing with ischaemic heart disease, found in 11% (n=1,925) of all rheumatic heart disease cases. Notably, the third most common CVM combination among women with rheumatic heart disease included four CVDs (rheumatic heart disease with ischaemic heart disease, atrial fibrillation, and heart failure), which was more common than most combinations of two or three CVDs among women with rheumatic heart disease (Figure 4.6).

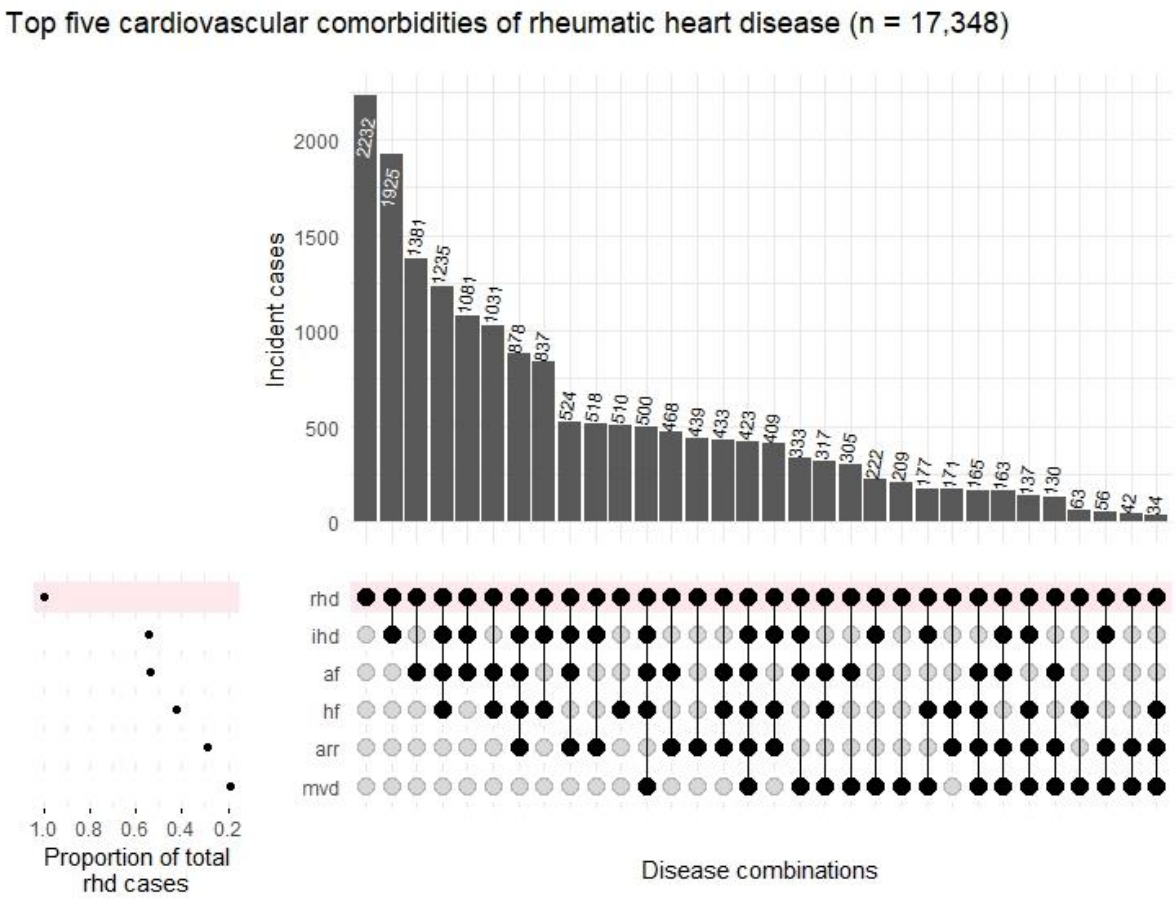


Figure 4.6 UpSet plot of the top five comorbidities of rheumatic heart disease (rhd)

af=atrial fibrillation, arr=other arrythmias, hf=heart failure, ihd=ischaemic heart disease, mvd=mitral valve disorders, rhd=rheumatic heart disease

In the Million Women Study, the most common 4-digit ICD-10 codes for a first rheumatic heart disease event in hospital admission data were, in descending order, I08.1 (disorders of both mitral and tricuspid valves), I08.0 (disorders of both mitral and aortic valves), I08.3 (combined disorders of all three valves), and I07.1 (tricuspid insufficiency) (Figure 4.7). These four codes accounted for over 80% of all recorded cases of rheumatic heart disease. The vast majority of cases (over 78% for each four-character ICD-10 code) were recorded as a secondary diagnosis. For both I08 (multiple valve diseases) and I07 (rheumatic tricuspid valve diseases), ICD-10 coding directives included diseases specified as rheumatic or of unspecified origin, and excluded diseases specified as nonrheumatic.

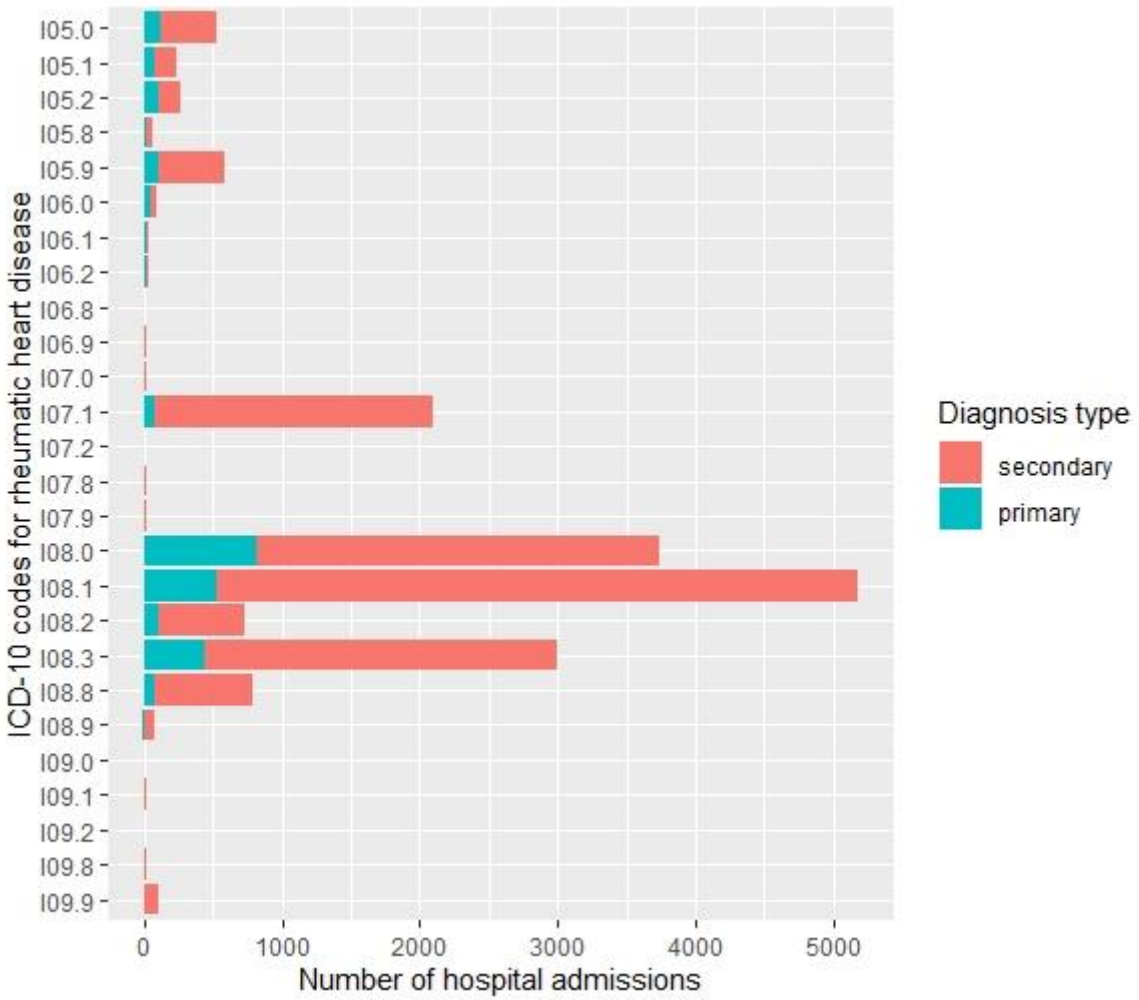


Figure 4.7 Distribution of 4-character ICD-10 codes for incident rheumatic heart disease in hospital admission data in the Million Women Study

There were 669 cases of incident mitral valve stenosis (I05.0) – the code that is thought to reliably represent rheumatic heart disease [20, 21] –in the Million Women Study (Figure 4.8). The five most common cardiovascular comorbidities of I05.0 were atrial fibrillation, ischaemic heart disease, heart failure, mitral valve disease, and other arrhythmias. Nearly 70% of all women who had I05.0 also had atrial fibrillation, and over 50% had ischaemic heart disease. Heart failure and mitral valve disease were also recorded in approximately 40% of women with I05.0, and more than 20% also had other arrhythmias.

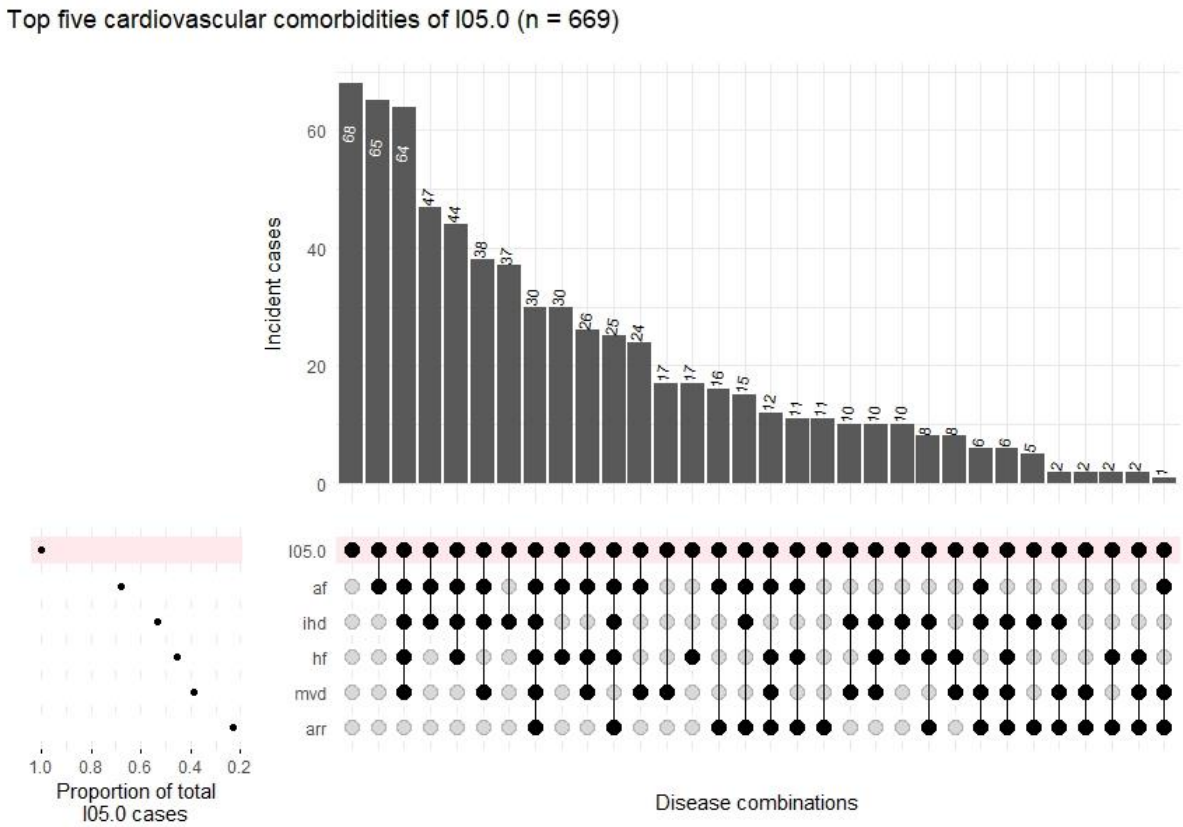


Figure 4.8 UpSet plot of the top five comorbidities of mitral valve stenosis (I05.0)

af=atrial fibrillation, arr=other arrythmias, hf=heart failure, ihd=ischaemic heart disease, mvd=mitral valve disorders

Defining CVM based on the co-occurrence of the 18 individual CVDs other than rheumatic heart disease (i.e. leaving out rheumatic heart disease) decreased the total number of CVM cases by 2.8% from 114,684 to 111,483, and the number of complex CVM by 13.1% from 26,896 to 23,381 (Table 4.7). Nevertheless, leaving out rheumatic heart disease did not substantially alter the composition or proportion of the CVDs comprising CVM or complex CVM, compared to before (see Table 4.2, page 105).

Table 4.7 Number and proportion of cardiovascular diseases that comprise CVM at the end of follow-up, after leaving out rheumatic heart disease from the definition of CVM

Disease	In women with CVM [n=111,483], % (n)	In women with complex CVM [n=23,381], % (n)
Heart diseases		
Hypertensive heart and renal disease	7.0 (7,843)	14.7 (3,437)
Ischaemic heart disease	55.5 (61,850)	74.9 (17,520)
Mitral valve disorders	9.8 (10,923)	22.0 (5,136)
Aortic stenosis	9.0 (10,076)	18.8 (4,391)
Other aortic valve disorders	4.0 (4,447)	9.5 (2,219)
Cardiomyopathy	3.2 (3,604)	8.1 (1,904)
Other arrhythmias	28.5 (31,823)	46.1 (10,780)
Atrial fibrillation	46.9 (52,324)	70.2 (16,421)
Heart failure	30.2 (33,640)	60.8 (14,208)
Cerebrovascular diseases		
Stroke	21.1 (23,504)	32.2 (7,539)
Transient cerebral ischaemic attacks	7.1 (7,958)	11.8 (2,753)
Other cerebrovascular disease	22.8 (25,463)	38.0 (8,885)
Vascular dementia	4.3 (4,801)	7.8 (1,831)
Other vascular diseases		
Venous thromboembolism	12.9 (14,363)	17.6 (4,122)
Aortic aneurysm	3.8 (4,181)	7.8 (1,820)
Peripheral vascular disease	8.8 (9,842)	17.6 (4,106)
Arterial embolism and thrombosis	3.0 (3,389)	6.5 (1,514)
Vascular disorders of the intestine	3.2 (3,514)	5.1 (1,183)

CVM=cardiovascular multimorbidity

4.4 Discussion

4.4.1 Key findings and interpretation of results

4.4.1.1 Overall summary

Among the 1.2 million healthy middle-aged UK women who were followed up in this study for an average of 16 years, 9% developed two or more CVDs (i.e. CVM) by the end of follow-up. The probability of having CVM by a certain age doubled every five years between age 60 and 80 years. CVM was recorded in over 40% of women who had developed any individual CVD by age 73 years (mean age at the end of follow-up). The four most common heart diseases – namely ischaemic heart disease (55% of CVM), atrial fibrillation (46%), other arrhythmias (28%), and heart failure (30%) – each contributed to more than a fourth to half of all cases of CVM. They often occurred together as comorbidities of one another. Stroke was found in approximately 1 in 5 cases of CVM. Rheumatic heart disease was particularly common in women with complex CVM (found in 1 in 3 cases). The proportion of cases occurring with cardiovascular comorbidities varied across individual CVDs – the lowest was around 50% for venous thromboembolism and highest was around 90% for rheumatic heart disease. Conditions which were most frequently found as isolated diagnoses (i.e. without CVM), tended to have a younger average age at first diagnosis, although this finding remained after age standardisation. In general, CVDs within the same disease class occurred together more frequently and were associated more strongly with each other than with diseases of different classes.

4.4.1.2 Findings in context

Few previous studies have described the prevalence and components of co-occurring CVDs [1, 2], and they investigated a limited selection of cardiovascular comorbidities of CVDs rather than a comprehensive range of CVDs as in this chapter. Analyses in US Medicare beneficiaries aged ≥65 years in 2012 investigated ischaemic heart disease, heart

failure, atrial fibrillation, and stroke as comorbid conditions of one another. [1] The study did not investigate any other CVDs. The results were consistent with the findings of this chapter in that the selected CVDs were common comorbidities of each other. For example, among 2.6 million patients with a claim indicating they received service or treatment for atrial fibrillation in 2012, 64% also had a claim for ischaemic heart disease and 51% for heart failure. [1] In this chapter, among 76,058 women with a diagnosis of incident atrial fibrillation during follow-up, 34% also had a record of incident ischaemic heart disease while 24% also had a record of heart failure. The prevalence estimates were generally greater in the US study than in this chapter, partly because this chapter did not include women with CVD prior to recruitment and will have missed CVD diagnoses that were not recorded in the linked hospital admissions and death data. In contrast, the study in this chapter was novel in its focus on CVM (rather than on comorbidities of a few CVDs) and the breadth and depth of the CVD subtypes investigated.

4.4.1.3 Mechanisms underlying cardiovascular multimorbidity

CVDs share risk factors and pathophysiology, and may cause one another (e.g. atrial fibrillation leading to cardioembolic stroke). [2, 22-24] Older age, higher body mass index, higher blood cholesterol, treatment for hypertension and diabetes, and smoking, are examples of well-established risk factors shared by the majority of CVDs. [25-27]

'Ischaemic heart disease & Atrial fibrillation' was the most common CVM pair (incidence 21.2 per 1000 persons over the duration of follow-up) in this cohort, and they were also the two most common individual CVDs. Previous studies reported that atrial fibrillation increases the risk of myocardial infarction. [28-31] In particular, the REGARDS prospective cohort study [28] found that having atrial fibrillation nearly doubled the risk of myocardial infarction among participants with no prior coronary heart disease, after adjustment for various cardiovascular risk factors. Furthermore, the Cardiovascular Health

Study [30] reported that having atrial fibrillation was associated with a higher risk of coronary heart disease in older women (HR [95% CI] 2.3 [1.8-2.8]) than in older men (1.6 [1.3-2.0]).

Similarly, atrial fibrillation is an independent risk factor for heart failure [32, 33], stroke [34, 35], and thromboembolism [36-38], and a marker for worse outcomes in heart failure and ischaemic stroke. [34, 39, 40] Heart failure increases short-term and long-term risk of all stroke subtypes [41], and can trigger atrial fibrillation [42]. One of the most common causes of heart failure is ischaemic heart disease. [43] Ischaemic heart disease can increase the risk of stroke, arrhythmias, and other CVDs. [44-47] As such, the frequent co-occurrence and strong associations among these common CVDs were as expected.

Strong associations observed between less common CVDs were also plausible. The most strongly associated CVM pair was between arterial thromboembolism (I74) and peripheral vascular diseases (I73.9: unspecified peripheral vascular disease including intermittent claudication and spasm of artery), which occurred 45 times more often than expected by chance; these conditions relate to largely the same area of the body and the disease process of atherosclerosis. [48-50] Aortic aneurysm and other aortic valve disease occurred together 26 times more often than expected; this may partly be because aortic regurgitation is often caused by aortic root or ascending thoracic aortic aneurysms and other structural abnormalities near the aortic valve. [51, 52]

In the Million Women Study, some CVD subtypes occurred with CVM at a much higher proportion than others. Women with rheumatic heart disease, heart valve diseases, heart failure, and cardiomyopathy had records of other CVDs (i.e. CVM) around 85% or more of the time. Rheumatic heart disease, heart valve diseases, and cardiomyopathy can lead to heart failure [43], and rheumatic heart disease is often found with non-rheumatic heart

valve disease [53]. The most common comorbidities of these CVDs were the common heart diseases mentioned above. In general, it appears that CVD subtypes with higher proportions of CVM tended to occur later in life because of having had other heart diseases and cardiovascular risk factors for a prolonged period of time.

Ischaemic heart disease, venous thromboembolism, and vascular diseases of the intestines only occurred with CVM around 50% of the time. Ischaemic heart disease may be both common in CVM and commonly found isolated for the same reason that many people are susceptible. Although venous thromboembolism and vascular disorders of the intestines share cardiovascular risk factors, they also have specific risk factors that are less strongly implicated in the development of other CVDs. For instance, major surgery, hospitalisation, active cancer, and hormonal factors (e.g. oral contraceptives, hormone therapy, pregnancy) are strong independent risk factors for incident venous thromboembolism. [54] Comorbidities (e.g. irritable bowel syndrome), use of oral contraceptives, and bacterial infections are associated with vascular ischaemia. [55, 56]

The CVD subtypes investigated in this thesis were chosen to represent distinct conditions wherever possible (see Section 3.3.1 of Chapter 3, page 79). Although some CVD subtypes may indeed represent (whether in whole or in part) common underlying pathophysiology such as atherosclerosis, the varied associations among the 19 selected CVD subtypes in this chapter supported that CVM as a whole is not a single disease (Figure 4.2, page 109). Moreover, previous studies have found different risk factor associations across different CVD subtypes [54-57], demonstrating that there are good reasons to further investigate risk factors for co-occurring CVDs (i.e. CVM).

4.4.1.4 Rheumatic heart disease in the UK and the Million Women Study

Rheumatic heart disease is a sequela of acute rheumatic fever from a group A streptococcal infection. [58] Acute rheumatic fever is a disease associated with poverty and poor sanitation. [59] Although rheumatic heart disease is believed to be largely eradicated in high income countries like the UK [19], it was not uncommon in this cohort of women (aged 77 years on average at time of writing). It was also associated with a particularly high proportion of multimorbidity (found in 1 in 3 cases of complex CVM), often recorded as occurring with multiple heart diseases and other mitral valve disease. Given that the women in this cohort were born between 1935-1950, some may have had acute rheumatic fever before socioeconomic improvements, widespread use of antibiotics, and the use of secondary prophylaxis treatment in the 1960s. [60]

It is also possible that incident rheumatic heart disease was overestimated in this study because hospital admission records identified past medical history (particularly given the high proportion of secondary diagnoses), and because nonrheumatic degenerative heart valve disorders were misclassified as rheumatic heart disease. A study in Australia [21] found that due to ICD-10 coding directives, rheumatic heart disease was vastly overestimated in hospital admission records – just 31% of those classified to have rheumatic heart disease had evidence of the condition in clinical records. The false positive cases were in fact tricuspid, multivalvular, and mitral valve disease with uncertain aetiology. In the Million Women Study, just four four-digit ICD-10 codes identified 80% of all first rheumatic heart disease cases: I08.1 (disorders of both mitral and tricuspid valves), I08.0 (disorders of both mitral and aortic valves), I08.3 (combined disorders of all three valves), and I07.1 (tricuspid insufficiency); this was a similar pattern to what was observed in the Australian study. Therefore, there may be reason to consider rheumatic heart disease (I05-I09) as largely representing multiple heart valve disease in the Million

Women Study. Nevertheless, ignoring rheumatic heart disease from the definition of CVM did not substantially alter the proportions of individual CVDs which comprised CVM in this chapter.

Further research is needed to understand the true distribution of rheumatic heart disease in UK women and its role in CVM. Even in developed countries where rheumatic heart disease is believed to be largely eradicated, some ethnic and socioeconomic groups suffer from higher rates of this disease (e.g. people living in poverty, immigrants from high risk countries). [60] Furthermore, women are more likely than men to develop rheumatic heart disease [61], and can be at risk of pregnancy related complications of rheumatic mitral valve disease [62]. Currently, 25% of women (50% in London) who give birth in the UK were not themselves born in the UK. [63] If some of these women were born and raised in countries where rheumatic heart disease is endemic, they may be at a higher risk of serious complications in pregnancy. [61] Therefore, there is an ongoing need for clinician training and public awareness of rheumatic fever and rheumatic heart disease in the UK, especially in high-risk subpopulations.

4.4.2 Strengths and limitations

As described in Chapter 3, the large size of the Million Women Study is a strength of this analysis. There were more than 4,000 incident cases of each of the 19 selected CVDs, which allowed for a comprehensive and detailed study of the incidence of CVM, complex CVM, and specific CVD pairs. Most cardiovascular research so far has been focussed on predominantly male populations and on a limited number of common CVDs like ischaemic heart disease and stroke. However, the incidence and prevalence estimates of CVM from these analyses are likely to be lower than the true population values. The participants of the Million Women Study are middle-aged and older women in the UK

who attended breast cancer screening, who tended to be healthier and more affluent than the average population. [64]

Some common CVD outcomes in hospital admission data in the Million Women Study have been compared with general practice records in a previous study. [65] Overall, 93% of women with a hospital record of coronary heart disease (I20-I25), cerebrovascular disease (G45, I60-I69), or venous thromboembolism (I26, I80-I82) had the diagnosis confirmed from their general practice records. Moreover, 97% of women with no hospital record of these diagnoses also did not have a general practice record of the diagnoses. This would represent moderate sensitivity in the Million Women Study, since from this it is possible to estimate that in 3% of 1.3 million women (39,000 women), hospital records may not have captured CVD diagnoses found in their primary care records. Similar validation studies have not been conducted for the other selected CVDs studied in this chapter, and there is some evidence that less common or less serious conditions may be less reliably coded in hospital admission data. [21] Finally, some records in secondary diagnosis fields are likely to represent past medical history rather than incident disease, particularly for less severe or chronic conditions compared to more severe or acute conditions.

4.4.3 Conclusions

The components of CVM in UK women had not been investigated in detail until now. Previous studies have often focussed on individual diagnoses and on a few selected conditions. The Million Women Study is a large prospective study linked to electronic health records that provided an important insight into how CVDs of the heart, head, and peripheral circulations occur together in middle-aged and older women. In this cohort, CVM was almost as common as ischaemic heart disease, which was by far the most common type of individual CVD. The risk of developing CVM nearly doubled every five

years between age 60 and 80 years, in women who had no CVDs at age 50 years. It is important for clinicians to keep in mind that each CVD is not an isolated disease, and the majority of people with any type of CVD will go on to develop at least one other CVD. A better understanding of which CVDs are most strongly associated with each other, and which risk factors are shared across or are specific to different CVDs, will help develop more personalised secondary prevention measures for patients presenting with particular types of CVM.

4.5 References

[1] Arnett DK, Goodman RA, Halperin JL, Anderson JL, Parekh AK, Zoghbi WA. AHA/ACC/HHS Strategies to Enhance Application of Clinical Practice Guidelines in Patients With Cardiovascular Disease and Comorbid Conditions: From the American Heart Association, American College of Cardiology, and U.S. Department of Health and Human Services. Journal of the American College of Cardiology. 2014 Oct;64(17):1851–1856. Available from: https://doi.org/10.1016/j.jacc.2014.07.012.

[2] Tran J, Norton R, Conrad N, Rahimian F, Canoy D, Nazarzadeh M, et al. Patterns and temporal trends of comorbidity among adult patients with incident cardiovascular disease in the UK between 2000 and 2014: A population-based cohort study. PLoS Med. 2018 Mar;15(3):e1002513. Available from: https://doi.org/10.1371/journal.pmed.1002513.

[3] The Million Women Study Collaborative Group. The Million Women Study: design and characteristics of the study population. Breast Cancer Res. 1999 Dec;1:73. Available from: https://doi.org/10.1186/bcr16.

[4] Green J, Reeves GK, Floud S, Barnes I, Cairns BJ, Gathani T, et al. Cohort Profile: the Million Women Study. Int J Epidemiol. 2019 Feb;48(1):28–29e. Available from: https://doi.org/10.1093/ije/dyy065.

[5] NHS Digital. Hospital Episode Statistics (HES) [Internet]; 2020. [cited 2021 Jan 13]. Available from: https://digital.nhs.uk/data-and-information/data-tools-and-services/-data-services/hospital-episode-statistics.

[6] Kendrick SW, Clarke JW. The Scottish Record Linkage System. Health bulletin. 1993 Mar;51(2):72–9.

[7] World Health Organization. ICD-10: international statistical classification of diseases and related health problems : tenth revision, 2nd ed. World Health Organization; 2004. Available from: https://apps.who.int/iris/handle/10665/42980.

[8] Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York; 2016. Available from: https://ggplot2.tidyverse.org.

[9] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2020. Available from: http://www.R-project.org/.

[10] Weisstein EW. Fisher's Exact Test [Internet];. [cited 2021 Jan 15]. Available from: https://mathworld.wolfram.com/FishersExactTest.html.

[11] Agresti A. A Survey of Exact Inference for Contingency Tables. Stat Sci. 1992 Feb;7(1):131–153.

[12] Fisher RA. On the Interpretation of $\chi 2$ from Contingency Tables, and the Calculation of P. J R Stat Soc. 1922 Jun;85(1):87–94. Available from: https://doi.org/-10.2307/2340521.

[13] Mehta CR, Patel NR, Tsiatis AA. Exact Significance Testing to Establish Treatment Equivalence with Ordered Categorical Data. Biometrics. 1984 Sep;40(3):819–825. Available from: https://doi.org/10.2307/2530927.

[14] Eurostat. Revision of the European Standard Population - Report of Eurostat's task force - 2013 edition. Luxembourg: Publications Office of the European Union; 2013. 121 p.

[15] Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: Visualization of Intersecting Sets. IEEE Trans Vis Comput Graph. 2014 Dec;20(12):1983–1992. Available from: https://doi.org/10.1109/TVCG.2014.2346248.

[16] Pagano M, Gauvreau K. Principles of biostatistics. 2nd ed. Boca Raton, FL: Chapman and Hall/CRC Press; 2018. 584 p.

[17] Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. J Am Stat Assoc. 1958;53(282):457–481. Available from: https://doi.org/10.2307/2281868.

[18] Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data. 2nd ed. John Wiley & Sons; 2002. 462 p.

[19] Gordis L. The virtual disappearance of rheumatic fever in the United States: lessons in the rise and fall of disease. T. Duckett Jones memorial lecture. Circulation. 1985 Dec;72:1155–62. Available from: https://doi.org/10.1161/01.cir.72.6.1155.

[20] Bonow RO, Carabello B, de Leon AC, Edmunds LH, Fedderly BJ, Freed MD, et al. Guidelines for the Management of Patients With Valvular Heart Disease. Circulation. 1998 Nov;98(18):1949–1984. Available from: https://doi.org/10.1161/01.CIR.98.18.1949.

[21] Katzenellenbogen JM, Nedkoff L, Cannon J, Kruger D, Pretty F, Carapetis JR, et al. Low positive predictive value of International Classification of Diseases, 10th Revision codes in relation to rheumatic heart disease: a challenge for global surveillance. Intern Med J. 2019 Mar;49(3):400–403. Available from: https://doi.org/10.1111/-imj.14221.

[22] Singh-Manoux A, Fayosse A, Sabia S, Tabak A, Shipley M, Dugravot A, et al. Clinical, socioeconomic, and behavioural factors at age 50 years and risk of cardiometabolic multimorbidity and mortality: A cohort study. PLoS Med. 2018 May;15(5):e1002571. Available from: https://doi.org/10.1371/journal.pmed.1002571.

[23] Kirchberger I, Meisinger C, Heier M, Zimmermann AK, Thorand B, Autenrieth CS, et al. Patterns of Multimorbidity in the Aged Population. Results from the KORA-Age Study. PLoS One. 2012 Jan;7(1):1–7. Available from: https://doi.org/10.1371/-journal.pone.0030556.

[24] Ashworth M, Durbaba S, Whitney D, Crompton J, Wright M, Dodhia H. Journey to multimorbidity: longitudinal analysis exploring cardiovascular risk factors and sociodemographic determinants in an urban setting. BMJ Open. 2019 Dec;9:e031649. Available from: http://dx.doi.org/10.1136/bmjopen-2019-031649.

[25] Kivimäki M, Kuosma E, Ferrie JE, Luukkonen R, Nyberg ST, Alfredsson L, et al. Overweight, obesity, and risk of cardiometabolic multimorbidity: pooled analysis of individual-level data for 120813 adults from 16 cohort studies from the USA and Europe. Lancet Public Health. 2017 Jun;2:e277–e285. Available from: https://doi.org/10.1016/-s2468-2667(17)30074-9.

[26] Yusuf S, Joseph P, Rangarajan S, Islam S, Mente A, Hystad P, et al. Modifiable risk factors, cardiovascular disease, and mortality in 155722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. Lancet. 2020 Mar;395:795–808. Available from: https://doi.org/10.1016/s0140-6736(19)32008-2.

[27] Emerging Risk Factors Collaboration, Sarwar N, Gao P, Seshasai SRK, Gobin R, Kaptoge S, et al. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet. 2010 Jun;375:2215–2222. Available from: https://doi.org/10.1016/s0140-6736(10)60484-9.

[28] Soliman EZ, Safford MM, Muntner P, Khodneva Y, Dawood FZ, Zakai NA, et al. Atrial fibrillation and the risk of myocardial infarction. JAMA Intern Med. 2014 Jan;174(1):107–114. Available from: https://doi.org/10.1001/jamainternmed.2013.11912.

[29] Chen LY, Sotoodehnia N, Bužková P, Lopez FL, Yee LM, Heckbert SR, et al. Atrial fibrillation and the risk of sudden cardiac death: the atherosclerosis risk in communities study and cardiovascular health study. JAMA Intern Med. 2013 Jan;173(1):29–35. Available from: https://doi.org/10.1001/2013.jamainternmed.744.

[30] O'Neal WT, Sangal K, Zhang ZM, Soliman EZ. Atrial fibrillation and incident myocardial infarction in the elderly. Clin Cardiol. 2014 Dec;37(12):750–755. Available from: https://dx.doi.org/10.1002/clc.22339.

[31] Soliman EZ, Lopez F, O'Neal WT, Chen LY, Bengtson L, Zhang ZM, et al. Atrial Fibrillation and Risk of ST-Segment-Elevation Versus Non-ST-Segment-Elevation Myocardial Infarction: The Atherosclerosis Risk in Communities (ARIC) Study. Circulation. 2015 May;131(21):1843–1850. Available from: https://doi.org/10.1161/-circulationaha.114.014145.

[32] Wang TJ, Larson MG, Levy D, Vasan RS, Leip EP, Wolf PA, et al. Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality: the Framingham Heart Study. Circulation. 2003 Jun;107:2920–2925. Available from: https://doi.org/10.1161/01.cir.0000072767.89944.6e.

[33] Miyasaka Y, Barnes ME, Gersh BJ, Cha SS, Bailey KR, Abhayaratna W, et al. Incidence and mortality risk of congestive heart failure in atrial fibrillation patients: a community-based study over two decades. Eur Heart J. 2006 Apr;27(8):936–941. Available from: https://doi.org/10.1093/eurheartj/ehi694.

[34] Bordignon S, Chiara Corti M, Bilato C. Atrial Fibrillation Associated with Heart Failure, Stroke and Mortality. J Atr Fibrillation. 2012 Jun;5(1):467. Available from: https://doi.org/10.4022/jafib.467.

[35] Bushnell C, McCullough LD, Awad IA, Chireau MV, Fedder WN, Furie KL, et al. Guidelines for the prevention of stroke in women: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2014 May;45(5):1545–1588. Available from: https://doi.org/10.1161/-01.str.0000442009.06663.48.

[36] Bekwelem W, Connolly SJ, Halperin JL, Adabag S, Duval S, Chrolavicius S, et al. Extracranial Systemic Embolic Events in Patients With Nonvalvular Atrial Fibrillation: Incidence, Risk Factors, and Outcomes. Circulation. 2015 Jul;132:796–803. Available from: https://doi.org/10.1161/circulationaha.114.013243.

[37] Frost L, Engholm G, Johnsen S, Møller H, Henneberg EW, Husted S. Incident thromboembolism in the aorta and the renal, mesenteric, pelvic, and extremity arteries after discharge from the hospital with a diagnosis of atrial fibrillation. Arch Intern Med. 2001 Jan;161(2):272–276. Available from: https://doi.org/10.1001/archinte.161.2.272.

[38] Andersen LV, Mortensen LS, Lip GYH, Lindholt JS, Faergeman O, Henneberg EW, et al. Atrial fibrillation and upper limb thromboembolectomy: a national cohort study. J Thromb Haemost. 2011 Sep;9(9):1738–1743. Available from: https://doi.org/10.1111/-j.1538-7836.2011.04435.x.

[39] Odutayo A, Wong CX, Williams R, Hunn B, Emdin CA. Prognostic Importance of Atrial Fibrillation Timing and Pattern in Adults With Congestive Heart Failure: A Systematic Review and Meta-Analysis. J Card Fail. 2017 Jan;23:56–62. Available from: https://doi.org/10.1016/j.cardfail.2016.08.005.

[40] Marini C, De Santis F, Sacco S, Russo T, Olivieri L, Totaro R, et al. Contribution of atrial fibrillation to incidence and outcome of ischemic stroke: results from a population-based study. Stroke. 2005 Jun;36:1115–1119. Available from: https://doi.org/-10.1161/01.str.0000166053.83476.4a.

[41] Adelborg K, Szépligeti S, Sundbøll J, Horváth-Puhó E, Henderson VW, Ording A, et al. Risk of Stroke in Patients With Heart Failure: A Population-Based 30-Year Cohort Study. Stroke. 2017 May;48:1161–1168. Available from: https://doi.org/10.1161/-strokeaha.116.016022.

[42] Bencivenga L, Komici K, Nocella P, Grieco FV, Spezzano A, Puzone B, et al. Atrial fibrillation in the elderly: a risk factor beyond stroke. Ageing Res Rev. 2020 Aug;61:101092. Available from: https://doi.org/10.1016/j.arr.2020.101092.

[43] Ziaeian B, Fonarow GC. Epidemiology and aetiology of heart failure. Nat Rev Cardiol. 2016 Jun;13:368–378. Available from: https://doi.org/10.1038/nrcardio.2016.25.

[44] Ghuran AV, Camm AJ. Ischaemic heart disease presenting as arrhythmias. Br Med Bull. 2001 Oct;59(1):193–210. Available from: https://doi.org/10.1093/bmb/59.1.193.

[45] Iung B. Interface between valve disease and ischaemic heart disease. Heart. 2000 Sep;84(3):347–352. Available from: https://dx.doi.org/10.1136/heart.84.3.347.

[46] Arboix A. Cardiovascular risk factors for acute stroke: Risk profiles in the different subtypes of ischemic stroke. World J Clin Cases. 2015 May;3(5):418–429. Available from: https://doi.org/10.12998/wjcc.v3.i5.418.

[47] Writing Committee Members, Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, et al. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. Circulation. 2013 Oct;128:e240–e327. Available from: https://doi.org/-10.1161/CIR.0b013e31829e8776.

[48] Lyaker MR, Tulman DB, Dimitrova GT, Pin RH, Papadimos TJ. Arterial embolism. Int J Crit Illn Inj Sci. 2013 Jan;3(23724391):77–87. Available from: https://-dx.doi.org/10.4103/2229-5151.109429.

[49] Cassar K. Intermittent claudication. BMJ. 2006 Nov;333(17095782):1002–1005. Available from: https://doi.org/10.1136/bmj.39001.562813.de.

[50] Olin JW, Sealove BA. Peripheral artery disease: current insight into the disease and its diagnosis and management. Mayo Clinic proceedings. 2010 Jul;85(20592174):678–692. Available from: https://dx.doi.org/10.4065/mcp.2010.0133.

[51] Kim DH, Handschumacher MD, Levine RA, Sun BJ, Jang JY, Yang DH, et al. Aortic valve adaptation to aortic root dilatation: insights into the mechanism of functional aortic regurgitation from 3-dimensional cardiac computed tomography. Circ Cardiovasc Imaging. 2014 Sep;7:828–835. Available from: https://doi.org/10.1161/-circimaging.113.001976.

[52] Underwood MJ, El Khoury G, Deronck D, Glineur D, Dion R. The aortic root: structure, function, and surgical reconstruction. Heart. 2000 Apr;83:376–380. Available from: http://dx.doi.org/10.1136/heart.83.4.376.

[53] Iung B, Vahanian A. Epidemiology of valvular heart disease in the adult. Nat Rev Cardiol. 2011 Mar;8:162–172. Available from: https://doi.org/10.1038/nrcardio.2010.202.

[54] Heit JA. Epidemiology of venous thromboembolism. Nat Rev Cardiol. 2015 Aug;12:464–474. Available from: https://dx.doi.org/10.1038/nrcardio.2015.83.

[55] Misiakos EP, Tsapralis D, Karatzas T, Lidoriki I, Schizas D, Sfyroeras GS, et al. Advents in the Diagnosis and Management of Ischemic Colitis. Front Surg. 2017 Sep;4. Available from: https://dx.doi.org/10.3389/fsurg.2017.00047.

[56] Higgins PDR, Davis KJ, Laine L. Systematic review: the epidemiology of ischaemic colitis. Aliment Pharmacol Ther. 2004 Apr;19:729–738. Available from: https://doi.org/10.1111/j.1365-2036.2004.01903.x.

[57] Ndumele CE, Matsushita K, Lazo M, Bello N, Blumenthal RS, Gerstenblith G, et al. Obesity and Subtypes of Incident Cardiovascular Disease. Journal of the American Heart Association. 2023 Jan;5(8):e003921. Available from: https://doi.org/10.1161/-JAHA.116.003921.

[58] Marijon E, Mirabel M, Celermajer DS, Jouven X. Rheumatic heart disease. Lancet. 2012 Mar;379:953–964. Available from: https://doi.org/10.1016/s0140-6736(11)61171-9.

[59] Carapetis JR, Steer AC, Mulholland EK, Weber M. The global burden of group A streptococcal diseases. Lancet Infect Dis. 2005 Nov;5:685–94. Available from: https://-doi.org/10.1016/s1473-3099(05)70267-x.

[60] Carapetis JR, Beaton A, Cunningham MW, Guilherme L, Karthikeyan G, Mayosi BM, et al. Acute rheumatic fever and rheumatic heart disease. Nat Rev Dis Primers. 2016 Jan;2:15084. Available from: https://doi.org/10.1038/nrdp.2015.84.

[61] Mutagaywa RK, Wind AM, Kamuhabwa A, Cramer MJ, Chillo P, Chamuleau S. Rheumatic heart disease anno 2020: Impacts of gender and migration on epidemiology and management. Eur J Clin Invest. 2020 Aug;50(12):e13374. Available from: https://doi.org/-10.1111/eci.13374.

[62] van Hagen IM, Thorne SA, Taha N, Youssef G, Elnagar A, Gabriel H, et al. Pregnancy Outcomes in Women With Rheumatic Mitral Valve Disease: Results From the Registry of Pregnancy and Cardiac Disease. Circulation. 2018 Feb;137:806–816. Available from: https://doi.org/10.1161/circulationaha.117.032561.

[63] Royal College of Obstetricians and Gynaecologists. Cardiac Disease and Pregnancy. London, UK: Royal College of Obstetricians and Gynaecologists; 2011. 18 p. Good Practice No. 3.

[64] Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, et al. Comparison of various characteristics of women who do and do not attend for breast cancer screening. Breast Cancer Res. 2001 Dec;4:R1. Available from: https://doi.org/10.1186/bcr418.

[65] Wright FL, Green J, Canoy D, Cairns BJ, Balkwill A, Beral V, et al. Vascular disease in women: comparison of diagnoses in hospital episode statistics and general practice records in England. BMC Med Res Methodol. 2012 Oct;12:161. Available from: https://doi.org/10.1186/1471-2288-12-161.

5 Established cardiovascular risk factors and risk of incident cardiovascular multimorbidity

5.1 Introduction

5.1.1 Background

Established cardiovascular risk factors include having diabetes, hypertension, and high blood cholesterol, nonmodifiable personal characteristics like age, and lifestyle behaviours such as smoking and physical inactivity, among others. [1] Many studies have investigated the association between established cardiovascular risk factors and individual cardiovascular diseases (CVDs). Little is known about how these risk factors are associated with incident cardiovascular multimorbidity (CVM) (see Chapter 2), even though CVM was present in over 40% of UK women who developed any incident CVD in middle-age (see Chapter 4). Investigating a range of known cardiovascular risk factors for CVM and whether specific CVDs may drive any risk factor associations with CVM is important to better characterise and understand CVM in UK women.

5.1.2 Research questions

Using data from the Million Women Study (described in Chapter 3), the following research questions were addressed in this chapter:

- What are the prospective associations of established cardiovascular risk factors with CVM and complex CVM?
- What are the prospective associations of established cardiovascular risk factors with selected pairs of CVDs (CVM pairs) that comprise CVM?
- Which, if any, individual CVDs might drive any observed risk factor associations with CVM outcomes?

5.2 Methods

In this section, methods that are specific to the analyses in this chapter are described. The methods that are common to other chapters are described in more detail in Section 3.3 of Chapter 3 (page 79).

5.2.1 Outcomes

The rationale and process of outcomes selection for this thesis has been described in detail in Section 3.3.1 of Chapter 3 (page 79). Briefly, the primary outcomes investigated in the main analyses were CVM, complex CVM, and CVM pairs. Incident CVM was defined as an individual developing two or more of the 19 selected CVDs (see Table 3.3 in Chapter 3, page 81) during the follow-up period. Complex CVM was defined as an individual developing four or more incident CVDs. The six selected CVM pairs were, in descending order of frequency: 'Ischaemic heart disease & Atrial fibrillation' (IHD-AF), 'Ischaemic heart disease & Heart failure' (IHD-HF), 'Ischaemic heart disease & Other Arrhythmias' (IHD-ARR), 'Atrial fibrillation & Heart failure' (AF-HF), 'Other arrhythmias & Heart failure' (ARR-HF), and 'Atrial fibrillation & Stroke' (AF-STR).

5.2.2 Risk factors

The cardiovascular risk factors investigated in these analyses were: (i) anthropometric factors, namely body mass index (BMI; <20, 20-24, 25-29, 30-34, 35+ kg/m^2), and height (<155, 155-159, 160-164, 165-169, 170+ cm); (ii) lifestyle behavioural factors, namely cigarette smoking (never, past, current <15 per day, current 15+ per day), alcohol consumption (1-2, 3-6, 7-14, 15+ units per week), and strenuous physical activity (none, up to once per week, 2-3 times per week, 4+ times per week); and (iii) whether individuals reported being treated for hypertension (yes or no), diabetes (yes or no), and high blood

cholesterol (yes or no) at study recruitment. All risk factor data were collected through the self-reported questionnaire completed at recruitment (available in Appendix B, page 298).

When alcohol consumption was the risk factor of interest, women who drank less than one alcoholic drink per week (i.e. non-drinkers) were excluded from the analyses. This was because six in seven non-drinkers at study recruitment were ex-drinkers, and the reasons they had stopped drinking were unknown. If the ex-drinkers had stopped drinking due to poor health associated with CVD development, this could potentially bias the observed association between alcohol consumption and CVM.

5.2.3 Statistical analyses

5.2.3.1 Cox proportional hazards regression

Cox proportional hazards regression using attained age as the underlying time variable was used to calculate hazards ratios (HRs) and 95% confidence intervals (CIs) for the CVM outcomes (i.e. CVM, complex CVM, and CVM pairs) in relation to the selected risk factors. Section 3.3.4.1 of Chapter 3 (page 85) describes Cox proportional hazards regression in more detail, including the dates of follow-up and censoring.

All regression models were stratified by age at recruitment (<53, 53-55, 56-58, 59-61, 62-64, 65+ years) to account for time in study and cohort effects, and recruitment region (ten geographic regions: Oxford, East Anglia, South West, Thames, West Midlands, North York, Trent, North West [Mersey], North West [Manchester/Lancaster], Scotland).

For multivariable adjusted HRs, models were adjusted for BMI, height, smoking, and strenuous physical activity as defined in Section 5.2.2, and for alcohol consumption (rarely/never, 1-6, 7-14, 15+ units per week), socioeconomic deprivation (fifths of the Townsend index), educational qualifications (none, technical qualifications, secondary school, tertiary [college or university]), parity (nulliparous, 1-2, 3+ full-term pregnancies),

and ever use of menopausal hormone therapy (yes or no). Multivariable adjusted HRs for treatment for hypertension, high cholesterol, and diabetes were not adjusted for BMI, since BMI can be a modifier in the relationship between these factors and some CVDs. [2] For each adjustment variable, missing values (<6% for each variable) were assigned to a separate category (Appendix D, Table D1, page 323).

Differences in HRs across the categories of each risk factor were tested using the likelihood ratio test for heterogeneity, with a significance level of $\alpha = 0.05$. The Benjamini-Hochberg procedure [3] (see Section 3.3.4.1 of Chapter 3, page 85) was applied to control the false discovery rate to account for multiple testing in this chapter. Log-log plots were assessed to check the proportional hazards assumption, which was found to be satisfied in all models.

To informally assess whether risk factor associations of CVM pairs were materially different from those of the component CVDs, effect estimates for each of the six selected CVM pairs were compared with the average of the effect estimates for the component individual CVDs. Both the simple average and the inverse variance weighted average, which could potentially have lower uncertainty and less bias towards the association for the rarer component condition, were computed. If the associations for a CVM pair are markedly different from the averaged effects of its component pairs, this may indicate that there are complex relationships between the component CVDs (beyond the shared risk factor) contributing to the observed association with the CVM pair. In this case, it may be warranted to consider further investigation.

5.2.3.2 Sensitivity analyses

5.2.3.2.1 Cardiovascular multimorbidity minus individual diseases

To assess whether individual CVDs may be driving any observed associations with CVM, risk factor associations with the six selected individual CVDs and with CVM leaving out each CVD were examined. The selected individual CVDs included the five most common CVDs in the cohort – ischaemic heart disease (n=119,847), atrial fibrillation (n=76,058), heart failure (n=38,172), other arrhythmias (n=45,241), and stroke (n=34,408). Venous thromboembolism was also selected as it was a common CVD in the cohort (n=31,309) and is known to be associated with reproductive factors (including use of oral contraceptives and hormone replacement therapy [4, 5]), making it particularly relevant to women's cardiovascular health.

To illustrate, leaving out ischaemic heart disease meant that participants were defined as having CVM if they developed two or more of 18 conditions (i.e. the 19 selected CVDs minus ischaemic heart disease). Comparing the associations with CVM minus ischaemic heart disease to the associations with overall CVM can provide an indication of the role of ischaemic heart disease in the overall CVM associations.

5.2.3.2.2 Addressing potential reverse causation bias

Reverse causation bias occurs when precursors to the disease outcome or unidentified disease have influenced the risk factor status of the participants, which can obscure the true association between the risk factor and incident disease. Reverse causality may be an issue in prospective studies if unidentified or subclinical disease or other early disease processes affect the related risk factors. For example, the incidence of myocardial infarction may be higher in past smokers than in current smokers in the short-term, if individuals experiencing chest pain or other symptoms sometimes quit smoking as a consequence of concern for their health. In this study, possible reverse causation bias was addressed in two

sensitivity analyses: (1) excluding the first five years of follow-up and (2) excluding the first three years of follow-up while simultaneously restricting to women who self-reported good or excellent self-rated health. Poor self-rated health may be an indirect indicator of subclinical disease or poor health not otherwise captured through healthcare records or other questionnaire responses, which can influence health-related behaviours. Therefore, restricting analyses to women who reported good or excellent health may help to minimise the potential bias associated with reverse causation for health-related behavioural risk factors.

5.2.3.2.3 Alcohol consumption including non-drinkers

As explained in Section 5.2.2, non-drinkers were excluded from the main analysis in relation to alcohol consumption because it was not possible to differentiate between never and former drinkers, which led to concerns of reverse causality. For completeness of results, a final sensitivity analysis was conducted including non-drinkers in the analyses of CVM outcomes in relation to alcohol consumption.

5.3 Results

5.3.1 Baseline characteristics

Table 5.1 shows the baseline characteristics of study participants overall and by their incident CVM status at the end of follow-up. Among all women, mean age at recruitment was 56.5 (SD 4.8) years, 20% were current smokers, mean BMI was 26.1 (SD 4.6) kg/m^2, 19% lived in the most deprived areas, and 15% reported being treated for hypertension. Women who developed CVM were, on average, older at recruitment and more likely to smoke, have a higher BMI, and live in the most deprived areas than those who did not develop CVM. Moreover, those who developed incident CVM were also much more likely to report being treated for hypertension, diabetes, and high blood cholesterol than those who did not develop CVM.

Table 5.1 Baseline characteristics in all women and by their incident CVM status at the end of follow-up

CVM status:	**No CVM**		**CVM**			**All women**
Number of incident CVDs:	**0 Disease**	**1 Disease**	**2 Diseases**	**3 Diseases**	**4+ Diseases**	
Women, n	968,758	156,826	60,022	27,766	26,896	1,240,268
Age at recruitment, mean (SD), years	56.0 (4.6)	57.7 (5.0)	58.6 (5.1)	59.3 (5.1)	60.0 (5.1)	56.5 (4.8)
Lifestyle and anthropometric factors						
Current smoker, % (n)	18.7 (171,107)	25.1 (37,061)	27.5 (15,475)	28.0 (7,287)	30.0 (7,575)	20.4 (238,505)
Drinker of alcoholic beverages, % (n)	66.9 (643,242)	59.9 (93,053)	57.8 (34,360)	56.3 (15,457)	53.9 (14,304)	65.0 (800,416)
No strenuous physical activity, % (n)	58.9 (551,717)	64.7 (97,061)	66.6 (38,075)	67.9 (17,888)	69.1 (17,577)	60.4 (722,318)
Body mass index, mean (SD), kg/m^2	25.9 (4.4)	26.8 (4.9)	27.0 (5.2)	27.4 (5.5)	27.7 (5.6)	26.1 (4.6)
Height, mean (SD), cm	162.1 (6.5)	162.0 (6.7)	162.1 (6.8)	162.1 (6.9)	162.1 (7.0)	162.1 (6.6)
Socioeconomic factors						
Living in the most deprived areas, % (n)	17.8 (171,299)	23.0 (35,796)	24.5 (14,580)	25.9 (7,125)	27.0 (7,231)	19.2 (236,031)
No educational qualifications, % (n)	40.8 (385,559)	49.5 (75,155)	51.8 (30,094)	54.4 (14,586)	56.7 (14,674)	43.1 (520,068)
Reproductive factors						
Age at menarche, mean (SD), years	13.0 (1.6)	13.0 (1.6)	13.0 (1.7)	13.0 (1.7)	13.0 (1.7)	13.0 (1.6)
Age at first birth, mean (SD), years	24.0 (4.3)	23.5 (4.2)	23.5 (4.2)	23.5 (4.2)	23.4 (4.2)	23.9 (4.3)
Age at natural menopause, mean (SD), years	49.2 (4.2)	49.0 (4.5)	49.0 (4.6)	48.9 (4.6)	48.9 (4.7)	49.1 (4.3)
Pre/perimenopausal, % (n)	14.7 (141,705)	8.3 (12,927)	6.2 (3,707)	5.0 (1,365)	4.1 (1,101)	13.0 (160,805)
Parity, mean (SD)	2.1 (1.2)	2.3 (1.3)	2.3 (1.4)	2.3 (1.4)	2.4 (1.4)	2.2 (1.2)
Ever breastfed children in parous women, % (n)	67.9 (467,247)	67.8 (73,876)	68.7 (28,393)	68.2 (12,763)	68.6 (12,278)	67.9 (594,557)
Ever use of contraceptive pills, % (n)	61.7 (592,104)	55.3 (85,509)	52.0 (30,722)	49.1 (13,400)	46.7 (12,328)	59.8 (734,063)
Ever use of HRT, % (n)	50.4 (482,859)	52.2 (80,659)	50.0 (29,533)	48.0 (13,079)	44.9 (11,838)	50.4 (617,968)
Healthcare factors						
Self-reported treatment for hypertension, % (n)	12.3 (119,320)	19.6 (30,664)	23.1 (13,879)	27.2 (7,546)	31.4 (8,433)	14.5 (179,842)
Self-reported treatment for diabetes mellitus, % (n)	1.4 (13,669)	3.2 (4,998)	4.7 (2,799)	6.4 (1,783)	8.9 (2,393)	2.1 (25,642)
Self-reported treatment for high cholesterol, % (n)	2.2 (20,974)	3.7 (5,755)	4.4 (2,626)	5.0 (1,374)	5.9 (1,594)	2.6 (32,323)

CVM=Cardiovascular multimorbidity, SD=Standard deviation, HRT=Hormone replacement therapy

5.3.2 Risk factors for cardiovascular multimorbidity

Figure 5.1 displays the minimally and multivariable adjusted HRs and 95% FCIs for incident CVM by categories of the selected cardiovascular risk factors. All risk factors were statistically significantly associated with incident CVM after controlling for the false discovery rate ($p<0.001$). Overall, multivariable adjusted HRs were smaller in magnitude compared to the minimally adjusted effect estimates. The focus of the following sections will be on the multivariable adjusted HRs.

5.3.2.1 Anthropometric factors

Both being underweight and being overweight were associated with a higher risk of CVM compared to having a healthy BMI (20-24 kg/m^2): the multivariable adjusted HR for CVM was 1.10 (95% CI 1.06-1.13) in underweight women (BMI <20 kg/m^2) and 2.54 (2.48-2.61) in severely obese women (BMI 35+ kg/m^2).

The relationship between height and CVM was inflated after multivariable adjustment ($\chi^2_{attenuation}$ = -114%). The negative attenuation in the LR statistic represents a stronger association, or in other words, a larger HR in absolute value. For example, compared to the shortest women (height <155 cm), women with a height of 165-169 cm had a 4% lower risk of CVM (0.96, 0.94-0.98) in the minimally adjusted model, but a 13% higher risk (1.13, 1.11-1.16) in the multivariable adjusted model. This may indicate that the positive association between height and CVM was initially masked by some of the covariates whose effects were controlled for in the multivariable adjusted model.

5.3.2.2 Lifestyle factors

Both past and current smoking was associated with a higher risk of CVM compared to never smoking. Moreover, current smokers who smoked 15+ cigarettes per day were more likely to develop CVM than those who smoked <15 cigarettes per day, with multivariable

adjusted HRs compared to never smokers of 2.42 (2.37-2.47) and 1.83 (1.79-1.87), respectively.

Compared to the lowest level of alcohol consumption (1-2 units per week), there was no difference in CVM risk associated with the highest level of consumption (15+ units per week), while consuming 3-14 units per week was associated with only a 5-7% lower risk of CVM.

Any level of strenuous physical activity was associated with a somewhat lower risk of CVM compared to no activity. However, multivariable adjustment attenuated the HRs and greatly reduced the LR statistic ($\chi^2_{attenuation}$ = 85%). There was no suggestion of a trend by frequency of activity since the HRs were similar across all groups with any activity (compared to no activity) – the multivariable adjusted HR for CVM was 0.87 (0.85-0.88) in women who reported physical activity of up to once per week, 0.84 (0.82-0.86) in those who reported 2-3 times of activity per week, and 0.91 (0.89-0.94) in women who reported 4 or more times of activity per week.

5.3.2.3 Treatment for hypertension, diabetes, and high blood cholesterol

The multivariable adjusted HRs for CVM risk associated with treatment for high blood pressure, diabetes mellitus, and high blood cholesterol were 1.90 (1.87-1.93), 2.98 (2.90-3.07). and 1.55 (1.51-1.60), respectively.

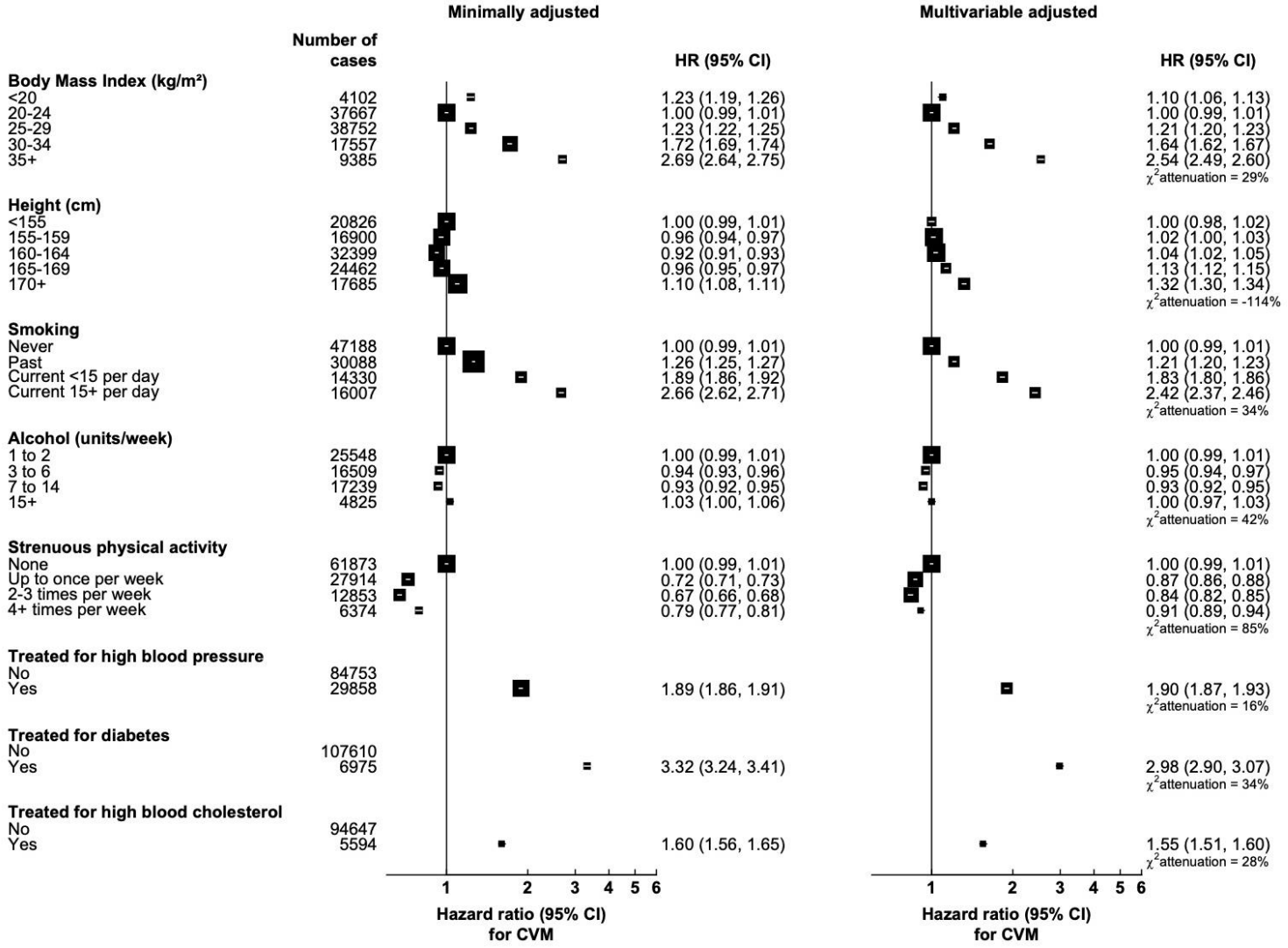


Figure 5.1 Relative risks of incident cardiovascular multimorbidity in relation to cardiovascular risk factors

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. Models for treatment for high blood pressure, diabetes, and high blood cholesterol were not adjusted for body mass index. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed where more than two groups were compared.

5.3.3 Risk factors for complex cardiovascular multimorbidity

Figure 5.2 displays the minimally and multivariable adjusted HRs and 95% FCIs for complex CVM (i.e. developing four or more CVDs) by categories of the selected cardiovascular risk factors. As with CVM, all risk factors were statistically significantly associated with complex CVM after controlling for the false discovery rate ($p<0.001$). Risk factor associations for complex CVM were similar in shape but often larger in magnitude than those for CVM. For instance, being treated for diabetes was a stronger risk factor for complex CVM risk (multivariable adjusted HR 4.13, 95% CI 3.93-4.33) than for CVM (2.98, 2.90-3.07). The exceptions were being underweight, which was not associated with complex CVM, and alcohol consumption and strenuous physical activity, for which the HRs were similar in magnitude for both CVM and complex CVM. The level of χ^2 attenuation for each risk factor was similar for CVM and complex CVM, except for height and alcohol consumption: the LR statistics for complex CVM (compared to those for CVM) had a smaller increase for height and a smaller reduction for alcohol consumption.

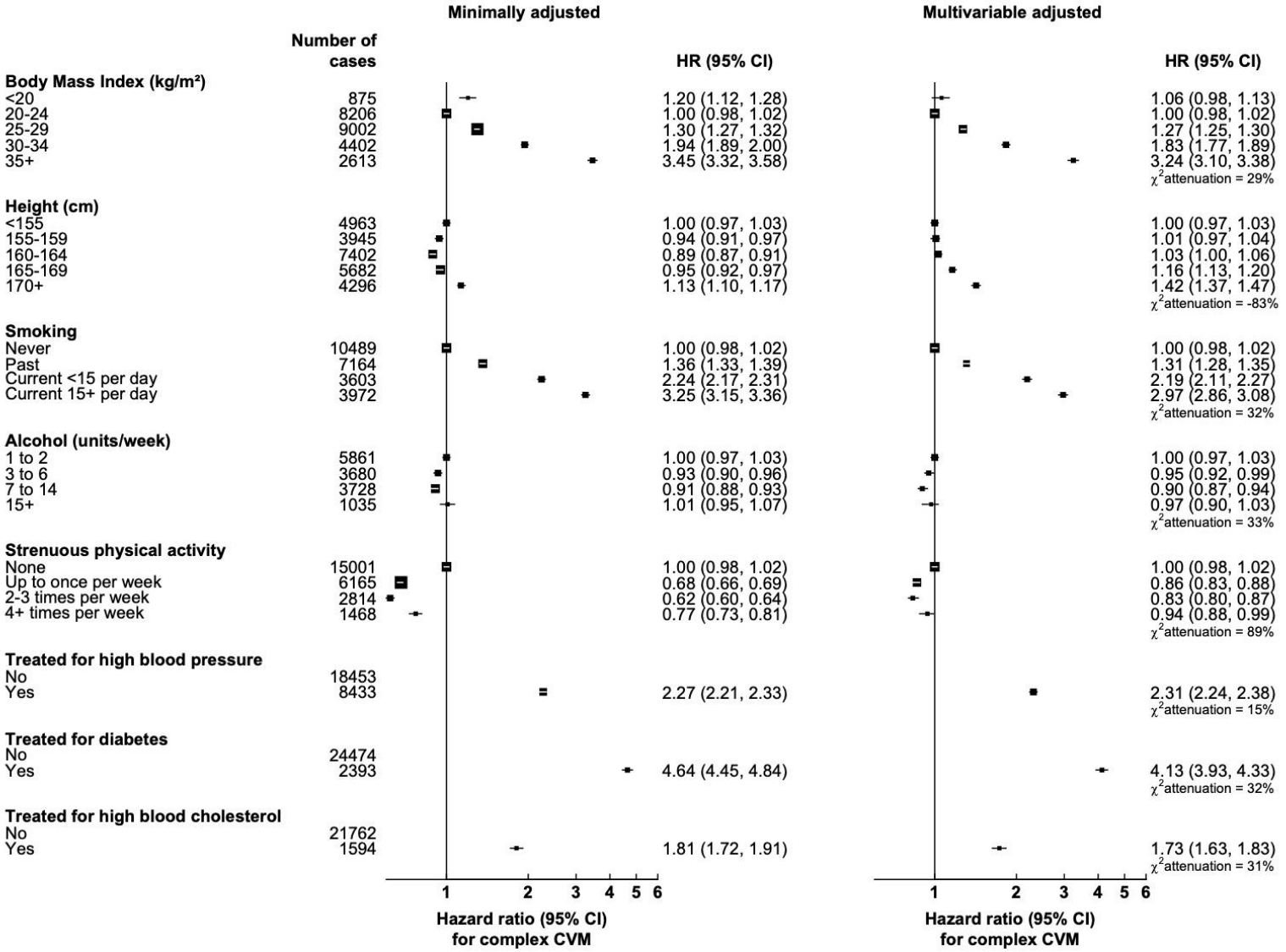


Figure 5.2 Relative risks of incident complex cardiovascular multimorbidity in relation to cardiovascular risk factors

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. Models for treatment for high blood pressure, diabetes, and high blood cholesterol were not adjusted for body mass index. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed where more than two groups were compared.

5.3.4 Risk factors for selected cardiovascular multimorbidity pairs

Figure 5.3 and Figure 5.4 show the multivariable adjusted HRs and 95% FCIs for six selected pairs of CVDs (i.e. CVM pairs) in relation to the risk factors of interest. Again, all risk factors were statistically significantly associated with each CVM pair after controlling for the false discovery rate ($p<0.001$). Figure 5.3 displays risk factor associations for the three most common CVM pairs – all of which included ischaemic heart disease – and in descending order are: 'Ischaemic heart disease & Atrial fibrillation' (IHD-AF), 'Ischaemic heart disease & Heart failure' (IHD-HF), 'Ischaemic heart disease & Other Arrhythmias' (IHD-ARR). Figure 5.4 displays risk factor associations for three other selected CVM pairs, all of which included cardiac arrhythmias: 'Atrial fibrillation & Heart failure' (AF-HF), 'Other arrhythmias & Heart failure' (ARR-HF), and 'Atrial fibrillation & Stroke' (AF-STR).

Generally, risk factor associations for CVM pairs were similar in shape and magnitude to associations for complex CVM (and therefore higher in magnitude than those for CVM). The CVM pairs including heart failure (namely IHD-HF, AF-HF, and ARR-HF) had larger relative risks associated with obesity (vs. healthy BMI) compared to those for either CVM or complex CVM. For instance, women who were severely obese had two to three times the risk of CVM and complex CVM compared to women of healthy BMI, but five times the risk of AF-HF (multivariable adjusted HR 5.25, 95% CI 5.01-5.50; Figure 5.4). Furthermore, women who self-reported treatment for diabetes had more than five times the risk of IHD-HF (5.37, 5.12-5.64; Figure 5.3) compared to those who did not report treatment, but only had three to four times the risk of CVM and complex CVM.

Exceptions were also observed for the risks of pairs including atrial fibrillation in relation to height and alcohol consumption. The relative risks associated with taller height for all

three pairs including atrial fibrillation (IHD-AF, AF-HF, and AF-STR) were higher than those observed for complex CVM. Compared to the shortest women, the tallest women had a 63% higher risk of IHD-AF (1.63, 1.58-1.68; Figure 5.3), 91% higher risk of AF-HF (1.91, 1.84-1.98; Figure 5.4), and 59% higher risk of AF-STR (1.59, 1.51-1.68; Figure 5.4), while they had a 42% higher risk of complex CVM. In terms of alcohol consumption, the highest level of consumption (compared to the lowest) was associated with somewhat higher risks of AF-HF (1.15, 1.06-1.24; Figure 5.4) and AF-STR (1.30, 1.18-2.08; Figure 5.4), while for all other CVM outcomes, women who consumed the highest level of alcohol either did not have any difference in risk or had a lower risk compared to women who consumed the lowest level.

Finally, the associations of AF-STR with physical activity and treatment for high blood cholesterol were weaker than those for all other CVM outcomes. Compared with women who did not exercise, women who exercised up to three times per week only had a 7-8% lower risk of AF-STR ($\chi^2_{attenuation}$ = 93%), in contrast to 11-24% lower risks for all other CVM outcomes (CVM, complex CVM, and all other CVM pairs). In women who reported being treated for high blood cholesterol compared with those who did not, the relative risk of AF-STR (1.19, 1.06-1.33; Figure 5.4) was lower than the relative risks of CVM (55% higher risk), complex CVM (73% higher risk), and all other CVM pairs (39-71% higher risks).

Risk factor associations with the six CVM pairs were similar to both the simple and the inverse variance weighted averages of associations with the component individual CVDs of each pair (Appendix D, Table D2-D7, page 327-332).

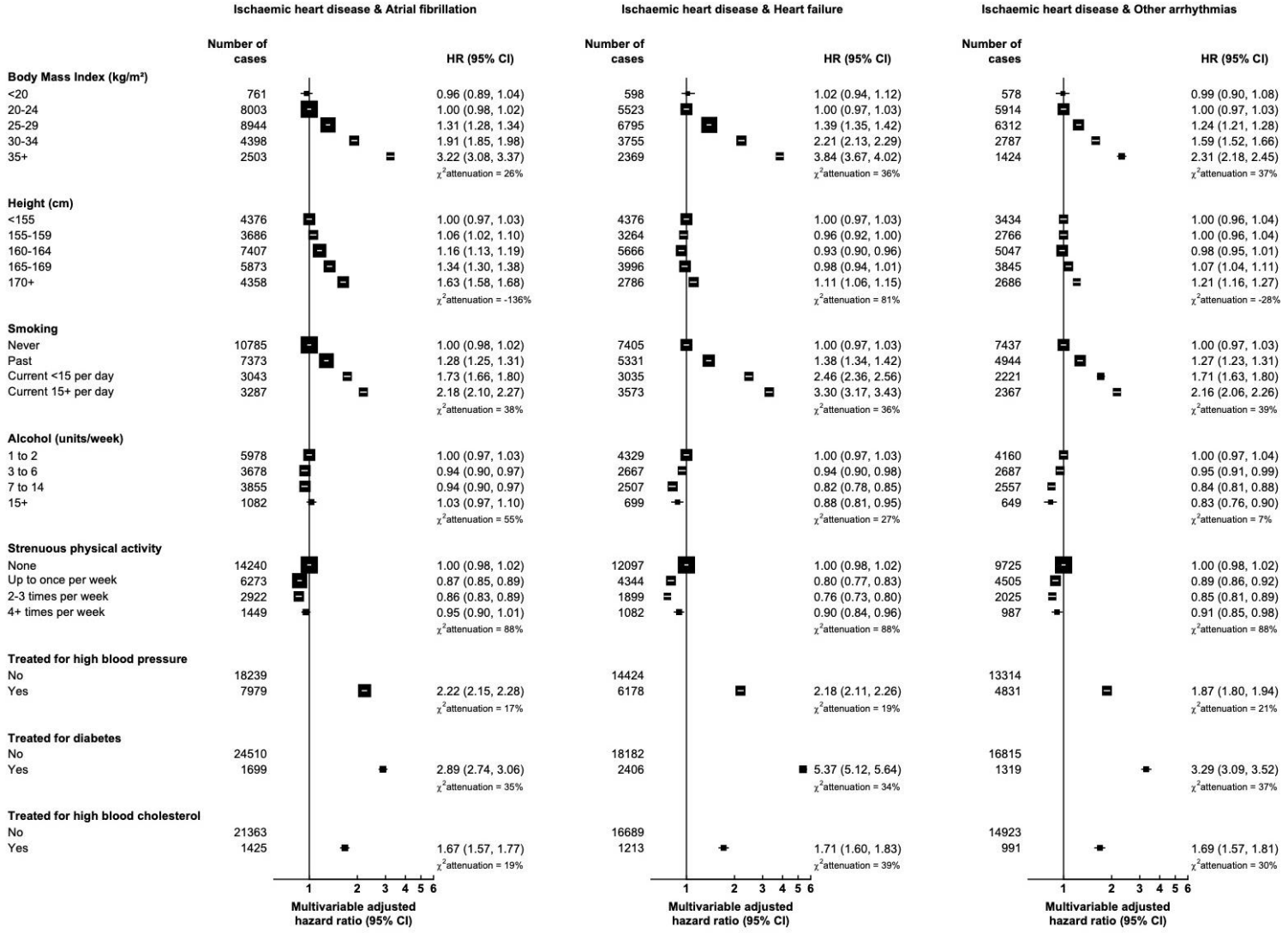


Figure 5.3 Relative risks of CVM pairs including ischaemic heart disease in relation to cardiovascular risk factors

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. Models for treatment for high blood pressure, diabetes, and high blood cholesterol were not adjusted for body mass index. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed where more than two groups were compared.

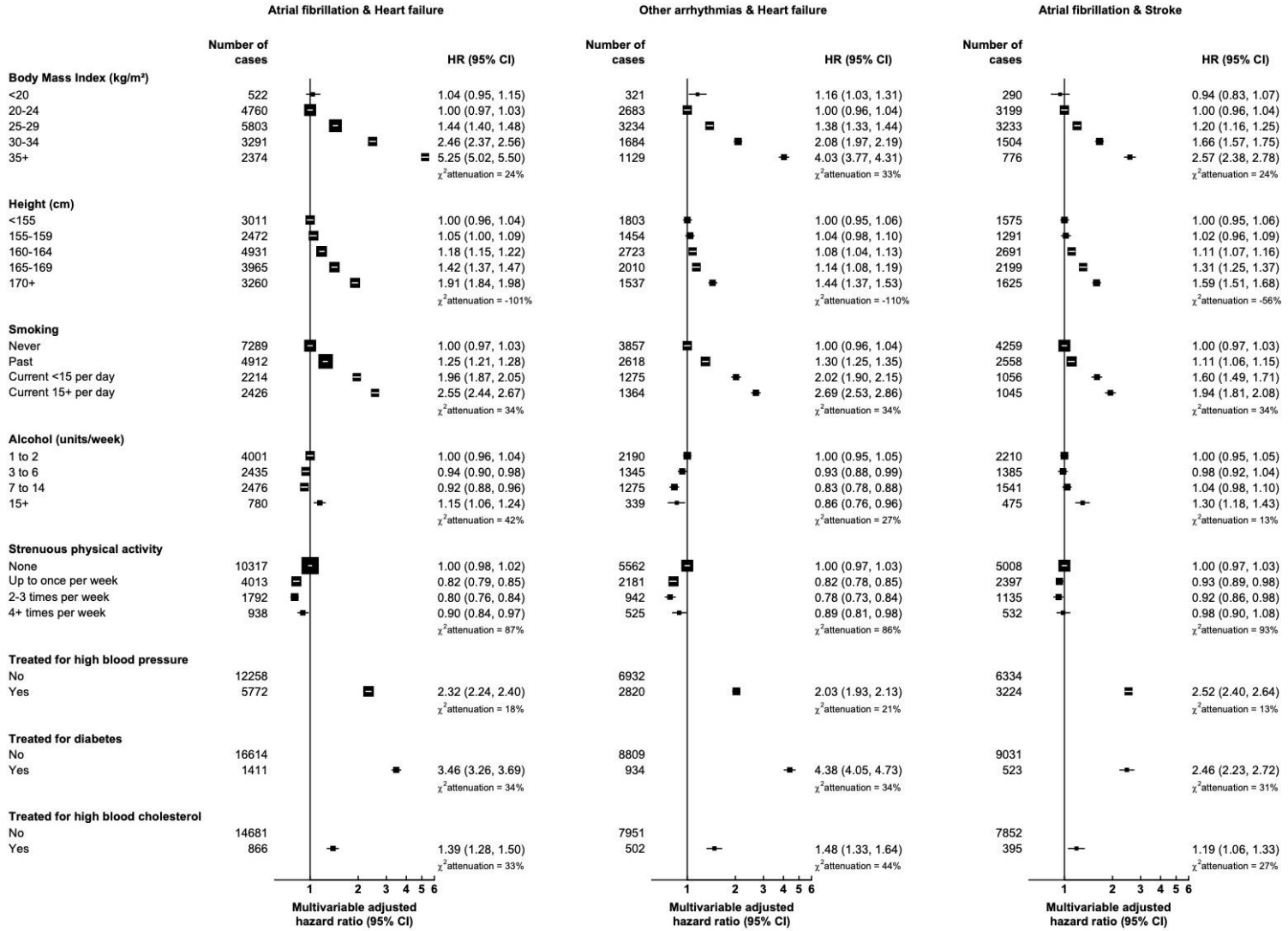


Figure 5.4 Relative risks of other CVM pairs including cardiac arrhythmias in relation to cardiovascular risk factors

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. Models for treatment for high blood pressure, diabetes, and high blood cholesterol were not adjusted for body mass index. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed where more than two groups were compared.

5.3.5 Sensitivity analyses

5.3.5.1 Assessing the role of individual cardiovascular diseases in the observed associations with cardiovascular multimorbidity

Risk factor associations with individual CVDs (Appendix D, Figure D1-D2, page 325-6) corresponded with the observations for their CVM pairs in the previous section. For example, compared to the shortest women, the tallest women had a larger relative risk of atrial fibrillation (multivariable adjusted HR 2.16, 95% CI 2.12-2.20; Figure D1, page 325) than any other individual CVD. This was consistent with the findings for the CVM pairs including atrial fibrillation (i.e. IHD-AF, AF-HF, and AF-STR). Similarly, compared to women of healthy BMI, severely obese women had a higher relative risk of heart failure (4.36, 4.22-4.50; Figure D1) than of any other CVD, consistent with the findings for the CVM pairs including heart failure (i.e. IHD-HF, AF-HF, and ARR-HF).

Sensitivity analyses leaving out each of the six selected individual CVDs from the definition of CVM were conducted to evaluate the relative importance of individual CVDs in the risk factor associations with CVM (Figure 5.5 & Figure 5.6). Leaving out individual CVDs from the definition of CVM did not materially change most risk factor associations with CVM. One exception was the risk of CVM minus atrial fibrillation in relation to height: the multivariable adjusted HRs for CVM minus atrial fibrillation were smaller than those for CVM in the main analysis (Figure 5.1, page 151).

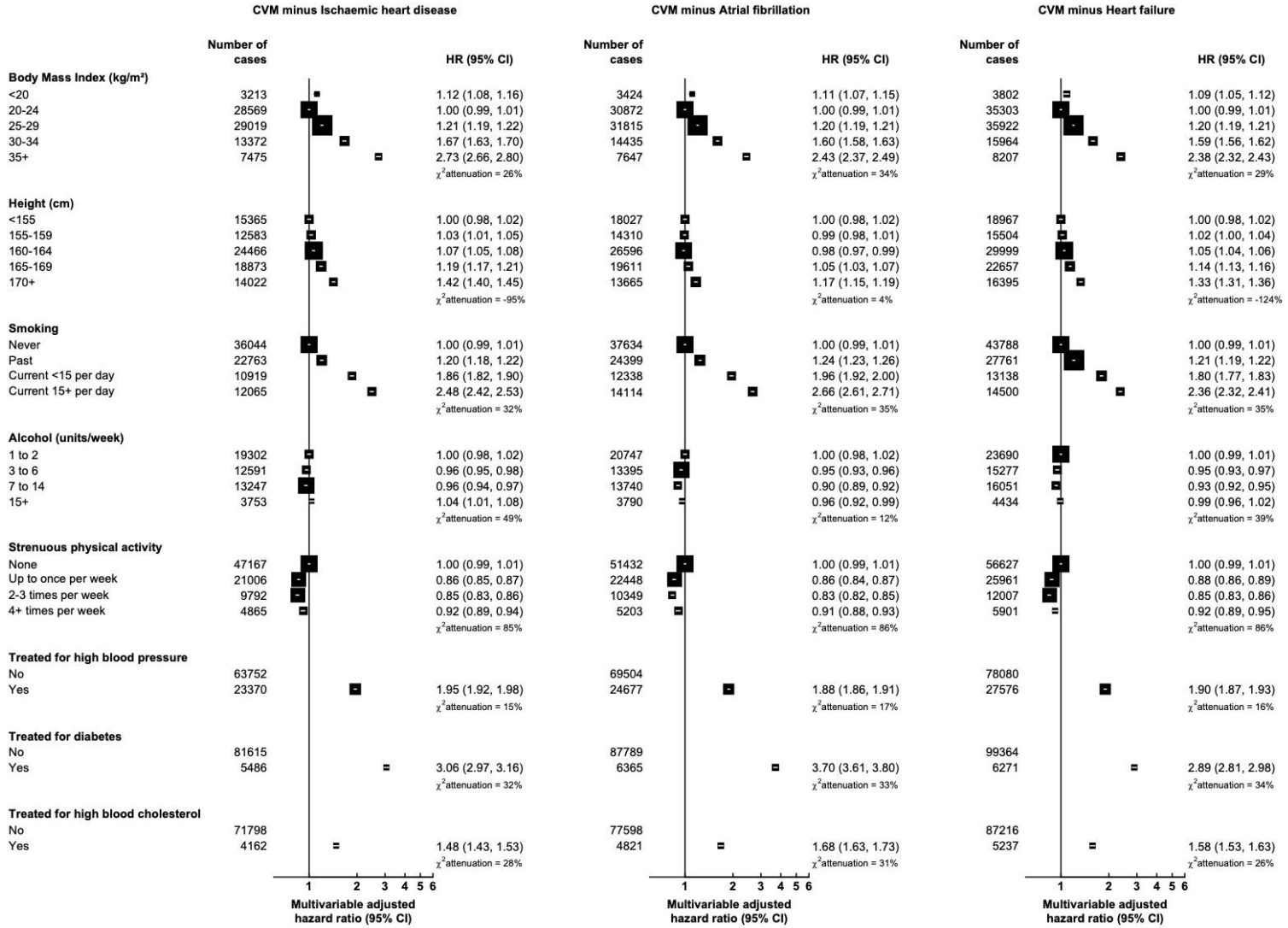


Figure 5.5 Risk factor associations of CVM leaving out ischaemic heart disease, atrial fibrillation, or heart failure as sensitivity analyses

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. Models for treatment for high blood pressure, diabetes, and high blood cholesterol were not adjusted for body mass index. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed where more than two groups were compared.

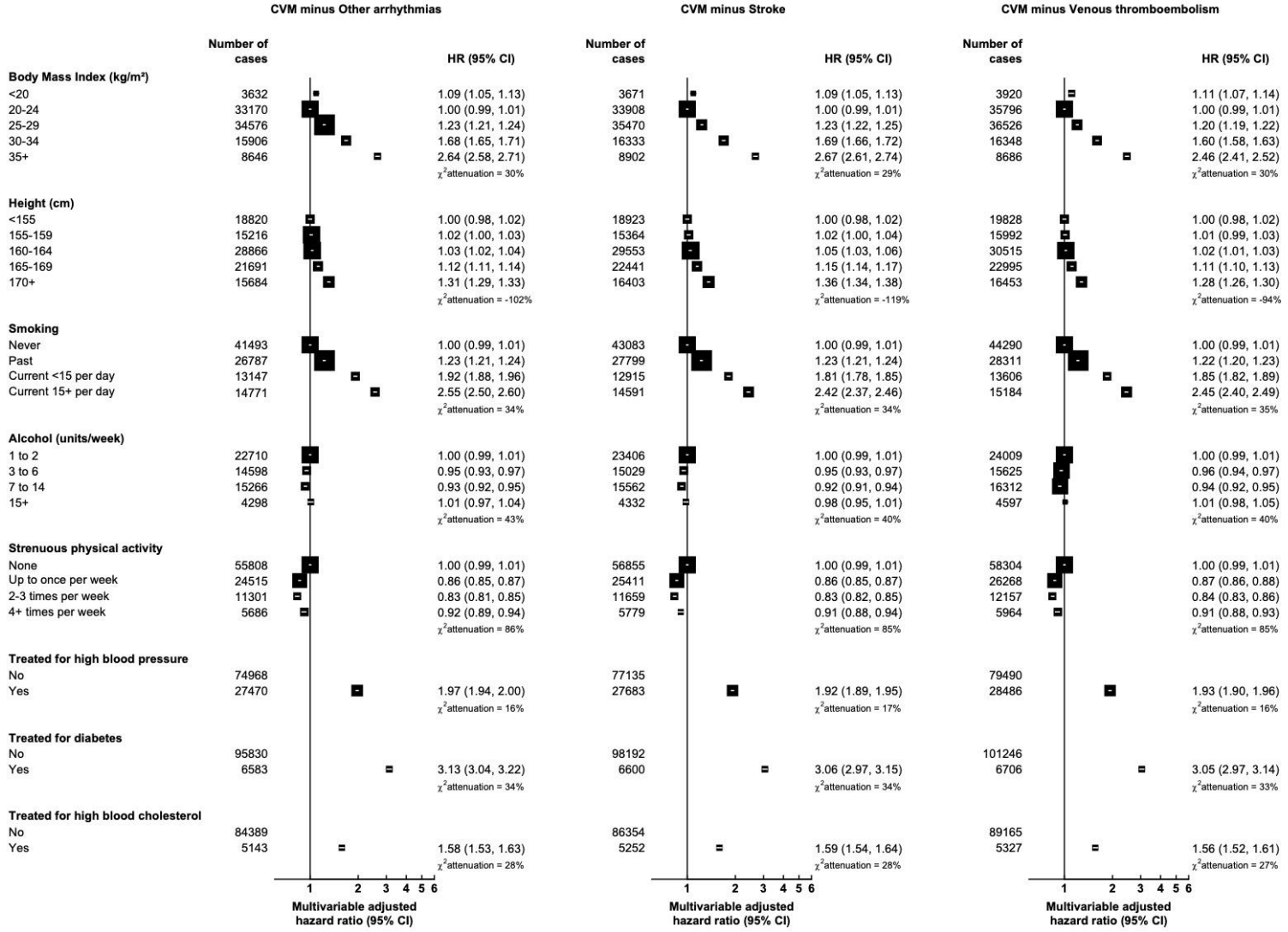


Figure 5.6 Risk factor associations of CVM leaving out other arrhythmias, stroke, or venous thromboembolism as sensitivity analyses

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. Models for treatment for high blood pressure, diabetes, and high blood cholesterol were not adjusted for body mass index. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed where more than two groups were compared.

5.3.5.2 Addressing potential reverse causation bias

Two sensitivity analyses were conducted to better understand the potential effect of reverse causation on risk factor associations with CVM (Table 5.2). Excluding the first five years of follow-up did not materially change the risk factor associations with CVM or complex CVM. Excluding both the first three years of follow-up and restricting to women who self-reported good or excellent self-rated health slightly attenuated the risk factor associations with CVM or complex CVM in relation to lifestyle behavioural risk factors, suggesting there may be some bias attributable to reverse causation.

Table 5.2 Sensitivity analyses controlling for potential reverse causation bias in comparison to the main results

	Multivariable adjusted hazard ratio (95% confidence interval)					
	Cardiovascular Multimorbidity			**Complex Cardiovascular Multimorbidity**		
	Main analysis[a]	**Sensitivity 1[b]**	**Sensitivity 2[c]**	**Main analysis[a]**	**Sensitivity 1[b]**	**Sensitivity 2[c]**
BMI (kg/m^2)						
<20	1.10 (1.06-1.13)	1.08 (1.04-1.13)	1.05 (0.99-1.11)	1.06 (0.98-1.14)	1.03 (0.95-1.13)	1.10 (0.96-1.25)
20-24	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)
25-29	1.21 (1.20-1.23)	1.20 (1.18-1.22)	1.21 (1.18-1.24)	1.27 (1.23-1.32)	1.25 (1.21-1.30)	1.26 (1.20-1.34)
30-34	1.64 (1.61-1.68)	1.62 (1.59-1.66)	1.57 (1.52-1.63)	1.83 (1.76-1.91)	1.83 (1.74-1.91)	1.75 (1.62-1.89)
35+	2.54 (2.48-2.61)	2.52 (2.45-2.60)	2.39 (2.27-2.51)	3.24 (3.08-3.40)	3.22 (3.04-3.41)	3.16 (2.84-3.50)
Height (cm)						
<155	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)
155-159	1.02 (0.99-1.04)	1.02 (0.99-1.04)	1.05 (1.01-1.09)	1.01 (0.96-1.06)	1.01 (0.95-1.06)	1.02 (0.93-1.12)
160-164	1.04 (1.02-1.06)	1.04 (1.02-1.06)	1.09 (1.05-1.12)	1.03 (0.99-1.07)	1.03 (0.98-1.08)	1.09 (1.01-1.17)
165-169	1.13 (1.11-1.16)	1.14 (1.12-1.17)	1.18 (1.14-1.23)	1.16 (1.11-1.21)	1.18 (1.12-1.23)	1.19 (1.10-1.29)
170+	1.32 (1.29-1.35)	1.34 (1.30-1.37)	1.41 (1.36-1.47)	1.42 (1.36-1.49)	1.44 (1.36-1.52)	1.55 (1.42-1.68)
Smoking						
Never	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)
Past	1.21 (1.19-1.23)	1.20 (1.18-1.22)	1.17 (1.14-1.21)	1.31 (1.27-1.36)	1.29 (1.24-1.33)	1.25 (1.18-1.32)
Current <15 per day	1.83 (1.79-1.87)	1.77 (1.73-1.82)	1.73 (1.66-1.80)	2.19 (2.09-2.28)	2.07 (1.97-2.18)	2.01 (1.84-2.19)
Current 15+ per day	2.42 (2.37-2.47)	2.38 (2.32-2.43)	2.38 (2.28-2.48)	2.97 (2.85-3.10)	2.93 (2.79-3.07)	2.83 (2.59-3.09)
Alcohol (units/week)						
Rarely/Never	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)
1-6	0.85 (0.84-0.86)	0.87 (0.85-0.88)	0.90 (0.88-0.93)	0.83 (0.81-0.86)	0.85 (0.82-0.88)	0.90 (0.85-0.95)
7-14	0.81 (0.79-0.83)	0.83 (0.81-0.85)	0.86 (0.83-0.89)	0.76 (0.73-0.80)	0.80 (0.76-0.84)	0.82 (0.77-0.89)
15+	0.87 (0.84-0.89)	0.89 (0.86-0.93)	0.92 (0.87-0.97)	0.82 (0.76-0.88)	0.87 (0.80-0.94)	0.90 (0.80-1.01)

Continued on the next page

Table 5.2 (continued): Sensitivity analyses controlling for potential reverse causation bias in comparison to the main results

	Multivariable adjusted hazard ratio (95% confidence interval)					
	Cardiovascular Multimorbidity			**Complex Cardiovascular Multimorbidity**		
	Main analysis[a]	**Sensitivity 1[b]**	**Sensitivity 2[c]**	**Main analysis[a]**	**Sensitivity 1[b]**	**Sensitivity 2[c]**
Strenuous physical activity						
None	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)
Up to once per week	0.87 (0.85-0.88)	0.88 (0.86-0.90)	0.95 (0.92-0.97)	0.86 (0.83-0.89)	0.88 (0.85-0.91)	0.93 (0.88-0.99)
2-3 times per week	0.84 (0.82-0.86)	0.86 (0.84-0.88)	0.94 (0.91-0.97)	0.83 (0.80-0.87)	0.87 (0.82-0.91)	0.96 (0.89-1.03)
4+ times per week	0.91 (0.89-0.94)	0.93 (0.90-0.96)	1.00 (0.96-1.05)	0.94 (0.88-0.99)	0.95 (0.89-1.02)	1.00 (0.90-1.10)
Being treated for high blood pressure						
No	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)
Yes	1.90 (1.87-1.93)	1.85 (1.82-1.88)	1.84 (1.79-1.89)	2.31 (2.24-2.38)	2.25 (2.18-2.33)	2.23 (2.11-2.35)
Being treated for diabetes						
No	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)
Yes	2.98 (2.90-3.07)	2.81 (2.72-2.90)	2.51 (2.35-2.69)	4.13 (3.93-4.33)	3.88 (3.67-4.11)	3.69 (3.29-4.15)
Being treated for high blood cholesterol						
No	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)
Yes	1.55 (1.51-1.60)	1.48 (1.43-1.53)	1.44 (1.36-1.52)	1.73 (1.63-1.83)	1.62 (1.51-1.73)	1.69 (1.51-1.89)

[a]Main analysis= no exclusions or restrictions – the same results as those displayed in Sections 5.3.2 and 5.3.3 (N = 1,240,268);
[b]Sensitivity 1= excluding the first five years of follow-up (N = 1,143,715);
[c]Sensitivity 2= restricting to women who reported good or excellent health. Self-rated health was reported in the first resurvey questionnaire, on average three years after study recruitment. Therefore, Sensitivity 2 also excluded the first three years of follow-up in effect. (N = 572,206);
Models were stratified by age at recruitment and area of residence and adjusted for socioeconomic deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. Models for treatment for high blood pressure, diabetes, and high blood cholesterol were not adjusted for body mass index.

5.3.5.3 Alcohol consumption including non-drinkers

In the main analyses, non-drinkers were excluded because including former drinkers who may have quit drinking due to poor health may lead to reverse causation bias. The results of sensitivity analyses including non-drinkers showed that reverse causation may indeed be a potential issue: including women who were non-drinkers showed that non-drinkers had a higher risk of various incident CVM outcomes compared to drinkers, with 11% to 26% higher risks in non-drinkers compared to women who drank 1-2 units of alcohol per week (Table 5.3). The only exception was AF-STR risk, where there was no statistically significant difference in risk between non-drinkers and women who drank 1 to 2 units per week.

Table 5.3 Relative risks of incident cardiovascular multimorbidity outcomes in relation to alcohol consumption, including non-drinkers as a sensitivity analysis

Outcome:	**CVM**	**Complex CVM**	**IHD-AF**	**IHD-ARR**	**IHD-HF**	**AF-HF**	**AF-STR**	**HF-ARR**
Alcohol (units/week)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Rarely/never	1.15 (1.13-1.17)	1.18 (1.14-1.22)	1.11 (1.07-1.15)	1.12 (1.07-1.16)	1.26 (1.21-1.31)	1.11 (1.07-1.16)	1.03 (0.97-1.09)	1.11 (1.05-1.18)
1 to 2	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)
3 to 6	0.95 (0.93-0.97)	0.95 (0.91-0.99)	0.93 (0.89-0.98)	0.95 (0.90-1.00)	0.94 (0.89-0.99)	0.94 (0.89-0.99)	0.98 (0.91-1.05)	0.93 (0.87-1.01)
7 to 14	0.93 (0.91-0.95)	0.90 (0.86-0.94)	0.93 (0.88-0.97)	0.84 (0.79-0.88)	0.81 (0.77-0.86)	0.92 (0.87-0.97)	1.04 (0.97-1.12)	0.83 (0.77-0.89)
15+	1.00 (0.96-1.03)	0.96 (0.89-1.03)	1.02 (0.95-1.09)	0.82 (0.75-0.89)	0.88 (0.80-0.96)	1.15 (1.05-1.25)	1.29 (1.16-1.44)	0.85 (0.75-0.96)

CVM=cardiovascular multimorbidity, IHD=ischaemic heart disease, AF=atrial fibrillation, HF=heart failure, STR=stroke, ARR=other arrhythmias
Multivariable adjusted models were stratified by age at recruitment and area of residence and additionally adjusted for area deprivation, educational attainment, physical activity, smoking, height, body mass index, parity, and ever use of menopausal hormone therapy.

5.4 Discussion

5.4.1 Key findings and interpretation of results

5.4.1.1 Risk factors for cardiovascular multimorbidity

5.4.1.1.1 Overall summary

In this cohort of 1.2 million UK women, BMI, height, smoking, and self-reported treatment for hypertension, diabetes, and high blood cholesterol were independent risk factors for incident CVM after multivariable adjustment. Severe obesity (compared to a healthy BMI), heavy smoking (compared to never smoking), and treatment for hypertension or diabetes (compared to not being treated) were associated with a two to three times higher risk of CVM. Treatment for high blood cholesterol was associated with a 55% higher risk of CVM. Compared to the shortest women, the tallest women had a 32% higher risk of CVM. The associations were similar in shape, but often stronger in magnitude for complex CVM. Leaving out individual CVDs from the definition of CVM did not materially change most risk factor associations, apart from atrial fibrillation in the relationship with height.

5.4.1.1.2 Context in the literature

Existing evidence on risk factors for CVM is scarce. The systematic review of prospective studies reported in Chapter 2 of this thesis did not find any studies that investigated CVM in relation to cardiovascular risk factors. The evidence most relevant to CVM were three studies on cardiometabolic multimorbidity, defined as developing at least two from coronary heart disease, stroke, and diabetes. [6-8] Since diabetes is not a CVD but rather a risk factor, the results from these studies were not directly comparable to the findings of this chapter.

Nevertheless, all three studies [6-8] found a positive association between BMI and cardiometabolic multimorbidity, which was consistent with the findings for CVM in this

chapter. However, the magnitudes of the relative risks reported for cardiometabolic multimorbidity in these studies were larger than those for CVM in this chapter. This may partly be due to the inclusion of diabetes in cardiometabolic multimorbidity, which is known to be strongly linked with obesity. [9] In fact, when Kivimäki et al. [6] reported associations between BMI and CVD only (i.e. coronary heart disease and/or stroke), the relative risks were of similar magnitudes to those for CVM in this chapter.

Current smoking, physical inactivity, and hypertension were also investigated in two [7, 8] of the three studies. Current smoking was associated with about twice the risk of cardiometabolic multimorbidity compared to never smoking (after multivariable adjustment), which was consistent with the findings for CVM in this chapter. Physical inactivity was not significantly associated with the risk of cardiometabolic multimorbidity in either study, although each study only had a few hundred incident cases. This was still consistent with the findings for CVM in this chapter, where the large attenuation in the LR statistic indicated that any remaining association was possibly due to residual confounding. Self-reported diagnosis of hypertension was associated with approximately twice the risk of cardiometabolic multimorbidity after multivariable adjustment in a cohort of middle-aged Australian women [7], which was consistent in direction and magnitude to the relationship between self-reported treatment for hypertension and CVM in this chapter. Although hypertension (measured blood pressure of ≥140/90 mmHg or use of antihypertensive medication) was associated with only a 50% higher risk of cardiometabolic multimorbidity in a study in the UK Whitehall II cohort [8], the effect estimate was not adjusted for any covariates apart from demographic factors, and the study had a smaller sample size. Therefore, it was less directly comparable to the results in this chapter than the Australian study.

5.4.1.1.3 Previous publications in the Million Women Study

The findings in this chapter were consistent with previous studies of risk factors for individual CVDs in the Million Women Study, which investigated BMI, smoking, and self-reported treatment for diabetes. Risk of coronary heart disease, [10] venous thromboembolism, [11] and ischaemic stroke [12, 13] increased progressively with increasing BMI. However, higher BMI was associated with a lower risk of haemorrhagic stroke (adjusted HR 0.89, 95% CI 0.86-0.92). [12] Previous studies also found that smoking was associated with a higher risk of coronary heart disease [14], cerebrovascular disease [14], all stroke subtypes [13], and venous thromboembolism [15]. Self-reported treatment for diabetes was associated with a higher risk of coronary heart disease, pulmonary embolism, ischaemic stroke, and intracerebral haemorrhage, but a lower risk of subarachnoid haemorrhage [13, 16], although stroke subtypes were not explored in this thesis.

5.4.1.2 Risk factors for selected cardiovascular multimorbidity pairs

5.4.1.2.1 Overall summary

Generally, risk factor associations for CVM pairs were similar in shape but larger in magnitude to those for CVM and were similar in magnitude to those for complex CVM. Comparing the magnitude of risk factor associations across CVM pairs, BMI and smoking-associated relative risks were greatest for pairs including heart failure, while height-associated relative risks were greatest for pairs including atrial fibrillation. Heaviest alcohol consumption (≥15 units/week) compared to the lowest (1-2 units/week) was associated with lower relative risks of all CVM pairs except for those including atrial fibrillation.

Risk factor associations with the six CVM pairs were not materially different from the averaged effects of the component individual conditions. While there was no evidence to

suggest strong complex relationships between the component CVDs (beyond shared risk factors) in the aetiology of the selected CVM pairs, the existence of such relationships could not be ruled out based on these pragmatic, if simple, analyses. Future methodological research could include exploration of suitable formal statistical analyses to examine in more detail what might be inferred about the possible aetiology of CVM in this way.

5.4.1.2.2 Context in the literature

No previous studies of risk factors for specific CVM pairs were found, but the risk factor associations for CVM pairs observed in this study were largely consistent with the literature on individual CVDs.

Obesity is a known risk factor for various individual CVDs, including coronary heart disease, stroke, and heart failure [17, 18]. A study of pooled individual-level data from ten prospective cohorts in the USA found that compared to women with a normal BMI (18.5-24.9 kg/m^2), obese women (BMI 30-39.9 kg/m^2) had a 75% higher risk of myocardial infarction, a 28% higher risk of stroke, and more than double the risk of heart failure (adjusted HR 2.28, 95% CI 2.00-2.60). [19] These relative risks were similar in magnitude to those estimated for individual CVDs (ischaemic heart disease, stroke, and heart failure) in this chapter. Furthermore, they were consistent with the finding in this chapter that obesity was associated with a greater risk of CVM pairs including heart failure than other pairs including ischaemic heart disease or stroke.

Smoking is an established risk factor for individual CVDs such as coronary heart disease, cerebrovascular disease, venous thromboembolism, and heart failure. [13-15, 20, 21] In a cohort of 1.93 million people in England aged ≥30 years (mean age in women with smoking data was 48 years, SD 16) followed-up between 1997-2010 via linkage to

electronic health records (covering primary care, hospitalisations, myocardial infarction registry and mortality), smoking had different associations with different subtypes of CVDs. [20] In that study, current smokers had a smaller relative risk of heart failure (age-adjusted HR 1.62, 1.47-1.79) than in this chapter, where current smokers had 2-3 times the risk of heart failure compared to never smokers after multivariable adjustment. This discrepancy may partly be due to differences in age and birth cohort, method of outcome ascertainment, and adjustment variables between studies. The English study also reported that current (vs. never) smokers had a higher risk of subarachnoid haemorrhage (2.70, 95% CI 2.27-3.21) than intracerebral haemorrhage (1.61, 1.37-1.89) or ischaemic stroke (1.90, 1.72-2.10). They also reported that current smoking was more strongly associated with myocardial infarction (2.32, 2.20-2.45) than unstable angina (1.54, 1.38-1.72). This was in line with the 2-2.5 times higher risk of stroke and the 1.7-2.1 times higher risk of ischaemic heart disease observed in current (vs. never) smokers in this chapter, although it was difficult to compare these estimates directly.

Observational studies have found a positive association between height and atrial fibrillation in many different populations [22-28], and a recent Mendelian randomisation study corroborated this relationship [29]. Similarly, height was a strong risk factor for CVM pairs including atrial fibrillation in this chapter. Moreover, the only CVM pair investigated in this study which did not include any heart arrhythmia was IHD-HF, which was also the only pair that did not have an evident association with height.

Most previous observational studies found that the risk of atrial fibrillation increases with higher levels of alcohol consumption. [30-33] Similarly, heavy alcohol consumption was associated with a higher risk of all stroke subtypes. [34-36] For heart failure, light drinking was associated with a lower risk compared to non-drinking according to observational evidence. [37-39] These findings were in line with the results in this chapter, in that heavy

alcohol consumption was associated with a higher risk of AF-STR and AF-HF. However, findings from Mendelian randomisation studies did not support causal associations between alcohol consumption and atrial fibrillation [37, 40] or heart failure [39], while the evidence on stroke subtypes was variable [36, 37, 41, 42].

5.4.2 Potential effect of changes in the prevalence of risk factors over time

When interpreting the findings of this chapter, it is important to consider the changes in the prevalence of key cardiovascular risk factors over the follow-up period from 1996/2001 to 2016. Since follow-up began for the participants of the Million Women Study in 1996-2001, there has been an increase in the prevalence of obesity, decrease in smoking, and increase in the prescription of medications to control cardiometabolic risk factors (e.g. statins) in the UK general population. [43-46] Such real changes over time as well as reporting errors in risk factors can result in regression dilution bias in the associations with CVM outcomes. Several previous studies in the Million Women Study have explored the changes in self-reported risk factors during follow-up and to what extent they may influence associations with disease outcomes.

The validity of self-reported anthropometry over time has previously been investigated in the Million Women Study by comparing self-reported values of weight, height, and derived BMI at four different time points over follow-up with measured values from an examination in 2008 (i.e. the mean year in which the examination was completed for the participants of this validation study, around 9 years after recruitment). [47] For height, the mean of the first reported value (mean year 1999) was greater by 1.2 cm (95% CI 1.1 to 1.3) than the measured value in 2008. For BMI, the mean of the first reported value was lower than the mean of the measured value. This can partly be due to real changes in height and BMI over time (as height tends to decrease, and adiposity tends to increase over

time in women), and the systematic overestimation of height and underestimation of weight. [48-50] The Pearson correlations between the first self-report and the measured value for height and BMI were 0.88 and 0.85, respectively.

Furthermore, the same study found that according to the estimated regression dilution ratios, reporting errors were unlikely to result in strongly biased estimates of associations between anthropometric variables and disease outcomes. [47] Regression dilution bias occurs when variation in the risk factor attenuates associations with the outcome to the null. Regression dilution ratios were calculated by linear regression of the measured values against the self-reported values in each questionnaire, providing an estimate of the relative attenuation of regression coefficients due to any source of variation in the risk factor (including random error, measurement error, and changes in population characteristics over time). The regression dilution ratios were 0.85 for height and 0.98 for BMI at the time the variables were first reported, and the ratios stayed similar over the follow-up period, indicating that changes in height and BMI over time are unlikely to result in substantial time-dependent regression dilution bias. Similarly, there was no suggestion of a trend in regression dilution ratios when examining height and BMI by five categories. These findings indicate that the effect estimates reported in this chapter in relation to BMI and height are unlikely to be strongly biased by reporting errors or changes in these variables over time.

Another study in the Million Women Study compared self-reported smoking habits in the recruitment questionnaire and in resurveys conducted about three and eight years later. [14] For never-smokers and past smokers, there was little misclassification or change in smoking habits; less than 0.1% of women who were never smokers and 2.2% of past smokers at recruitment reported being current smokers at the 3-year resurvey. Among women who reported being current smokers at recruitment, 23% and 44% reported having

stopped smoking at the 3-year and 8-year resurvey, respectively. Smoking cessation was more common among women who smoked fewer cigarettes per day. Such changes in smoking habits are likely to have resulted in attenuated estimates of relative risks associated with continued smoking in the analyses presented in this chapter; in other words, the relative risks of current smoking in this chapter are possibly underestimates of the true association.

Finally, a study in the Million Women Study found that there was a decrease in the agreement of self-reported physical activity measures as longer time elapsed between initial and repeat questionnaires, which highlighted a need for regular updates of physical activity data over long follow-up periods. [51] Since the analyses in this chapter only used physical activity data from the recruitment questionnaire despite the long-term follow-up, there is likely to be regression dilution bias. Although the magnitude of the potential attenuation of the relative risks is unclear, the associations reported in this chapter are likely to be closer to the null than the actual relationship.

5.4.3 Potential role of reverse causation

Another important consideration in the interpretation of the study findings is bias due to reverse causation. Reverse causation is of concern because early symptoms of CVDs can influence some health behaviours and metrics like BMI, such that the risk factor measurements collected at recruitment may have been the result, rather than the cause, of CVD outcomes. For example, women who are already ill with CVD may not be able to engage in much physical activity due to symptoms, and lower BMI may be indicative of underlying pathology rather than a healthy lifestyle. The potential for this bias was addressed in two sensitivity analyses: (1) excluding the first five years of follow-up and (2) excluding the first three years of follow-up and also restricting to women who reported good or excellent health at the first re-survey questionnaire. The reduction in the

magnitude of the effect estimates indicated that there may be slight reverse causation bias in the results for physical activity.

Reverse causation is of particular concern for the analyses of alcohol consumption since women may stop or reduce drinking due to poor health. In the three-year resurvey of the Million Women Study, women were asked if they were lifelong non-drinkers or had stopped drinking for whatever reason. More than 60% of women who had reported no alcohol drinking at study recruitment had not been lifelong teetotallers, but there was no information available on why or when these women had stopped drinking. Therefore, although sensitivity analyses including non-drinkers showed that non-drinkers had a higher risk of various CVM outcomes compared to drinkers, this does not necessarily suggest drinking alcohol is protective against CVM.

5.4.4 Potential role of residual confounding

Another important source of bias in observational studies is residual confounding. This occurs when the effects of unknown confounders or confounders which are not fully measured remain in the observed risk factor associations despite multivariable adjustment. In the Million Women Study, information on potential confounders was collected by self-report at a single point in time, making it susceptible to measurement error. The presence of measurement error means that adjustment will not completely control for the biases introduced by the confounders of interest.

Residual confounding appeared to be a particular issue in the associations with physical activity, as indicated by the large reduction in the LR statistics after multivariable adjustment for virtually all CVM outcomes. Over 86% of all the association between physical activity and CVM outcomes may be accounted for by adjustment for only the known and relatively crudely measured confounders. Therefore, it is plausible that much

of the remaining associations may be explained by residual confounding, and the weak inverse associations observed between physical activity and various CVM outcomes should not be interpreted as suggesting a genuine (let alone causal) relationship.

Potential residual confounding must be considered in all risk factor associations in this study. It is important to note that the LR attenuation method does not intend to quantify the exact contribution of well-measured confounders. Instead, LR attenuation measures how much of the information about the outcome contained in the risk factor (i.e. the improvement in goodness-of-fit attributable to the risk factor) is accounted for by the adjusted covariates. Since the covariates are measured with error, the higher the LR attenuation, the greater the likelihood that residual confounding (e.g. due to measurement error) may explain any remaining association even if the attenuation is less than 100%. Therefore, the LR attenuation method has a somewhat different approach from other methods commonly used in the literature to examine the contribution of covariates to risk factor associations (e.g. the absolute scale difference method). [52]

It is possible to apply the absolute scale difference method to the associations reported in this chapter. To illustrate, the HR point estimates for the association between treatment for diabetes and CVM in this chapter was 3.32 in the minimally adjusted model and 2.98 in the multivariable adjusted model (Figure 5.1, page 151). The contribution of additional covariates to the association between treatment for diabetes and CVM can be calculated as: $100 \times (\log(3.32) - \log(2.98))/\log(3.32) = 9\%$. The LR attenuation reported in this chapter for the same association was 34%. The interpretation of both these results is that the observed relationship between treatment for diabetes and CVM was unlikely to be substantially explained by the adjusted covariates.

5.4.5 Strengths and limitations

The strengths of this study included the large sample size and number of events, which allowed reliable estimation of risk factor associations with CVM outcomes. Most prospective studies do not have enough CVM events across all categories of risk factors to reliably investigate these relationships. In the Million Women Study, there was substantially complete follow-up and only a small proportion of missing values for the risk factor variables (<6% for each variable). Furthermore, despite the possibility of residual confounding, the availability of information on a range of important confounders allowed for their adjustment in these analyses.

Due to the observational design, the analyses in this chapter were susceptible to residual confounding and reverse causation bias as discussed in the two previous sections. A further limitation was the use of self-reported risk factor variables which may be less accurate than direct measurements, and in some cases (such as self-reported weight) may be prone to social desirability bias. However, studies in the Million Women Study have examined the validity and repeatability of the risk factors investigated in this chapter and have shown them to be broadly reliable. [14, 47, 51] The use of self-reported treatment for diabetes, hypertension, and high blood cholesterol as proxies for the risk factors themselves is another limitation. It is possible that participants diagnosed with these conditions were not receiving treatment at recruitment and hence incorrectly counted as not having the condition. If so, this would have biased the observed associations towards the null. Nonetheless, a previous publication in the Million Women Study found that there was high agreement between general practice prescription data and self-reported treatment for diabetes and hypertension. [53]

5.4.6 Conclusions

Smoking, BMI, height, and being treated for diabetes, hypertension, or hyperlipidaemia were independently associated with CVM outcomes, while the evidence was not as strong for physical activity and alcohol consumption. Some individual CVDs appeared to be more important drivers of the risk factor associations with CVM than others (in particular, atrial fibrillation in the association with height). The findings in this chapter provided novel insights into the risk factors for CVM and its component CVM pairs in a large prospective cohort with substantially complete follow-up, thereby addressing an important gap in the evidence.

5.5 References

[1] Wood D. Established and emerging cardiovascular risk factors. American Heart Journal. 2001 Feb;141(2, Supplement):S49–S57. Available from: *https://-www.sciencedirect.com/science/article/pii/S0002870301011516*.

[2] Lu Y, Hajifathalian K, Ezzati M, Woodward M, Rimm EB, Danaei G. Metabolic mediators of the effects of body-mass index, overweight, and obesity on coronary heart disease and stroke: a pooled analysis of 97 prospective cohorts with 1.8 million participants. Lancet. 2014 Mar;383:970–83. Available from: *https://doi.org/10.1016/-s0140-6736(13)61836-x*.

[3] Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Series B Stat Methodol. 1995;57(1):289–300. Available from: *https://doi.org/10.1111/j.2517-6161.1995.tb02031.x*.

[4] Stegeman BH, de Bastos M, Rosendaal FR, van Hylckama Vlieg A, Helmerhorst FM, Stijnen T, et al. Different combined oral contraceptives and the risk of venous thrombosis: systematic review and network meta-analysis. BMJ. 2013 Sep;347:f5298. Available from: *https://doi.org/10.1136/bmj.f5298*.

[5] Canonico M, Plu-Bureau G, Lowe GDO, Scarabin PY. Hormone replacement therapy and risk of venous thromboembolism in postmenopausal women: systematic review and meta-analysis. BMJ. 2008 May;336(7655):1227. Available from: *https://-doi.org/10.1136/bmj.39555.441944.be*.

[6] Kivimäki M, Kuosma E, Ferrie JE, Luukkonen R, Nyberg ST, Alfredsson L, et al. Overweight, obesity, and risk of cardiometabolic multimorbidity: pooled analysis of individual-level data for 120813 adults from 16 cohort studies from the USA and Europe. Lancet Public Health. 2017 Jun;2:e277–e285. Available from: *https://doi.org/10.1016/-s2468-2667(17)30074-9*.

[7] Xu X, Mishra GD, Dobson AJ, Jones M. Progression of diabetes, heart disease, and stroke multimorbidity in middle-aged women: A 20-year cohort study. PLOS Medicine. 2018 Mar;15(3):e1002516. Available from: *https://doi.org/10.1371/-journal.pmed.1002516*.

[8] Singh-Manoux A, Fayosse A, Sabia S, Tabak A, Shipley M, Dugravot A, et al. Clinical, socioeconomic, and behavioural factors at age 50 years and risk of cardiometabolic multimorbidity and mortality: A cohort study. PLoS Med. 2018 May;15(5):e1002571. Available from: *https://doi.org/10.1371/journal.pmed.1002571*.

[9] Abdullah A, Peeters A, de Courten M, Stoelwinder J. The magnitude of association between overweight and obesity and the risk of diabetes: A meta-analysis of prospective cohort studies. Diabetes Research and Clinical Practice. 2010 Sep;89(3):309–319. Available from: *https://doi.org/10.1016/j.diabres.2010.04.012*.

[10] Canoy D, Cairns BJ, Balkwill A, Wright FL, Green J, Reeves G, et al. Body mass index and incident coronary heart disease in women: a population-based prospective study. BMC Med. 2013 Apr;11(1):87. Available from: *https://doi.org/10.1186/1741-7015-11-87*.

[11] Parkin L, Sweetland S, Balkwill A, Green J, Reeves G, Beral V. Body mass index, surgery, and risk of venous thromboembolism in middle-aged women: a cohort study.

Circulation. 2012 Apr;125:1897–904. Available from: *https://doi.org/10.1161/-circulationaha.111.063354*.

[12] Kroll ME, Green J, Beral V, Sudlow CLM, Brown A, Kirichek O, et al. Adiposity and ischemic and hemorrhagic stroke. Neurology. 2016 Oct;87(14):1473–1481. Available from: *https://doi.org/10.1212/WNL.0000000000003171*.

[13] Price AJ, Wright FL, Green J, Balkwill A, Kan SW, Yang TO, et al. Differences in risk factors for 3 types of stroke. Neurology. 2018 Jan;90(4):e298–e306. Available from: *https://doi.org/10.1212/WNL.0000000000004856*.

[14] Pirie K, Peto R, Reeves GK, Green J, Beral V. The 21st century hazards of smoking and benefits of stopping: a prospective study of one million women in the UK. The Lancet. 2013 Jan;381(9861):133–141. Available from: *https://doi.org/10.1016/S0140-6736(12)61720-6*.

[15] Sweetland S, Parkin L, Balkwill A, Green J, Reeves G, Beral V. Smoking, surgery, and venous thromboembolism risk in women: United Kingdom cohort study. Circulation. 2013 Mar;127(12):1276–82. Available from: *https://doi.org/10.1161/-circulationaha.113.001428*.

[16] Spencer EA, Pirie KL, Stevens RJ, Beral V, Brown A, Liu B, et al. Diabetes and modifiable risk factors for cardiovascular disease: the prospective Million Women Study. Eur J Epidemiol. 2008 Nov;23(12):793. Available from: *https://doi.org/10.1007/s10654-008-9298-3*.

[17] Alpert MA, Lavie CJ, Agrawal H, Aggarwal KB, Kumar SA. Obesity and heart failure: epidemiology, pathophysiology, clinical manifestations, and management. Translational Research. 2014 Apr;164(4):345–356. Available from: *https://doi.org/-10.1016/j.trsl.2014.04.010*.

[18] Kenchaiah S, Evans JC, Levy D, Wilson PWF, Benjamin EJ, Larson MG, et al. Obesity and the Risk of Heart Failure. N Engl J Med. 2002 Aug;347(5):305–313. Available from: *https://doi.org/10.1056/NEJMoa020245*.

[19] Khan SS, Ning H, Wilkins JT, Allen N, Carnethon M, Berry JD, et al. Association of Body Mass Index With Lifetime Risk of Cardiovascular Disease and Compression of Morbidity. JAMA Cardiol. 2018 Apr;3(4):280–287. Available from: *https://doi.org/-10.1001/jamacardio.2018.0022*.

[20] Pujades-Rodriguez M, George J, Shah AD, Rapsomaniki E, Denaxas S, West R, et al. Heterogeneous associations between smoking and a wide range of initial presentations of cardiovascular disease in 1937360 people in England: lifetime risks and implications for risk prediction. International journal of epidemiology. 2015 Feb;44:129–41. Available from: *https://doi.org/10.1093/ije/dyu218*.

[21] Aune D, Schlesinger S, Norat T, Riboli E. Tobacco smoking and the risk of heart failure: A systematic review and meta-analysis of prospective studies. Eur J Prev Cardiol. 2019 Feb;26(3):279–288. Available from: *https://doi.org/10.1177/2047487318806658*.

[22] Marott JL, Skielboe AK, Dixen U, Friberg JB, Schnohr P, Jensen GB. Increasing population height and risk of incident atrial fibrillation: the Copenhagen City Heart Study. Eur Heart J. 2018 Dec;39(45):4012–4019. Available from: *https://doi.org/10.1093/-eurheartj/ehy367*.

[23] Persson CE, Adiels M, Björck L, Rosengren A. Young women, body size and risk of atrial fibrillation. Eur J Prev Cardiol. 2018 Jan;25(2):173–180. Available from: *https://-doi.org/10.1177/2047487317740644*.

[24] Andersen K, Rasmussen F, Neovius M, Tynelius P, Sundström J. Body size and risk of atrial fibrillation: a cohort study of 1.1 million young men. J Intern Med. 2018 Apr;283(4):346–355. Available from: *https://doi.org/10.1111/joim.12717*.

[25] Park YM, Moon J, Hwang IC, Lim H, Cho B. Height is associated with incident atrial fibrillation in a large Asian cohort. International Journal of Cardiology. 2020;304:82–84. Available from: *https://doi.org/10.1016/j.ijcard.2020.01.017*.

[26] Schmidt M, Bøtker HE, Pedersen L, Sørensen HT. Adult height and risk of ischemic heart disease, atrial fibrillation, stroke, venous thromboembolism, and premature death: a population based 36-year follow-up study. European Journal of Epidemiology. 2014;29(2):111–118. Available from: *https://doi.org/10.1007/s10654-013-9867-y*.

[27] Rosenberg MA, Patton KK, Sotoodehnia N, Karas MG, Kizer JR, Zimetbaum PJ, et al. The impact of height on the risk of atrial fibrillation: the Cardiovascular Health Study. Eur Heart J. 2012 Nov;33(21):2709–2717. Available from: *https://doi.org/-10.1093/eurheartj/ehs301*.

[28] Lai FY, Nath M, Hamby SE, Thompson JR, Nelson CP, Samani NJ. Adult height and risk of 50 diseases: a combined epidemiological and genetic analysis. BMC Medicine. 2018 Oct;16(1):187. Available from: *https://doi.org/10.1186/s12916-018-1175-7*.

[29] Levin MG, Judy R, Gill D, Vujkovic M, Verma SS, Bradford Y, et al. Genetics of height and risk of atrial fibrillation: A Mendelian randomization study. PLOS Medicine. 2020 Oct;17(10):e1003288. Available from: *https://doi.org/10.1371/-journal.pmed.1003288*.

[30] Larsson SC, Drca N, Wolk A. Alcohol consumption and risk of atrial fibrillation: a prospective study and dose-response meta-analysis. Journal of the American College of Cardiology. 2014 Jul;64:281–9. Available from: *https://doi.org/10.1016/-j.jacc.2014.03.048*.

[31] Gallagher C, Hendriks JML, Elliott AD, Wong CX, Rangnekar G, Middeldorp ME, et al. Alcohol and incident atrial fibrillation - A systematic review and meta-analysis. International journal of cardiology. 2017 Nov;246:46–52. Available from: *https://doi.org/-10.1016/j.ijcard.2017.05.133*.

[32] Conen D, Tedrow UB, Cook NR, Moorthy MV, Buring JE, Albert CM. Alcohol consumption and risk of incident atrial fibrillation in women. JAMA. 2008 Dec;300:2489–96. Available from: *https://dx.doi.org/10.1001/jama.2008.755*.

[33] Csengeri D, Sprünker NA, Di Castelnuovo A, Niiranen T, Vishram-Nielsen JK, Costanzo S, et al. Alcohol consumption, cardiac biomarkers, and risk of atrial fibrillation and adverse outcomes. European heart journal. 2021 Mar;42:1170–1177. Available from: *https://doi.org/10.1093/eurheartj/ehaa953*.

[34] Reynolds K, Lewis B, Nolen JDL, Kinney GL, Sathya B, He J. Alcohol consumption and risk of stroke: a meta-analysis. JAMA. 2003 Feb;289:579–88. Available from: *https://doi.org/10.1001/jama.289.5.579*.

[35] Larsson SC, Wallin A, Wolk A, Markus HS. Differing association of alcohol consumption with different stroke types: a systematic review and meta-analysis. BMC Medicine. 2016;14(1):178. Available from: *https://doi.org/10.1186/s12916-016-0721-4*.

[36] Christensen AI, Nordestgaard BG, Tolstrup JS. Alcohol Intake and Risk of Ischemic and Haemorrhagic Stroke: Results from a Mendelian Randomisation Study. Journal of stroke. 2018 May;20(29886720):218–227. Available from: *https://dx.doi.org/-10.5853/jos.2017.01466*.

[37] Larsson SC, Burgess S, Mason AM, Michaëlsson K. Alcohol Consumption and Cardiovascular Disease. Circulation: Genomic and Precision Medicine. 2020 Jun;13(3):e002814. Available from: *https://doi.org/10.1161/CIRCGEN.119.002814*.

[38] Larsson SC, Wallin A, Wolk A. Alcohol consumption and risk of heart failure: Meta-analysis of 13 prospective studies. Clinical nutrition (Edinburgh, Scotland). 2018 Aug;37:1247–1251. Available from: *https://doi.org/10.1016/j.clnu.2017.05.007*.

[39] van Oort S, Beulens JWJ, van Ballegooijen AJ, Handoko ML, Larsson SC. Modifiable lifestyle factors and heart failure: A Mendelian randomization study. American Heart Journal. 2020 Sep;227:64–73. Available from: *https://doi.org/10.1016/-j.ahj.2020.06.007*.

[40] Jiang Q, Wang K, Shi J, Li M, Chen M. No association between alcohol consumption and risk of atrial fibrillation: A two-sample Mendelian randomization study. Nutrition, metabolism, and cardiovascular diseases : NMCD. 2020 Jul;30:1389–1396. Available from: *https://doi.org/10.1016/j.numecd.2020.04.014*.

[41] Millwood IY, Walters RG, Mei XW, Guo Y, Yang L, Bian Z, et al. Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500000 men and women in China. The Lancet. 2019 May;393(10183):1831–1842. Available from: *https://doi.org/10.1016/S0140-6736(18)31772-0*.

[42] Holmes MV, Dale CE, Zuccolo L, Silverwood RJ, Guo Y, Ye Z, et al. Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data. BMJ. 2014 Jul;349:g4164. Available from: *https://doi.org/-10.1136/bmj.g4164*.

[43] Agha M, Agha R. The rising prevalence of obesity: part A: impact on public health. International journal of surgery Oncology. 2017 Aug;2(29177227):e17–e17. Available from: *https://dx.doi.org/10.1097/IJ9.0000000000000017*.

[44] Jackson SE, Beard E, Kujawski B, Sunyer E, Michie S, Shahab L, et al. Comparison of Trends in Self-reported Cigarette Consumption and Sales in England, 2011 to 2018. JAMA network open. 2019 Aug;2(31461148):e1910161–e1910161. Available from: *https://dx.doi.org/10.1001/jamanetworkopen.2019.10161*.

[45] Simpson CR, Hippisley-Cox J, Sheikh A. Trends in the epidemiology of smoking recorded in UK general practice. The British journal of general practice : the journal of the Royal College of General Practitioners. 2010 Mar;60(20202355):e121–e127. Available from: *https://doi.org/10.3399/bjgp10X483544*.

[46] O'Keeffe AG, Nazareth I, Petersen I. Time trends in the prescription of statins for the primary prevention of cardiovascular disease in the United Kingdom: a cohort study

using The Health Improvement Network primary care data. Clinical epidemiology. 2016 May;8(27313477):123–132. Available from: *https://doi.org/10.2147/clep.s104258*.

[47] Wright FL, Green J, Reeves G, Beral V, Cairns BJ, on behalf of the Million Women Study collaborators. Validity over time of self-reported anthropometric variables during follow-up of a large cohort of UK women. BMC Med Res Methodol. 2015 Oct;15:81. Available from: *https://doi.org/10.1186/s12874-015-0075-1*.

[48] Peter RS, Fromm E, Klenk J, Concin H, Nagel G. Change in height, weight, and body mass index: longitudinal data from Austria. American journal of human biology : the official journal of the Human Biology Council. 2014 Sep;26:690–6. Available from: *https://doi.org/10.1002/ajhb.22582*.

[49] Drøyvold WB, Nilsen TIL, Krüger O, Holmen TL, Krokstad S, Midthjell K, et al. Change in height, weight and body mass index: Longitudinal data from the HUNT Study in Norway. International journal of obesity (2005). 2006 Jun;30:935–9. Available from: *https://doi.org/10.1038/sj.ijo.0803178*.

[50] Kuczmarski MF, Kuczmarski RJ, Najjar M. Effects of age on validity of self-reported height, weight, and body mass index: findings from the Third National Health and Nutrition Examination Survey, 1988-1994. Journal of the American Dietetic Association. 2001 Jan;101:28–34. Available from: *https://doi.org/10.1016/S0002-8223(01)00008-6*.

[51] Armstrong MEG, Cairns BJ, Green J, Reeves GK, Beral V, the Million Women Study Collaborators. Reported frequency of physical activity in a large epidemiological study: relationship to specific activities and repeatability over time. BMC Med Res Methodol. 2011 Jun;11:97. Available from: *https://doi.org/10.1186/1471-2288-11-97*.

[52] Petrovic D, de Mestral C, Bochud M, Bartley M, Kivimäki M, Vineis P, et al. The contribution of health behaviors to socioeconomic inequalities in health: A systematic review. Preventive Medicine. 2018 Aug;113:15–31. Available from: *https://doi.org/-10.1016/j.ypmed.2018.05.003*.

[53] Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, et al. Agreement between general practice prescription data and self-reported use of hormone replacement therapy and treatment for various illnesses. J Epidemiol Biostat. 2001 Feb;6:357–363. Available from: *https://doi.org/10.1080/13595220152601837*.

6 Socioeconomic status and risk of incident cardiovascular multimorbidity

6.1 Introduction

6.1.1 Background

Lower socioeconomic status (SES) has been consistently associated with a higher risk of cardiovascular disease (CVD) in various populations. [1-4] Lower SES has also been associated with a higher prevalence of modifiable cardiovascular risk factors including smoking, obesity, and high blood pressure [5-7], and these factors are thought to partly explain the social gradient in CVD risk [8]. In fact, previous work in the Million Women Study found that more than 70% of the social gradient in coronary heart disease incidence was potentially accounted for by four health-related behavioural factors: smoking, alcohol consumption, physical inactivity, and body mass index (BMI). [9]

In contrast, few prospective studies have examined the relationship between SES and multimorbidity [10, 11], and none have specifically studied the possible social gradient in the risk of cardiovascular multimorbidity (CVM) as defined in this thesis. It is expected that the relationship between SES and CVM will be mediated and/or confounded by known cardiovascular risk factors (as has been the case for individual CVDs), but it is not clear to what extent. Furthermore, although reproductive factors may play a role in women's CVD risk [12, 13], whether they may be able to account for any relationship between SES and CVM has not been investigated. Confirming the potential mediation of SES by health behaviours and other factors will be important for developing public health policies aimed at reducing any social inequalities in CVM risk.

6.1.2 Research questions

Using data from the Million Women Study (described in Chapter 3), this chapter addresses the following research questions:

- What are the prospective associations of socioeconomic factors with CVM and complex CVM?
- What are the prospective associations of socioeconomic factors with selected pairs of CVDs (CVM pairs) that comprise CVM?
- How much of the associations between socioeconomic factors and CVM outcomes (i.e. CVM, complex CVM, and selected CVM pairs) are accounted for by known cardiovascular risk factors and reproductive factors?
- Which, if any, individual CVDs potentially drive any observed associations between socioeconomic factors and CVM?

6.2 Methods

This section mainly describes the methods that are specific to the analyses presented in this chapter. Methods that are common to multiple chapters are described only briefly; further details on such methods are available in Section 3.3 of Chapter 3 (page 79).

6.2.1 Outcomes

The outcomes investigated in this chapter were CVM, complex CVM, and selected CVM pairs. The selection and definition of outcomes have been described in detail in Section 3.3.1 of Chapter 3 (page 79). Briefly, CVM and complex CVM were defined as an individual developing ≥2 or ≥4 of the selected 19 CVDs (see Table 3.3 in Chapter 3, page 81), respectively. The selected pairs were, in descending order of frequency: 'Ischaemic heart disease & Atrial fibrillation' (IHD-AF), 'Ischaemic heart disease & Heart failure' (IHD-HF), 'Ischaemic heart disease & Other Arrhythmias' (IHD-ARR), 'Atrial fibrillation & Heart failure' (AF-HF), 'Other arrhythmias & Heart failure' (ARR-HF), and 'Atrial fibrillation & Stroke' (AF-STR).

6.2.2 Risk factors

The two risk factors of primary interest in this chapter were measures of SES, namely the level of attained education and area deprivation level. Risk factor data were collected through the self-reported recruitment questionnaire (Appendix B, page 298).

Participants were asked about their educational qualifications with the following question: "What qualification(s) do you have from school, college or the equivalent? (please put a cross in the most appropriate box(es))." Women could select the following boxes: (1) clerical or commercial qualifications (e.g. secretarial, hairdressing etc), (2) nursing or teaching, (3) "O" level (or equivalent), (4) "A" level (or equivalent), (5) college/university

degree (or equivalent), (6) none of these. Women also reported the age when they finished full-time schooling.

The education variable was categorised into five groups, according to each participant's highest educational attainment: (i) did not complete compulsory schooling (no qualifications), (ii) completed compulsory schooling (no qualifications), (iii) technical qualifications (nursing, teaching, clerical, or commercial), (iv) secondary qualifications (A levels or O levels), and (v) tertiary qualifications (college or university) [referent]. Whether women completed compulsory schooling was determined using their reported date of birth, age at finishing full-time schooling, and the compulsory school leaving age at the time. The Education Act 1944 (in effect from 1 April 1947) increased the compulsory school leaving age in England from 14 to 15 years old during the time when some of the study participants were at school.

The Townsend index of area deprivation has been a standard indicator of socioeconomic deprivation in the UK in the past three decades and has been widely used in population health research. [14-16] It is calculated using census information on the percentage of unemployment, home and car ownership, and overcrowding in relevant geographical units (census enumeration district in England and census output area in Scotland). [14] In the Million Women Study, participants reported their postcode of residence in the recruitment questionnaire, which was used to derive the Townsend score for each individual at the time. The scores were categorised into fifths before any exclusions were made to the cohort. The least deprived fifth was the reference group in the analyses in this chapter.

It should be noted that participants of the Million Women Study were less likely to live in socioeconomically deprived areas compared to non-participants. This is because on average, attenders of breast screening (via which Million Women Study participants were

recruited) were less likely to be from deprived areas than non-attenders. In a sample of 1,064 women invited to breast screening from two general practices, 83% (n=882) had attended screening, and 30% of attenders versus 50% of non-attenders came from the most deprived areas. [17] Among those who attended screening, there were no differences in area deprivation level between Million Women Study participants and non-participants. There were no significant differences in age or prescriptions for common illnesses including hypertension, heart disease, hypercholesterolaemia, and diabetes mellitus between attenders and non-attenders of breast screening. This indicates that the differences in area deprivation level in Million Women Study participants compared to non-participants were unlikely to have resulted in substantial differences in CVD incidence between them.

6.2.3 Statistical analyses

6.2.3.1 Cox proportional hazards regression

Cox proportional hazards regression with attained age as the underlying time variable was used to calculate hazards ratios (HRs) and 95% confidence intervals (CIs) for the CVM outcomes in relation to the SES measures. Section 3.3.4.1 of Chapter 3 (page 85) describes this method in more detail.

Minimally adjusted models were stratified by age at recruitment (<53, 53-55, 56-58, 59-61, 62-64, 65+ years) and recruitment region (ten geographic regions: Oxford, East Anglia, South West, Thames, West Midlands, North York, Trent, North West [Mersey], North West [Manchester/Lancaster], Scotland).

Multivariable adjusted models were additionally adjusted for BMI (<20, 20-24, 25-29, 30-34, 35+ kg/m^2), height (<155, 155-159, 160-164, 165-169, 170+ cm), smoking (never, past, current <15 per day, current 15+ per day), alcohol consumption (rarely/never, 1-6, 7-

14, 15+ units per week), strenuous physical activity (none, up to once per week, 2-3 times per week, 4+ times per week), parity (nulliparous, 1-2, 3+), and ever use of menopausal hormone therapy (yes or no). Multivariable adjusted models for educational attainment and area deprivation were adjusted for each other, as they were only weakly correlated (Spearman's rho=0.23). For each adjustment variable, missing values (<6% for each variable) were assigned to a separate category.

Differences in HRs across the categories of each risk factor were tested using the likelihood ratio test for heterogeneity, with a significance level of $\alpha = 0.05$. The Benjamini-Hochberg procedure [18] (see Section 3.3.4.1 of Chapter 3, page 85) was applied to control the false discovery rate to account for multiple testing in this chapter. Log-log plots were assessed to check the proportional hazards assumption, which was found to be satisfied in all models.

6.2.3.2 Likelihood-ratio χ^2 test statistic attenuation

The likelihood-ratio χ^2 test (LR) statistic attenuation method [9, 19] was used to measure the extent to which the relationships between socioeconomic factors and CVM outcomes were accounted for by other variables. Section 3.3.4.2 of Chapter 3 (page 86) describes this method in more detail.

6.2.3.3 Sensitivity analyses

Three sensitivity analyses of the relationship between the two SES measures and CVM were conducted. First, to assess whether some individual CVDs may be driving any of the associations between SES and CVM, the relationship of education and area deprivation with six common individual CVDs and then with CVM minus each of the CVDs were examined. The six individual CVDs were ischaemic heart disease, atrial fibrillation, heart failure, other arrhythmias, stroke, and venous thromboembolism. The rationale for the

selection of these CVDs and the definition of CVM minus each CVD are described in Section 5.2.3.2 of Chapter 5 (page 145). Second, sensitivity analysis excluding women who were treated for diabetes or hypertension at recruitment was conducted to assess whether these conditions could be substantial confounders and/or mediators of the observed associations between SES and CVM. Diabetes and hypertension are potential pathways through which area deprivation and education may affect CVM, so risk estimates may be attenuated if those pathways are accounted for by removing the effects of those conditions. Finally, sensitivity analysis restricting to women who were never smokers at recruitment was conducted to assess the potential for residual confounding by smoking.

6.3 Results

6.3.1 Baseline characteristics by socioeconomic status

6.3.1.1 Educational attainment

Baseline characteristics of the study participants by educational attainment are presented in Table 6.1. Approximately 13% (n=163,123) of women had tertiary educational qualifications, while 42% (n=520,068) had only received compulsory schooling or less (i.e. no qualifications). On average, women with no educational qualifications were older and more likely to live in the most deprived areas compared to women with qualifications. Women with no educational qualifications were more likely to be current smokers, non-drinkers of alcohol, had a higher BMI, and shorter height. They also had a higher number of full-term pregnancies (i.e. parity) on average, and were less likely to have breastfed children or used hormone replacement therapy. Women with no educational qualifications were also more likely to report treatment for hypertension, diabetes, and high blood cholesterol than those with qualifications.

Table 6.1 Baseline characteristics of study participants by educational attainment

Educational attainment				**No qualifications**	
	Tertiary	**Secondary**	**Technical**	**Compulsory schooling**	**< Compulsory schooling**
Women, n	163,123	321,451	203,018	492,035	28,033
Age at recruitment, mean (SD), years	55.9 (4.7)	55.9 (4.6)	56.6 (4.8)	56.9 (4.8)	57.9 (4.9)
Most deprived fifth % (n)	10.0 (16,249)	11.6 (37,051)	14.1 (28,452)	27.6 (134,814)	34.9 (9,725)
Lifestyle and anthropometric factors					
Current smoker, % (n)	11.2 (17,475)	15.3 (47,336)	17.8 (34,292)	27.0 (123,801)	28.0 (7,221)
Drinker of alcoholic beverages, % (n)	77.6 (126,166)	72.9 (233,708)	67.4 (136,228)	56.2 (274,260)	47.7 (13,192)
No strenuous physical activity, % (n)	48.1 (77,224)	53.8 (169,843)	59.0 (116,221)	68.7 (321,622)	72.0 (18,946)
Body mass index, mean (SD), kg/m^2	25.3 (4.3)	25.7 (4.4)	26.0 (4.4)	26.7 (4.8)	26.9 (5.0)
Height, mean (SD), cm	163.5 (6.5)	162.9 (6.4)	162.1 (6.5)	161.2 (6.6)	160.5 (6.9)
Reproductive factors					
Age at menarche, mean (SD), years	12.9 (1.5)	12.9 (1.5)	13.0 (1.6)	13.1 (1.6)	13.2 (1.7)
Age at first birth, mean (SD), years	26.2 (4.4)	24.6 (4.2)	24.3 (4.2)	22.7 (4.0)	22.7 (4.2)
Age at natural menopause, mean (SD), years	49.5 (4.0)	49.4 (4.1)	49.2 (4.2)	48.9 (4.4)	48.6 (4.8)
Pre/perimenopausal, % (n)	17.3 (28,118)	15.6 (49,868)	12.7 (25,635)	10.5 (51,348)	9.0 (2,481)
Parity, mean (SD)	1.9 (1.2)	2.0 (1.1)	2.1 (1.1)	2.3 (1.3)	2.6 (1.6)
Ever breastfed children in parous women, % (n)	82.2 (94,474)	73.2 (172,504)	66.7 (103,690)	60.1 (199,939)	65.7 (11,617)
Ever use of contraceptive pills, % (n)	69.0 (112,104)	65.9 (210,969)	59.9 (120,809)	53.9 (261,282)	45.8 (12,514)
Ever use of hormone replacement therapy, % (n)	51.3 (83,205)	52.1 (166,520)	52.1 (104,949)	48.8 (236,479)	44.5 (12,175)
Healthcare factors					
Self-reported treatment for hypertension, % (n)	11.1 (18,174)	12.7 (40,928)	14.8 (30,119)	16.4 (80,470)	17.7 (4,961)
Self-reported treatment for diabetes mellitus, % (n)	1.4 (2,231)	1.6 (4,984)	1.8 (3,720)	2.6 (12,595)	4.1 (1,148)
Self-reported treatment for high cholesterol, % (n)	1.8 (3,000)	2.1 (6,624)	2.9 (5,794)	3.0 (14,934)	3.5 (984)

SD = standard deviation, CVM = cardiovascular multimorbidity

6.3.1.2 Area deprivation

Baseline characteristics of study participants by level of area deprivation are displayed in Table 6.2. There was no difference in mean age by deprivation level. Women living in more deprived areas were less likely to have any educational qualifications beyond compulsory schooling than those living in more affluent areas. They were more likely to be current smokers, non-drinkers, physically inactive, higher in BMI, and shorter in stature. Women living in more deprived areas also had higher parity and were less likely to have breastfed children or used menopausal hormone replacement therapy. They were also more likely to report treatment for hypertension, diabetes, and high blood cholesterol than those living in more affluent areas.

Table 6.2 Baseline characteristics of participant by area deprivation level

Deprivation level	Fifths of the Townsend index of area deprivation				
	1 (Least deprived)	**2**	**3**	**4**	**5 (Most deprived)**
Women, n	252,383	250,544	247,443	244,832	236,031
Age at recruitment, mean (SD), years	56.4 (4.7)	56.5 (4.8)	56.6 (4.8)	56.6 (4.8)	56.5 (4.8)
No educational qualifications, % (n)	29.1 (72,188)	35.5 (87,119)	40.4 (97,568)	48.5 (115,248)	63.9 (144,539)
Lifestyle and anthropometric factors					
Current smoker, % (n)	13.1 (31,228)	15.6 (36,902)	18.1 (42,148)	22.9 (52,802)	33.5 (73,901)
Drinker of alcoholic beverages, % (n)	72.0 (180,885)	69.1 (172,278)	66.1 (162,635)	62.0 (150,455)	55.1 (128,204)
No strenuous physical activity, % (n)	54.7 (134,820)	56.7 (137,962)	58.6 (140,225)	62.9 (147,782)	70.2 (156,533)
Body mass index, mean (SD), kg/m^2	25.6 (4.2)	25.8 (4.3)	26.0 (4.5)	26.4 (4.7)	27.0 (5.2)
Height, mean (SD), cm	162.6 (6.4)	162.5 (6.5)	162.3 (6.6)	161.9 (6.6)	161.1 (6.8)
Reproductive factors					
Age at menarche, mean (SD), years	13.0 (1.5)	13.0 (1.6)	13.0 (1.6)	13.0 (1.6)	13.1 (1.7)
Age at first birth, mean (SD), years	24.8 (4.2)	24.4 (4.2)	24.1 (4.3)	23.6 (4.3)	22.6 (4.1)
Age at natural menopause, mean (SD), years	49.4 (4.1)	49.4 (4.1)	49.3 (4.2)	49.0 (4.3)	48.6 (4.5)
Pre/perimenopausal, % (n)	14.3 (35,865)	13.7 (34,041)	13.3 (32,642)	12.6 (30,569)	11.4 (26,662)
Parity, mean (SD)	2.0 (1.1)	2.1 (1.1)	2.1 (1.2)	2.2 (1.3)	2.4 (1.5)
Ever breastfed children in parous women, % (n)	70.6 (128,790)	70.6 (127,115)	69.8 (123,213)	66.6 (113,797)	61.1 (97,098)
Ever use of contraceptive pills, % (n)	62.0 (155,295)	60.6 (150,705)	59.9 (146,819)	59.0 (142,651)	57.4 (133,074)
Ever use of hormone replacement therapy, % (n)	51.5 (128,830)	51.0 (126,525)	50.4 (123,530)	49.8 (120,467)	49.0 (113,725)
Healthcare factors					
Self-reported treatment for hypertension, % (n)	12.8 (32,258)	13.6 (34,094)	14.2 (35,193)	15.2 (37,234)	16.9 (39,783)
Self-reported treatment for diabetes mellitus, % (n)	1.4 (3,617)	1.6 (4,072)	1.9 (4,677)	2.3 (5,573)	3.2 (7,527)
Self-reported treatment for high cholesterol, % (n)	2.2 (5,542)	2.4 (5,913)	2.5 (6,191)	2.7 (6,607)	3.3 (7,796)

SD = standard deviation, CVM = cardiovascular multimorbidity

6.3.2 Risk of cardiovascular multimorbidity and complex cardiovascular multimorbidity in relation to socioeconomic status

Figure 6.1 shows the minimally and multivariable adjusted HRs and 95% floated confidence intervals (FCIs) for incident CVM and complex CVM by categories of educational attainment and area deprivation level. All associations were statistically significant after controlling for the false discovery rate ($p<0.01$). Risks of both CVM and complex CVM increased with lower educational attainment and higher area deprivation level. Although the associations remained statistically significant, HRs for both CVM outcomes were greatly attenuated after multivariable adjustment ($\chi^2_{attenuation}$ =86-89%).

With minimal adjustment (only stratifying by age and region), women who did not complete compulsory schooling had almost twice the risk of CVM compared to women with tertiary qualifications (HR 1.92, 95% CI 1.85-1.99). After multivariable adjustment, the HR was attenuated to 1.36 (95% CI 1.31-1.42) [$\chi^2_{attenuation}$ =89%]. Similarly, compared to women living in the least deprived areas, women living in the most deprived areas had a 76% higher risk of CVM (1.76, 1.73-1.79) with minimal adjustment. After multivariable adjustment, the HR was attenuated to 1.27 (1.25-1.30) [$\chi^2_{attenuation}$ =86%].

The relative risks for complex CVM in relation to SES were greater in magnitude than those for CVM. With minimal adjustment, women who did not complete compulsory schooling had more than twice the risk of developing complex CVM compared to women with tertiary qualifications (2.27, 2.11-2.44), which was attenuated to a 50% higher risk after multivariable adjustment (1.51, 1.39-1.64) [$\chi^2_{attenuation}$ =87%]. Compared to women living in the least deprived areas, women living in the most deprived areas had a minimally adjusted HR of 1.97 (1.90-2.05) for complex CVM, which was attenuated to a HR of 1.30 (1.24-1.36) after multivariable adjustment ($\chi^2_{attenuation}$ =88%).

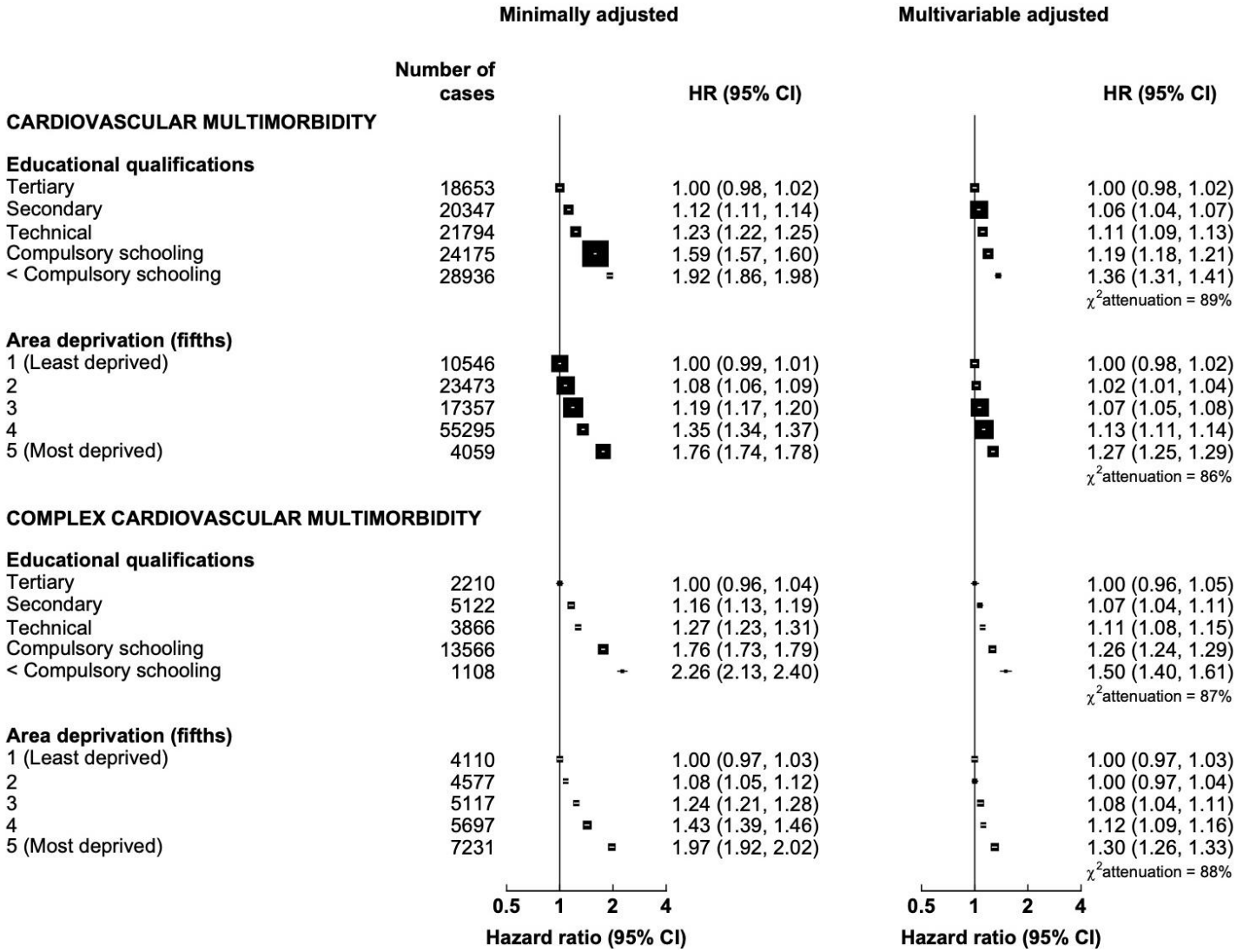


Figure 6.1 Relative risks of incident CVM and complex CVM in relation to education and area deprivation

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation or educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed.

6.3.3 Risk of selected cardiovascular multimorbidity pairs in relation to socioeconomic status

Figure 6.2 and Figure 6.3 show the multivariable adjusted HRs and 95% FCIs for the six selected pairs of CVDs (i.e. CVM pairs) in relation to education and area deprivation level. Both socioeconomic factors were statistically significantly associated with each CVM pair after controlling for the false discovery rate ($p<0.01$). Figure 6.2 displays the associations for the three most common CVM pairs – all of which included ischaemic heart disease – which in descending number of cases are: 'Ischaemic heart disease & Atrial fibrillation' (IHD-AF), 'Ischaemic heart disease & Heart failure' (IHD-HF), 'Ischaemic heart disease & Other Arrhythmias' (IHD-ARR). Figure 6.3 displays socioeconomic factor associations for the three other selected CVM pairs, all of which included cardiac arrhythmias: 'Atrial fibrillation & Heart failure' (AF-HF), 'Other arrhythmias & Heart failure' (ARR-HF), and 'Atrial fibrillation & Stroke' (AF-STR).

The HRs for CVM pairs were often greater in magnitude than those for CVM, and similar to those for complex CVM (see Section 6.3.2, page 195). When stratifying by age and region only, women in the lowest SES level had 1.5-3 times the risk of each CVM pair compared to those in the highest level. After multivariable adjustment, the attenuation compared to the minimally adjusted model was very substantial ($\chi^2_{attenuation}$ =73-91%). Women with the lowest level of education had a 42-58% higher risk of each CVM pair compared to those with the highest level of education, while women living in the most deprived areas had a 23-42% higher risk of each CVM pair compared to women living in the least deprived areas.

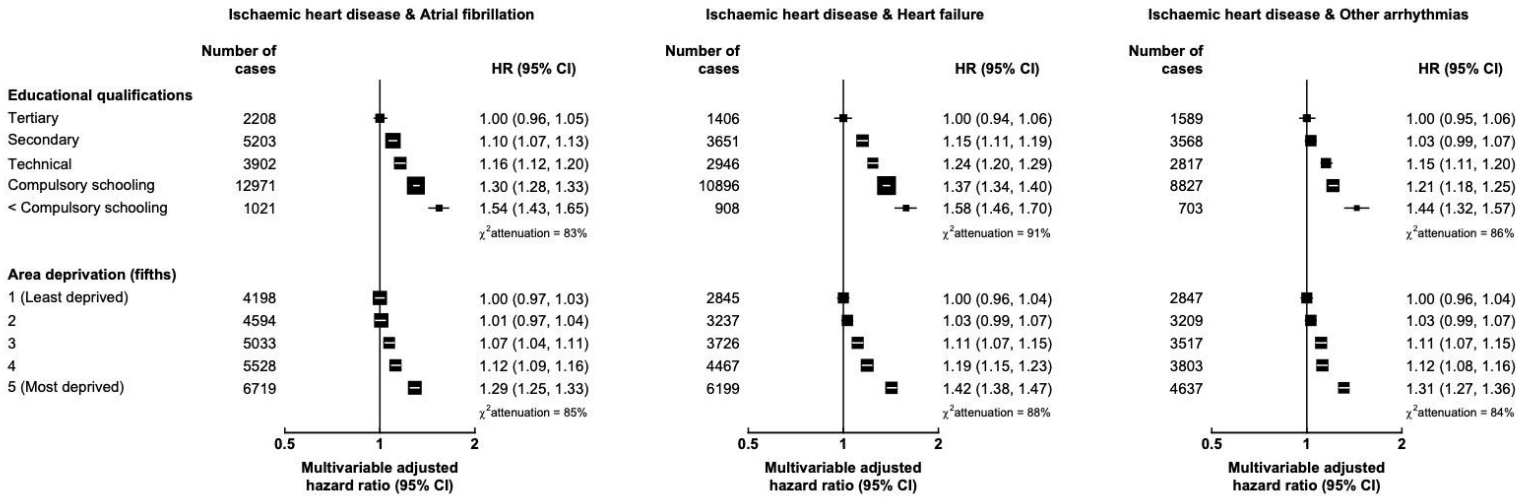


Figure 6.2 Relative risks of CVM pairs including ischaemic heart disease in relation to education and area deprivation

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation or educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed.

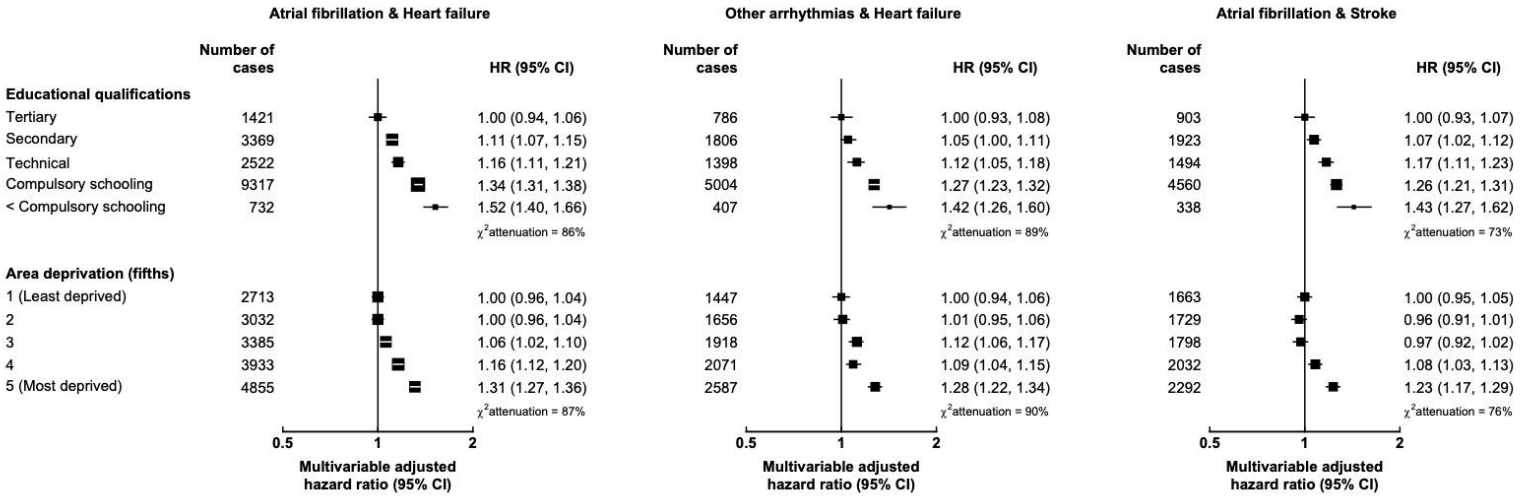


Figure 6.3 Relative risks of other CVM pairs including cardiac arrhythmias in relation to education and area deprivation

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation or educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed.

6.3.4 The role of covariates in accounting for the social gradient in the risk of cardiovascular multimorbidity outcomes

Overall, the associations between SES and all CVM outcomes were greatly attenuated after multivariable adjustment ($\chi^2_{attenuation}$ =73-91%, Figure 6.1-Figure 6.3, page 196-9). The role of potential mediators and confounders in these associations was explored in more detail by assessing the changes in the LR statistics after adjusting for covariates, both individually and in different combinations.

6.3.4.1 Cardiovascular multimorbidity and complex cardiovascular multimorbidity

6.3.4.1.1 The role of covariates in the educational gradient in risk

Table 6.3 presents the relationship of educational attainment with the risk of CVM and complex CVM after adjusting for area deprivation, anthropometric, behavioural, and reproductive factors in addition to minimal adjustment (stratifying by age and region).

The attenuation in the LR statistic compared to the minimally adjusted model indicated that additionally adjusting for area deprivation alone accounted for about 40% of the relationship between education and the CVM outcomes. Adjusting for smoking alone resulted in the greatest attenuation in the LR statistics ($\chi^2_{attenuation}$ =45% for CVM and 43% for complex CVM). In contrast, adjusting individually for height, parity, and menopausal hormone therapy had smaller effects on the LR statistics ($\chi^2_{attenuation}$≤9%).

Adjusting simultaneously for four health-related behavioural factors (i.e. smoking, alcohol consumption, physical activity, and BMI) accounted for over 80% of the associations between education and the CVM outcomes ($\chi^2_{attenuation}$ =85% for CVM and 83% for complex CVM). Adjusting for the four health behavioural factors plus area deprivation accounted for approximately 90% of the associations ($\chi^2_{attenuation}$ =91% for CVM and 89% for complex CVM).

Table 6.3 Relationship of education with risk of CVM and complex CVM after adjustment for various covariates

Education	LR[1]	attenuation in LR[1] (%)	Tertiary	Secondary	Technical	Compulsory schooling	< Compulsory schooling
INCIDENT CVM (n)			10,546	23,473	17,357	55,295	4,059
HR (95% CI), adjusted for,							
Age, region (minimally adjusted)	3803	-	1.00 (ref.)	1.12 (1.10 - 1.15)	1.23 (1.20 - 1.26)	1.59 (1.55 - 1.62)	1.92 (1.85 - 1.99)
Age, region, deprivation	2245	41	1.00 (ref.)	1.11 (1.09 - 1.14)	1.21 (1.18 - 1.24)	1.46 (1.43 - 1.49)	1.71 (1.65 - 1.77)
Age, region, smoking	2081	45	1.00 (ref.)	1.09 (1.07 - 1.12)	1.18 (1.15 - 1.21)	1.42 (1.39 - 1.45)	1.69 (1.63 - 1.76)
Age, region, alcohol	2854	25	1.00 (ref.)	1.11 (1.09 - 1.14)	1.20 (1.17 - 1.23)	1.51 (1.48 - 1.54)	1.79 (1.73 - 1.86)
Age, region, physical activity	2410	37	1.00 (ref.)	1.09 (1.07 - 1.12)	1.19 (1.16 - 1.22)	1.46 (1.43 - 1.49)	1.75 (1.68 - 1.81)
Age, region, body mass index	2463	35	1.00 (ref.)	1.10 (1.08 - 1.13)	1.20 (1.17 - 1.23)	1.47 (1.44 - 1.50)	1.75 (1.69 - 1.82)
Age, region, height	3800	0	1.00 (ref.)	1.13 (1.10 - 1.16)	1.25 (1.22 - 1.28)	1.61 (1.57 - 1.64)	1.95 (1.88 - 2.02)
Age, region, parity	3505	8	1.00 (ref.)	1.12 (1.10 - 1.15)	1.24 (1.21 - 1.27)	1.57 (1.54 - 1.60)	1.87 (1.80 - 1.94)
Age, region, menopausal hormone therapy	3738	2	1.00 (ref.)	1.12 (1.10 - 1.15)	1.23 (1.20 - 1.26)	1.58 (1.55 - 1.62)	1.93 (1.86 - 2.00)
Age, region, health behaviours[2]	551	85	1.00 (ref.)	1.05 (1.02 - 1.08)	1.10 (1.07 - 1.13)	1.21 (1.18 - 1.24)	1.39 (1.34 - 1.45)
Age, region, health behaviours[2], plus deprivation	347	91	1.00 (ref.)	1.05 (1.02 - 1.07)	1.09 (1.06 - 1.12)	1.17 (1.14 - 1.20)	1.33 (1.28 - 1.39)
Fully adjusted[3]	406	89	1.00 (ref.)	1.06 (1.03 - 1.08)	1.11 (1.08 - 1.14)	1.19 (1.17 - 1.22)	1.36 (1.31 - 1.42)
INCIDENT COMPLEX CVM (n)			2,210	5,122	3,866	13,566	1,108
HR (95% CI), adjusted for,							
Age, region (minimally adjusted)	1344	-	1.00 (ref.)	1.16 (1.10 - 1.22)	1.27 (1.21 - 1.34)	1.76 (1.68 - 1.84)	2.27 (2.11 - 2.44)
Age, region, deprivation	810	40	1.00 (ref.)	1.14 (1.09 - 1.20)	1.24 (1.17 - 1.30)	1.59 (1.52 - 1.66)	1.97 (1.83 - 2.12)
Age, region, smoking	764	43	1.00 (ref.)	1.12 (1.07 - 1.18)	1.21 (1.14 - 1.27)	1.55 (1.48 - 1.63)	1.96 (1.82 - 2.12)
Age, region, alcohol	992	26	1.00 (ref.)	1.14 (1.09 - 1.20)	1.23 (1.17 - 1.30)	1.65 (1.57 - 1.72)	2.08 (1.93 - 2.23)
Age, region, physical activity	881	34	1.00 (ref.)	1.13 (1.07 - 1.18)	1.21 (1.15 - 1.28)	1.60 (1.53 - 1.68)	2.03 (1.88 - 2.19)
Age, region, body mass index	876	35	1.00 (ref.)	1.14 (1.08 - 1.20)	1.22 (1.16 - 1.29)	1.61 (1.53 - 1.68)	2.03 (1.89 - 2.19)
Age, region, height	1364	-1	1.00 (ref.)	1.17 (1.11 - 1.23)	1.29 (1.22 - 1.36)	1.79 (1.71 - 1.88)	2.33 (2.17 - 2.51)
Age, region, parity	1253	7	1.00 (ref.)	1.16 (1.11 - 1.22)	1.28 (1.21 - 1.35)	1.75 (1.67 - 1.83)	2.21 (2.05 - 2.38)
Age, region, menopausal hormone therapy	1287	4	1.00 (ref.)	1.16 (1.10 - 1.22)	1.27 (1.20 - 1.34)	1.75 (1.67 - 1.83)	2.25 (2.09 - 2.42)
Age, region, health behaviours[2]	224	83	1.00 (ref.)	1.07 (1.01 - 1.12)	1.10 (1.04 - 1.16)	1.28 (1.21 - 1.34)	1.56 (1.43 - 1.69)
Age, region, health behaviours[2], plus deprivation	153	89	1.00 (ref.)	1.06 (1.00 - 1.12)	1.09 (1.03 - 1.15)	1.23 (1.17 - 1.29)	1.47 (1.36 - 1.60)
Fully adjusted[3]	171	87	1.00 (ref.)	1.07 (1.02 - 1.13)	1.12 (1.05 - 1.18)	1.26 (1.20 - 1.33)	1.51 (1.39 - 1.64)

[1]Likelihood-ratio test statistic; [2]smoking, alcohol consumption, physical activity, body mass index

[3]adjusted for all health behaviours, deprivation, height, parity, and ever use of menopausal hormone therapy; CVM=cardiovascular multimorbidity; ref. means reference category

6.3.4.1.2 The role of covariates in the gradient in risk by area deprivation

Table 6.4 shows the attenuation in the LR statistics of the minimally adjusted models for the relationship between area deprivation with CVM and complex CVM, after additionally adjusting for various covariates (both separately and simultaneously).

Adjusting separately for education accounted for about 40% of the associations between area deprivation level and both CVM outcomes ($\chi^2_{attenuation}$ =38% for CVM and 40% for complex CVM). Smoking again had the largest individual effect, accounting for almost half of the associations ($\chi^2_{attenuation}$ =48% for CVM and 49% for complex CVM), while reproductive factors and height had little individual effect ($\chi^2_{attenuation}$ =2-11%).

Adjusting for the four health-related behaviours together accounted for 80% or more of the associations ($\chi^2_{attenuation}$ =80% for CVM and 87% for complex CVM). Adjusting for the four health-related behavioural factors plus education accounted for 85% or more of the associations ($\chi^2_{attenuation}$ =85% for CVM and 87% for complex CVM).

Table 6.4 Relationship of area deprivation with risk of CVM and complex CVM after adjustment for various covariates

Area deprivation (fifths)	LR[1]	attenuation in LR[1] (%)	1 (Least deprived)	2	3	4	5 (Most deprived)
INCIDENT CVM (n)			18,653	20,347	21,794	24,175	28,936
HR (95% CI), adjusted for,							
Age, region (minimally adjusted)	4641	-	1.00 (ref.)	1.08 (1.06 - 1.10)	1.19 (1.16 - 1.21)	1.35 (1.33 - 1.38)	1.76 (1.73 - 1.79)
Age, region, education	2875	38	1.00 (ref.)	1.06 (1.04 - 1.08)	1.15 (1.12 - 1.17)	1.27 (1.25 - 1.30)	1.59 (1.56 - 1.62)
Age, region, smoking	2405	48	1.00 (ref.)	1.05 (1.03 - 1.08)	1.14 (1.11 - 1.16)	1.26 (1.23 - 1.28)	1.53 (1.50 - 1.56)
Age, region, alcohol	3830	17	1.00 (ref.)	1.07 (1.05 - 1.09)	1.17 (1.14 - 1.19)	1.32 (1.29 - 1.34)	1.68 (1.65 - 1.72)
Age, region, physical activity	3467	25	1.00 (ref.)	1.07 (1.05 - 1.09)	1.17 (1.14 - 1.19)	1.31 (1.29 - 1.34)	1.66 (1.63 - 1.69)
Age, region, body mass index	3067	34	1.00 (ref.)	1.07 (1.04 - 1.09)	1.15 (1.13 - 1.17)	1.29 (1.26 - 1.31)	1.62 (1.58 - 1.65)
Age, region, height	4482	3	1.00 (ref.)	1.08 (1.06 - 1.10)	1.19 (1.16 - 1.21)	1.35 (1.33 - 1.38)	1.76 (1.73 - 1.79)
Age, region, parity	4150	11	1.00 (ref.)	1.08 (1.05 - 1.10)	1.18 (1.16 - 1.20)	1.34 (1.31 - 1.36)	1.72 (1.68 - 1.75)
Age, region, menopausal hormone therapy	4549	2	1.00 (ref.)	1.08 (1.06 - 1.10)	1.19 (1.16 - 1.21)	1.35 (1.33 - 1.38)	1.76 (1.73 - 1.80)
Age, region, health behaviours[2]	929	80	1.00 (ref.)	1.03 (1.01 - 1.05)	1.08 (1.06 - 1.10)	1.15 (1.13 - 1.18)	1.33 (1.30 - 1.35)
Age, region, health behaviours[2], plus education	689	85	1.00 (ref.)	1.02 (1.00 - 1.04)	1.07 (1.05 - 1.09)	1.13 (1.11 - 1.15)	1.28 (1.25 - 1.31)
Fully adjusted[3]	637	86	1.00 (ref.)	1.02 (1.00 - 1.04)	1.07 (1.04 - 1.09)	1.13 (1.10 - 1.15)	1.27 (1.25 - 1.30)
INCIDENT COMPLEX CVM (n)			4,110	4,577	5,117	5,697	7,231
HR (95% CI), adjusted for,							
Age, region (minimally adjusted)	1583	-	1.00 (ref.)	1.08 (1.04 - 1.13)	1.24 (1.19 - 1.30)	1.43 (1.37 - 1.48)	1.97 (1.90 - 2.05)
Age, region, education	957	40	1.00 (ref.)	1.05 (1.01 - 1.10)	1.19 (1.14 - 1.24)	1.31 (1.26 - 1.37)	1.73 (1.66 - 1.80)
Age, region, smoking	808	49	1.00 (ref.)	1.05 (1.01 - 1.10)	1.18 (1.13 - 1.23)	1.30 (1.25 - 1.36)	1.66 (1.59 - 1.72)
Age, region, alcohol	1298	18	1.00 (ref.)	1.07 (1.02 - 1.12)	1.21 (1.16 - 1.27)	1.37 (1.32 - 1.43)	1.86 (1.79 - 1.93)
Age, region, physical activity	1187	25	1.00 (ref.)	1.07 (1.02 - 1.12)	1.21 (1.16 - 1.27)	1.37 (1.32 - 1.43)	1.84 (1.76 - 1.91)
Age, region, body mass index	994	37	1.00 (ref.)	1.06 (1.02 - 1.11)	1.18 (1.13 - 1.23)	1.33 (1.27 - 1.38)	1.75 (1.68 - 1.82)
Age, region, height	1514	4	1.00 (ref.)	1.08 (1.04 - 1.13)	1.24 (1.19 - 1.29)	1.42 (1.37 - 1.48)	1.96 (1.89 - 2.04)
Age, region, parity	1402	11	1.00 (ref.)	1.08 (1.04 - 1.13)	1.23 (1.18 - 1.29)	1.41 (1.35 - 1.46)	1.91 (1.83 - 1.98)
Age, region, menopausal hormone therapy	1532	3	1.00 (ref.)	1.08 (1.04 - 1.13)	1.23 (1.18 - 1.29)	1.41 (1.36 - 1.47)	1.96 (1.89 - 2.04)
Age, region, health behaviours[2]	284	82	1.00 (ref.)	1.01 (0.97 - 1.06)	1.10 (1.05 - 1.15)	1.16 (1.11 - 1.22)	1.37 (1.31 - 1.43)
Age, region, health behaviours[2], plus education	199	87	1.00 (ref.)	1.01 (0.96 - 1.05)	1.08 (1.04 - 1.13)	1.13 (1.08 - 1.18)	1.31 (1.25 - 1.37)
Fully adjusted[3]	183	88	1.00 (ref.)	1.00 (0.96 - 1.05)	1.08 (1.03 - 1.13)	1.12 (1.08 - 1.18)	1.30 (1.24 - 1.36)

[1]Likelihood-ratio test statistic; [2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, education, height, parity, and ever use of menopausal hormone therapy; CVM=cardiovascular multimorbidity; ref. means reference category

6.3.4.2 Cardiovascular multimorbidity pairs

The role of various covariates in the relationships between the two SES measures and the six selected CVM pairs (Appendix E, Tables E1-12, page 334-45) were broadly similar to what was observed for CVM and complex CVM (see Section 6.3.4.1, page 200): further adjusting for the other SES measure alone accounted for about 40% of the associations observed in the minimally adjusted models; smoking often had the largest individual effect; reproductive factors or height had minimal effect; and jointly adjusting for the four health behaviours (i.e. smoking, alcohol consumption, physical activity, and BMI) accounted for most of the associations ($\chi^2_{attenuation}$ =69-85%). Adjusting for the four health behaviours plus the other SES factor accounted for an even greater proportion of the associations ($\chi^2_{attenuation}$ =77-90%).

Exceptions were the three pairs including atrial fibrillation. For the associations between education and IHD-AF (Appendix E, Table E1, page 334), AF-HF (Table E7, page 340), and AF-STR (Table E11, page 344), adjusting only for height (in addition to minimal adjustment) led to an increase of the LR statistics by 8-19%, resulting in a slightly steeper educational gradient in risk. However, the associations between these CVM pairs and area deprivation were little changed by adjusting for height. Furthermore, for both SES measures and AF-STR, adjusting only for alcohol consumption (in addition to minimal adjustment) accounted for a relatively small proportion of the associations (Table E11-E12, page 344-45), resulting in a reduction of the LR statistic of 9% for education and 8% for deprivation (compared to the 15-30% reductions seen for the other pairs).

6.3.5 Sensitivity analyses

6.3.5.1 Assessing the role of individual cardiovascular diseases in potentially driving the observed associations with cardiovascular multimorbidity

Sensitivity analyses leaving out each of the six selected individual CVDs from the definition of CVM were conducted to evaluate whether some individual CVDs may play a major role in driving the associations between SES and CVM. Leaving out any one of the individual CVDs when defining CVM did not materially change CVM risks in relation to either SES measure (Appendix E, Figure E1-2, page 346-7). This may indicate that no individual CVD was driving the social gradient observed in CVM risk, although the social gradient was likely to be largely accounted for by other factors (as indicated by the level of attenuation in the LR statistics both in the main analysis and in these sensitivity analyses).

The associations between SES and the selected individual CVDs were also assessed to gain a more complete picture of the contribution of individual CVDs to the social gradient in CVM risk (Appendix E, Figure E3-4, page 348-9). Lower SES was associated with a higher risk of all individual CVDs, even after multivariable adjustment. The magnitude of the relative risks in relation to SES was greatest for heart failure and smallest for other arrhythmias.

6.3.5.2 Assessing the potential for mediation and residual confounding

The sensitivity analysis excluding women treated for hypertension or diabetes at recruitment (n=1,045,607) showed similar results to the main findings for CVM, for both education (Appendix E, Table E13, page 350) and deprivation (Table E14, page 351). It was therefore unlikely that a large proportion of the observed association between SES and CVM risk was mediated by diabetes or hypertension, after accounting for other covariates.

In the sensitivity analysis restricted to never-smokers (n=603,782), the associations of both education and area deprivation with CVM (Appendix E, Table E15-16, page 352-3) were

not materially different from those in the main analysis (which adjusted for smoking in the whole cohort). This indicated that substantial residual confounding by smoking was unlikely to be present in this study.

6.4 Discussion

6.4.1 Key findings and interpretation of results

6.4.1.1 Overall summary

In this large prospective cohort of middle-aged UK women, there was a social gradient in the risk of all CVM outcomes. After minimal adjustment, women with the lowest level (vs. highest) of educational attainment had almost twice the risk of CVM, and 2.3 times the risk of complex CVM. Women living in the most deprived areas (vs. least) had a 76% higher risk of CVM and two times the risk of complex CVM.

Much of this social gradient in risk appeared to be explained by the social patterning of modifiable cardiovascular risk factors. After full adjustment, women with the lowest level (vs. highest) of educational attainment had a 36% higher risk of CVM and 50% higher risk of complex CVM. Similarly, women living in the most deprived areas (vs. least) had a 27% higher risk of CVM and 30% higher risk of complex CVM. Over 80% of the observed associations in the minimally adjusted models was accounted for by four health behaviours measured at recruitment: smoking, alcohol consumption, physical activity, and BMI. At least 43% of the observed associations was accounted for by smoking behaviour alone. Adjusting for reproductive factors, in contrast, had little effect.

The findings were similar for CVM pairs. The lowest level (vs. highest) of both SES factors was associated with 1.6 to three times the risk of each selected CVM pair in the minimally adjusted models, and at least 69% of these associations could potentially be explained by simultaneous adjustment for the four health behaviours. Adjusting for the four health behaviours plus the other SES factor accounted for 77-90% of all associations between both SES measures and all CVM outcomes.

6.4.1.2 Socioeconomic status and risk of multimorbidity

There is currently a lack of reliable prospective evidence on the relationship between SES and multimorbidity. Most previous studies were small, cross-sectional, and did not investigate the effect of potential mediators/confounders. [10, 11] Moreover, it was difficult to compare findings across the studies or with the findings of this chapter due to the heterogeneous selection of diseases to define multimorbidity and the variability in SES measures. The proportion of the multimorbidity outcome constituted by CVDs was not reported in most studies, although the list of conditions to define multimorbidity almost always included common CVDs like ischaemic heart disease and stroke.

6.4.1.2.1 Education and risk of cardiovascular multimorbidity

As identified in the systematic review presented in Chapter 2, no prospective study has specifically investigated CVM. Only one study has investigated incident cardiometabolic multimorbidity in relation to education. [20] This study was conducted in a cohort of 13,714 middle-aged Australian women, and defined cardiometabolic multimorbidity as developing two or more of coronary heart disease, stroke, and diabetes. After adjusting for various sociodemographic and lifestyle behavioural factors as well as other chronic conditions, the odds ratio for cardiometabolic multimorbidity in those with no educational qualifications was 1.45 (95% CI 1.01-2.10) compared to women with university/higher degree qualifications. This was similar in direction and magnitude to the relative risks associated with the lowest level of educational attainment (compared to the highest) observed for CVM and individual ischaemic heart disease in this chapter. Nevertheless, the inclusion of diabetes – a cardiovascular risk factor rather than a CVD – meant that the findings of this study did not strictly address CVM as I have defined it in this thesis, and the Australian study only had 423 cases of cardiometabolic multimorbidity in their analysis.

Three other prospective studies explored the relationship between education and incident multimorbidity, although not CVM specifically. [21-23] One study [21] in 6,425 British civil servants with 1,694 cases of multimorbidity found that civil servants with low educational attainment (lower secondary school or less) had a 26% higher risk of multimorbidity compared to those with high educational attainment (university or higher degree) after adjusting for age, sex, race, marital status, and birth cohort. This was similar to the 23% higher risk of CVM in women who have attained technical qualifications (vs. tertiary qualifications) after adjusting only for age and region in this chapter. The other two studies [22, 23] adjusted for health-related behavioural factors (e.g. BMI, smoking, physical activity) and other potential confounders and mediators (e.g. blood pressure, blood cholesterol, social detachment). They did not find a statistically significant relationship between educational attainment and multimorbidity risk. This was consistent with the fully adjusted results reported in this chapter and the finding that that health-related behavioural factors accounted for much of the educational gradient in CVM risk. However, both of these studies had fewer than 500 cases of incident multimorbidity and lacked precision to estimate these moderate relative risks.

6.4.1.2.2 Area deprivation and risk of cardiovascular multimorbidity

No previous prospective study has investigated incident CVM in relation to area deprivation. Findings from other types of observational studies (e.g. cross-sectional or retrospective studies) on multimorbidity (rather than CVM) were qualitatively consistent with the findings of this chapter, according to a systematic review conducted in 2020. [11] The review identified 17 studies published since 2010 which examined the association between area deprivation and multimorbidity. The number of selected diseases to define multimorbidity across the 17 studies ranged from five to 146, and most studies were cross-sectional (n=13). Pooling data from seven cross-sectional studies which contained

available and relevant data, the review reported that the prevalence of multimorbidity was higher in participants living in the most deprived areas versus the least (pooled odds ratio 1.42, 95% CI 1.41-1.42). However, it was not appropriate to quantitatively compare this pooled effect estimate (which was based on multimorbidity prevalence) to the relative risks calculated based on incident multimorbidity in this chapter.

6.4.1.3 The role of covariates in accounting for social gradients in the risk of individual cardiovascular diseases

No previous study has formally quantified the contribution of covariates to the socioeconomic gradient in CVM risk. Existing evidence on individual CVDs have reported that health behaviours contributed to the socioeconomic gradients in CVD risk to varying degrees. A 2018 systematic review investigated the contribution of smoking, alcohol consumption, physical activity, and dietary patterns to the socioeconomic inequalities in cardiometabolic outcomes (defined as cardiovascular disease, hypertension, coronary heart disease, stroke, diabetes, impaired glucose tolerance, metabolic syndrome, allostatic load, or obesity). [8] The median contribution of multiple health behaviours to educational inequalities in risk was 18% (minimum -59%, maximum 56%) across 21 studies, while for inequalities associated with other SES indicators (including area deprivation) the median contribution of health behaviours was 30% (-16%, 69%) across 15 studies.

The review calculated the contribution of health behaviours using the absolute scale difference method [24], which tends to yield lower estimates of the proportional contribution of covariates compared to the LR attenuation method used in this chapter. As explained in Section 5.4.4 of Chapter 5 (page 174), the purpose of the LR attenuation method is not to quantify the exact contribution of confounders/mediators but to assess whether residual confounding is likely to account for the remaining associations. Therefore, estimates derived using these two different methods are not directly

comparable. Applying the absolute scale difference method to the findings of this chapter, it was estimated that behavioural factors (i.e. smoking, alcohol consumption, physical activity and BMI) contributed to 50% of the association between both SES factors and incident CVM. The LR attenuation was 80-85% for those relationships. In contrast to the example given in Chapter 5 (page 175), here the interpretation differed by the method used: while the absolute scale difference method indicated that substantial proportions of the associations were yet unexplained, the LR attenuation method suggested that residual confounding could account for any remaining associations.

Heterogeneity in demographic characteristics of study populations, study type (e.g. cross-sectional or longitudinal), and the measures of SES, health behaviours, and cardiovascular outcomes are some possible reasons for the differences in the findings between the papers in the review and the results of this chapter. The review also found that on average, the contribution of health behaviours in socioeconomic gradients in CVD risk was higher in longitudinal studies compared to cross-sectional studies. This could be because the measures of health behaviours more accurately reflect true behaviour over time in longitudinal studies.

6.4.1.3.1 Longitudinal evidence on the role of covariates in the educational gradient in risk

Previous work in the Million Women Study on socioeconomic factors and coronary heart disease found that 76% of the association between education and coronary heart disease risk was accounted for by the four health behaviours explored in this chapter (i.e. smoking, alcohol consumption, physical inactivity, and BMI), using the same LR statistic attenuation method. [9] This was similar to the 80% contribution of these four health behaviours to the educational gradient in CVM risk in this chapter.

Other prospective studies on education have also estimated that over half of the educational gradient in risk of coronary heart disease and stroke could potentially be explained by modifiable cardiovascular risk factors. [25-27] For example, a collaborative analysis of multiple cohorts estimated that traditional cardiovascular risk factors (i.e. smoking, alcohol consumption, BMI, cholesterol levels, systolic blood pressure, and diabetes) accounted for 58% of the social gradient in stroke incidence in middle-aged UK women. [25] Furthermore, a prospective cohort study of 40,011 Dutch men and women found that diet and lifestyle (mainly smoking and alcohol) accounted for over 70% of the educational gradient in both coronary heart disease and stroke incidence. [27] These results were consistent with the findings of this chapter in that a substantial proportion of the educational inequalities in CVM risk was accounted for by the social patterning of health-related behaviours.

A recent MR study of 217,013 adults in the UK Biobank estimated that lifetime smoking exposure, systolic blood pressure, and BMI each mediated up to 34% (95% CI 17% to 50%), 21% (15% to 26%), and 18% (14% to 23%) of the protective effect of higher education on coronary heart disease, together accounting for approximately 36% (5% to 68%) of the association. [28] The estimated contribution of smoking and BMI to the educational gradient in CVD risk were smaller than those for CVM in this chapter, but the confidence intervals were wide. The authors speculated that differences in health seeking behaviour and prescription of medications such as statins may explain more of the observed gradient in risk, although such information was not available in their analyses. Sensitivity analyses assessing the effect of diet and exercise did not further account for the educational gradient in risk. Similar results were found for stroke and composite CVD (i.e. coronary heart disease, stroke, or myocardial infarction) risk.

6.4.1.3.2 *Longitudinal evidence on the role of covariates in the gradient in risk by area deprivation*

Previous work in the Million Women Study also found that simultaneously adjusting for the four health behaviours accounted for 71% of the association with area deprivation and incident coronary heart disease, similar to the estimates of ~80% for area deprivation CVM in this chapter. [9]

Only one longitudinal study [29] examined the effect of adjusting for lifestyle behaviours in the relationship between area deprivation and multimorbidity. This study recruited 4,510 Scottish men and women in 1987 and re-surveyed them in four follow-up waves over 20 years. At recruitment and each follow-up wave, information on each participant's area deprivation (using Carstairs scores, calculated based on car ownership, male unemployment, overcrowding, and low social class), health-related behaviours, and self-reported chronic conditions were collected. Multimorbidity was defined as the presence of two or more of 40 chronic conditions, using the same selection of diseases as Barnett et al. [30]. Multimorbidity at each wave was modelled based on information on area deprivation and covariates collected from the previous wave.

The study found that participants living in the most deprived areas had a 46% higher risk of developing multimorbidity compared to those living in the least deprived areas (OR 1.46, 95% CI 1.26-1.68), after adjusting for various age-related covariates, prior multimorbidity, time between waves, sex, and sex*cohort interaction. Adjusting for smoking alone attenuated the relationship by 15.4%, followed by diet, which attenuated the relationship by 11.9%. Adjusting for smoking, BMI, diet, physical activity, and alcohol consumption altogether resulted in a 40.8% attenuation. These estimates were smaller than those found in this chapter. This may partly be due to the inclusion of people with prior

disease or multimorbidity in the analyses and of many non-CVDs in the multimorbidity definition.

In summary, although previous studies consistently demonstrated that substantial proportions of the social gradient in CVD/multimorbidity risk were likely to be mediated by known cardiovascular risk factors, the magnitude of their contributions was variable across studies. This was possibly due to reasons such as differences in the choice and measurement of the potential mediators, statistical methods to quantify mediation, and distribution of health-related behavioural factors across SES levels in each study population.

6.4.1.4 Potential mechanisms behind the social gradient in cardiovascular multimorbidity

Far from being randomly distributed in the population, health-damaging behaviours tend to cluster in socioeconomically disadvantaged groups. [31] This was true of the study participants in this chapter: women of lower SES were more likely to smoke, have higher BMI, and be physically inactive. The social gradient in CVM risk appeared to be largely mediated by the differences in behavioural profile by socioeconomic position.

Many theories exist on how higher education can lead to healthier behaviour. [32] Certain cognitive skills learned through education (e.g. ability to think abstractly and assess long-term benefits and harms) may help people make healthier choices. For adolescents, spending time in education can limit time spent engaging in unhealthy behaviour such as smoking and alcohol use, which can have lifelong consequences on exposure. Education is often the means to secure occupations, higher income, and social connections which can be used to achieve better health.

Both physical and social mechanisms underlie the association between area deprivation and health-related behaviours. [33] Physical characteristics of the area such as access to public transport, recreational facilities, healthy food options, and green spaces can influence lifestyle choices such as those regarding physical exercise and diet. Social factors like the safety and social cohesion of the area can also influence health-related behaviours. More deprived areas have a higher prevalence of various cardiovascular risk factors including obesity [6], hypertension [7], and depression [34]. Socioeconomic disadvantage has also been associated with poor medication adherence for CVD treatment [35] and previous studies have reported potential social inequalities in the quality and access to cardiovascular care in the UK [36].

In this chapter, more than 80% of the social gradient in CVM risk was accounted for by four health behaviours: smoking, BMI, physical activity, and alcohol consumption, in order of relative importance. Since these behaviours were imperfectly measured, some or all of the remaining association may also be accounted for by these factors. [37]

6.4.2 Strengths and limitations

The strengths of this study include the uniquely large sample size, nearly complete follow-up, and the ability to adjust for a wide range of potential confounders and mediators. It was possible to reliably assess the social gradient in multiple CVM outcomes due to the high number of cases. Another strength is that this study investigated the impact of selected reproductive factors and height, which have not been studied in the context of social gradients previously. Overall, this study contributes substantial new prospective evidence on social gradient in CVM incidence, where there is currently a significant research gap.

There were several limitations to the study presented in this chapter. Education and area deprivation are not comprehensive measures of SES, but information on other important

measures of SES such as occupation or household income and other relevant characteristics of socioeconomic well-being (e.g. social isolation) were not available in the recruitment questionnaire. It was not possible to control for the potential impact of reverse causation; for example, ill health during childhood can have a negative effect on educational attainment and ill health in adulthood can adversely affect job prospects and social circumstances, causing people to live in more deprived areas. [32] The recruitment questionnaire also did not have information on some potential confounders which may play a role in CVM development such as dietary patterns, mental health, or quality of care.

6.4.3 Conclusions

This study formally quantified the associations between two common SES measures with CVM risk and assessed the extent of the associations which may be explained by known cardiovascular risk factors. There was a social gradient for all CVM outcomes, which was largely attenuated by adjusting for smoking, alcohol consumption, physical activity, and BMI. Height and reproductive factors accounted for little or none of the social gradient in CVM outcomes. The associations which remained after multivariable adjustment may potentially be explained by residual confounding or mediation, since the covariates have been measured with error and there may still be unmeasured mediators/confounders of small effect. Future studies could explore the impact of using directly measured health behaviours as well as taking multiple measurements over time to minimise potential residual confounding, to gain a more accurate estimate of the contribution of various factors to the social gradient in CVM risk.

6.5 References

[1] Pujades-Rodriguez M, Timmis A, Stogiannis D, Rapsomaniki E, Denaxas S, Shah A, et al. Socioeconomic deprivation and the incidence of 12 cardiovascular diseases in 1.9 million women and men: implications for risk prediction and prevention. PloS one. 2014 Aug;9:e104671. Available from: *https://doi.org/10.1371/journal.pone.0104671*.

[2] Kivimäki M, Batty GD, Pentti J, Shipley MJ, Sipilä PN, Nyberg ST, et al. Association between socioeconomic status and the development of mental and physical health conditions in adulthood: a multi-cohort study. The Lancet Public health. 2020 Mar;5:e140–e149. Available from: *https://doi.org/10.1016/S2468-2667(19)30248-8*.

[3] Kerr GD, Slavin H, Clark D, Coupar F, Langhorne P, Stott DJ. Do vascular risk factors explain the association between socioeconomic status and stroke incidence: a meta-analysis. Cerebrovascular diseases (Basel, Switzerland). 2011 Oct;31:57–63. Available from: *https://doi.org/10.1159/000320855*.

[4] Manrique-Garcia E, Sidorchuk A, Hallqvist J, Moradi T. Socioeconomic position and incidence of acute myocardial infarction: a meta-analysis. J Epidemiol Community Health. 2011 Apr;65(4):301. Available from: *https://doi.org/10.1136/jech.2009.104075*.

[5] Hiscock R, Bauld L, Amos A, Fidler JA, Munafò M. Socioeconomic status and smoking: a review. Annals of the New York Academy of Sciences. 2012 Feb;1248:107–23. Available from: *https://doi.org/10.1111/j.1749-6632.2011.06202.x*.

[6] Mohammed SH, Habtewold TD, Birhanu MM, Sissay TA, Tegegne BS, Abuzerr S, et al. Neighbourhood socioeconomic status and overweight/obesity: a systematic review and meta-analysis of epidemiological studies. BMJ Open. 2019 Nov;9(11):e028238. Available from: *https://doi.org/10.1136/bmjopen-2018-028238*.

[7] Leng B, Jin Y, Li G, Chen L, Jin N. Socioeconomic status and hypertension: a meta-analysis. Journal of Hypertension. 2015 Feb;33(2):221–229. Available from: *https://-doi.org/10.1097/hjh.0000000000000428*.

[8] Petrovic D, de Mestral C, Bochud M, Bartley M, Kivimäki M, Vineis P, et al. The contribution of health behaviors to socioeconomic inequalities in health: A systematic review. Preventive Medicine. 2018 Aug;113:15–31. Available from: *https://doi.org/-10.1016/j.ypmed.2018.05.003*.

[9] Floud S, Balkwill A, Moser K, Reeves GK, Green J, Beral V, et al. The role of health-related behavioural factors in accounting for inequalities in coronary heart disease risk by education and area deprivation: prospective study of 1.2 million UK women. BMC Med. 2016 Oct;14(1):145. Available from: *https://doi.org/10.1186/s12916-016-0687-2*.

[10] Pathirana TI, Jackson CA. Socioeconomic status and multimorbidity: a systematic review and meta-analysis. Australian and New Zealand journal of public health. 2018 Apr;42:186–194. Available from: *https://doi.org/10.1111/1753-6405.12762*.

[11] Ingram E, Ledden S, Beardon S, Gomes M, Hogarth S, McDonald H, et al. Household and area-level social determinants of multimorbidity: a systematic review. J Epidemiol Community Health. 2021 Mar;75(3):232. Available from: *https://doi.org/-10.1136/jech-2020-214691*.

[12] Li W, Ruan W, Lu Z, Wang D. Parity and risk of maternal cardiovascular disease: A dose-response meta-analysis of cohort studies. European journal of preventive cardiology. 2019 Apr;26:592–602. Available from: *https://doi.org/10.1177/-2047487318818265*.

[13] Kim JE, Chang JH, Jeong MJ, Choi J, Park J, Baek C, et al. A systematic review and meta-analysis of effects of menopausal hormone therapy on cardiovascular diseases. Scientific Reports. 2020 Nov;10(1):20631. Available from: *https://doi.org/10.1038/-s41598-020-77534-9*.

[14] Townsend P, Phillimore P, Beatti A. Health Deprivation: Inequality and the North. London: Croom Helm; 1988.

[15] Adams J, Ryan V, White M. How accurate are Townsend Deprivation Scores as predictors of self-reported health? A comparison with individual level data. J Public Health (Oxf). 2005 Mar;27(1):101–106. Available from: *https://doi.org/10.1093/pubmed/fdh193*.

[16] Shavers VL. Measurement of socioeconomic status in health disparities research. J Natl Med Assoc. 2007 Sep;99(17913111):1013–1023. Available from: *https://-pubmed.ncbi.nlm.nih.gov/17913111/*.

[17] Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, et al. Comparison of various characteristics of women who do and do not attend for breast cancer screening. Breast Cancer Res. 2001 Dec;4:R1. Available from: *https://doi.org/10.1186/bcr418*.

[18] Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Series B Stat Methodol. 1995;57(1):289–300. Available from: *https://doi.org/10.1111/j.2517-6161.1995.tb02031.x*.

[19] Parish S, Peto R, Palmer A, Clarke R, Lewington S, Offer A, et al. The joint effects of apolipoprotein B, apolipoprotein A1, LDL cholesterol, and HDL cholesterol on risk: 3510 cases of acute myocardial infarction and 9805 controls. Eur Heart J. 2009 Sep;30(19520708):2137–2146. Available from: *https://doi.org/10.1093/eurheartj/ehp221*.

[20] Xu X, Mishra GD, Dobson AJ, Jones M. Progression of diabetes, heart disease, and stroke multimorbidity in middle-aged women: A 20-year cohort study. PLOS Medicine. 2018 Mar;15(3):e1002516. Available from: *https://doi.org/10.1371/-journal.pmed.1002516*.

[21] Dugravot A, Fayosse A, Dumurgier J, Bouillon K, Rayana TB, Schnitzler A, et al. Social inequalities in multimorbidity, frailty, disability, and transitions to mortality: a 24-year follow-up of the Whitehall II cohort study. The Lancet Public Health. 2020 Jan;5(1):e42–e50. Available from: *https://doi.org/10.1016/S2468-2667(19)30226-9*.

[22] Mounce LTA, Campbell JL, Henley WE, Tejerina Arreal MC, Porter I, Valderas JM. Predicting Incident Multimorbidity. Annals of family medicine. 2018 Jul;16:322–329. Available from: *https://doi.org/10.1370/afm.2271*.

[23] Wikström K, Lindström J, Harald K, Peltonen M, Laatikainen T. Clinical and lifestyle-related risk factors for incident multimorbidity: 10-year follow-up of Finnish population-based cohorts 1982-2012. European journal of internal medicine. 2015 Apr;26:211–6. Available from: *https://doi.org/10.1016/j.ejim.2015.02.012*.

[24] Stringhini S, Sabia S, Shipley M, Brunner E, Nabi H, Kivimäki M, et al. Association of Socioeconomic Position With Health Behaviors and Mortality. JAMA. 2010 Mar;303(12):1159–1166. Available from: *https://doi.org/10.1001/jama.2010.297*.

[25] Ferrario MM, Veronesi G, Kee F, Chambless LE, Kuulasmaa K, Jørgensen T, et al. Determinants of social inequalities in stroke incidence across Europe: a collaborative analysis of 126,635 individuals from 48 cohort studies. J Epidemiol Community Health. 2017 Dec;71(12):1210. Available from: *https://doi.org/10.1136/jech-2017-209728*.

[26] Kershaw KN, Droomers M, Robinson WR, Carnethon MR, Daviglus ML, Monique Verschuren WM. Quantifying the contributions of behavioral and biological risk factors to socioeconomic disparities in coronary heart disease incidence: the MORGEN study. European Journal of Epidemiology. 2013 Oct;28(10):807–814. Available from: *https://doi.org/10.1007/s10654-013-9847-2*.

[27] Méjean C, Droomers M, van der Schouw YT, Sluijs I, Czernichow S, Grobbee DE, et al. The contribution of diet and lifestyle to socioeconomic inequalities in cardiovascular morbidity and mortality. International Journal of Cardiology. 2013 Jul;168(6):5190–5195. Available from: *https://doi.org/10.1016/j.ijcard.2013.07.188*.

[28] Carter AR, Gill D, Davies NM, Taylor AE, Tillmann T, Vaucher J, et al. Understanding the consequences of education inequality on cardiovascular disease: mendelian randomisation study. BMJ. 2019 May;365:l1855. Available from: *https://-doi.org/10.1136/bmj.l1855*.

[29] Katikireddi SV, Skivington K, Leyland AH, Hunt K, Mercer SW. The contribution of risk factors to socioeconomic inequalities in multimorbidity across the lifecourse: a longitudinal analysis of the Twenty-07 cohort. BMC Medicine. 2017 Aug;15(1):152. Available from: *https://doi.org/10.1186/s12916-017-0913-6*.

[30] Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet. 2012 Jul;380(9836):37–43. Available from: *https://doi.org/-10.1016/S0140-6736(12)60240-2*.

[31] Shankar A, McMunn A, Steptoe A. Health-Related Behaviors in Older Adults: Relationships with Socioeconomic Status. American Journal of Preventive Medicine. 2010 Jan;38(1):39–46. Available from: *https://doi.org/10.1016/j.amepre.2009.08.026*.

[32] Berkman LF, Kawachi I, Glymour MM, editors. Social Epidemiology. 2nd ed. Oxford University Press; 2014. 17-62 p.

[33] Schultz WM, Kelli HM, Lisko JC, Varghese T, Shen J, Sandesara P, et al. Socioeconomic Status and Cardiovascular Outcomes. Circulation. 2018 May;137(20):2166–2178. Available from: *https://doi.org/10.1161/-CIRCULATIONAHA.117.029652*.

[34] Remes O, Lafortune L, Wainwright N, Surtees P, Khaw KT, Brayne C. Association between area deprivation and major depressive disorder in British men and women: a cohort study. BMJ Open. 2019 Nov;9(11):e027530. Available from: *https://doi.org/-10.1136/bmjopen-2018-027530*.

[35] Gast A, Mathes T. Medication adherence influencing factors–an (updated) overview of systematic reviews. Systematic Reviews. 2019 May;8(1):112. Available from: *https://doi.org/10.1186/s13643-019-1014-8*.

[36] Asthana S, Moon G, Gibson A, Bailey T, Hewson P, Dibben C. Inequity in cardiovascular care in the English National Health Service (NHS): a scoping review of the literature. Health Soc Care Community. 2018 May;26(3):259–272. Available from: *https://doi.org/10.1111/hsc.12384*.

[37] Fewell Z, Davey Smith G, Sterne JAC. The Impact of Residual and Unmeasured Confounding in Epidemiologic Studies: A Simulation Study. Am J Epidemiol. 2007 Sep;166(6):646–655. Available from: *https://doi.org/10.1093/aje/kwm165*.

7 Reproductive factors and risk of incident cardiovascular multimorbidity

7.1 Introduction

7.1.1 Background

A growing body of literature suggests that women's cardiovascular disease (CVD) risk may be influenced by their reproductive history, such as age at menarche, number of children, and breastfeeding, among others. [1] Menarche signals the initiation of changes in ovarian and endocrine functions that enable reproduction. [2] There is evidence that age at menarche is associated with cardiometabolic health in adolescents and young adults [3, 4], and potential implications for CVD risk in middle-age have been reported in the Million Women Study [5]. Pregnancy accompanies metabolic changes which are unfavourable for cardiovascular health, including weight gain and increasing levels of cholesterol. [6, 7] It has been suggested that breastfeeding may counteract some of these adverse effects. [8] However, findings of previous studies on the associations between these reproductive factors and individual CVDs have been inconsistent, and often had major limitations like small sample size or lack of information on important mediators and/or confounders such as body mass index (BMI). [1, 9-11] Until now, no study has investigated the relationship between reproductive factors and incident cardiovascular multimorbidity (CVM).

7.1.2 Research questions

Using data from the Million Women Study (described in Chapter 3), this chapter addresses the following research questions:

- What are the prospective associations of reproductive factors with CVM and complex CVM?
- What are the prospective associations of reproductive factors with selected CVM pairs that comprise CVM?

- How much of the associations between reproductive factors and incident CVM and complex CVM are accounted for by known cardiovascular risk factors?
- Which, if any, individual CVDs potentially drive any observed associations between reproductive factors and CVM?

7.2 Methods

This section describes the methods that are specific to the analyses presented in this chapter. Methods that are common to multiple chapters are described only briefly; further details on such methods are available in Section 3.3 of Chapter 3 (page 79).

7.2.1 Outcomes

The primary outcomes investigated were CVM, complex CVM, and selected CVM pairs. CVM and complex CVM were defined as an individual developing ≥2 or ≥4 of the selected 19 CVDs (see Table 3.3 in Chapter 3, page 81), respectively. The CVM pairs were, 'Ischaemic heart disease & Atrial fibrillation' (IHD-AF), 'Ischaemic heart disease & Heart failure' (IHD-HF), 'Ischaemic heart disease & Other Arrhythmias' (IHD-ARR), 'Atrial fibrillation & Heart failure' (AF-HF), 'Other arrhythmias & Heart failure' (ARR-HF), and 'Atrial fibrillation & Stroke' (AF-STR). Section 3.3.1 of Chapter 3 (page 79) describes and defines these outcomes in more detail.

7.2.2 Risk factors

The reproductive factors investigated in this chapter were: age at menarche (<12, 12, 13, 14, 15+ years), parity (nulliparous, 1-2, 3+), and total breastfeeding duration (0, 1, 2-5, 6-11, 12+ months). Total breastfeeding duration was analysed in parous women only. All risk factor data were collected through the self-reported recruitment questionnaire (Appendix B, page 298). Participants were asked, "About how old were you when your periods started?" and reported their age at menarche in years. For information on parity, participants were asked, "How many children have you had? (please include stillbirths; it is not necessary to include miscarriages)." Other studies in the scientific literature may refer to parity in terms of number of births or pregnancies, but in this chapter, parity will be referred to as the number of children of each participant, consistent with the primary

phrasing of the question. Participants were also asked, "When was each child born, and for how many months did you breastfeed each child, if at all?" and reported the day, month, and year of each child's birth and the number of months they breastfed each child. They were asked to answer "0 months" if they did not breastfeed the child and "1 month" if they breastfed for one month or less.

7.2.3 Statistical analyses

7.2.3.1 Cox proportional hazards regression

Cox proportional hazards regression with attained age as the underlying time variable was used to calculate hazard ratios (HRs) and 95% confidence intervals (CIs) for the CVM outcomes, in relation to the selected reproductive factors. Section 3.3.4.1 of Chapter 3 (page 85) describes this method in detail, including the dates of follow-up and censoring.

Minimally adjusted regression models were stratified by age at recruitment (<53, 53-55, 56-58, 59-61, 62-64, 65+ years) and recruitment region (ten geographic regions: Oxford, East Anglia, South West, Thames, West Midlands, North York, Trent, North West [Mersey], North West [Manchester/Lancaster], Scotland). The minimally adjusted models for breastfeeding duration were additionally adjusted for parity. Women with more children have a longer time in which they may potentially breastfeed their children, so all else being equal, they are expected to have breastfed their children for a longer duration on average. Therefore, it was judged important to control for parity in the minimally adjusted models to separate out the effects of breastfeeding from other potential parity-related effects on the risk of CVM outcomes.

For multivariable adjusted HRs, models were additionally adjusted for socioeconomic deprivation (fifths, based on the Townsend index of area deprivation), educational attainment (none, technical qualifications, secondary school, tertiary [college or

university]), BMI (<20, 20-24, 25-29, 30-34, 35+ kg/m^2), height (<155, 155-159, 160-164, 165-169, 170+ cm), cigarette smoking (never, past, current <15 per day, current 15+ per day), alcohol consumption (rarely/never, 1-6, 7-14, 15+ units per week), strenuous physical activity (none, up to once per week, 2-3 times per week, 4+ times per week), parity (nulliparous, 1-2, 3+), and ever use of menopausal hormone therapy (yes or no). For each adjustment variable, missing values (<6% for each variable) were assigned to a separate category.

To examine the combined effect of parity and breastfeeding on the risk of CVM, HRs for CVM and complex CVM were also compared across three groups: nulliparous women (reference group), parous women who ever breastfed, and parous women who never breastfed. Since nulliparous women will have never breastfed children, the HR in parous women who never breastfed (vs. nulliparous women) represents the association between parity and CVM/complex CVM risk without the effect of breastfeeding. On the other hand, the HR in parous women who ever breastfed (vs. nulliparous women) represents the combined effect of parity and breastfeeding on CVM/complex CVM risk.

Differences in HRs across the categories of each risk factor were tested using the likelihood ratio test for heterogeneity, with a significance level of $\alpha = 0.05$. The Benjamini-Hochberg procedure (see Section 3.3.4.1 of Chapter 3, page 85) [12] was applied to control the false discovery rate to account for multiple testing in this chapter. Log-log plots were assessed to check the proportional hazards assumption, which was found to be satisfied in all models.

7.2.3.2 Likelihood-ratio χ^2 test statistic attenuation

In this chapter, the likelihood-ratio χ^2 test (LR) statistic of the minimally adjusted model was compared to various models further adjusted for covariates, individually and in

combinations. Attenuation in LR statistics was calculated to quantify the extent to which relationships between reproductive factors and CVM outcomes were accounted for by other variables. Section 3.3.4.2 of Chapter 3 (page 86) describes the LR statistic method in more detail.

7.2.3.3 Sensitivity analyses

To assess whether some individual CVDs may be driving any observed associations with CVM, prospective associations of reproductive factors with the six selected individual CVDs and with CVM minus each of the selected CVDs were examined. The six individual CVDs were ischaemic heart disease, atrial fibrillation, heart failure, other arrhythmias, stroke, and venous thromboembolism. The rationale for the selection of these CVDs and the definition of CVM minus each CVD are described in Section 5.2.3.2 of Chapter 5 (page 145).

7.3 Results

7.3.1 Baseline characteristics by reproductive factors

7.3.1.1 Age at menarche

The baseline characteristics of study participants by age at menarche are presented in Table 7.1. The mean self-reported age at menarche in this study cohort was 13 years (SD 1.6). There were 245,606 (20%) women who reported early menarche (i.e. age at menarche of <12 years) and 204,980 (17%) women who reported late menarche (i.e. age at menarche of 15 years or above). Women with early menarche were more likely to have a higher BMI and be younger at recruitment than women with late menarche. Women with early menarche were also less likely to smoke and more likely to live in affluent areas, have any educational qualifications, and report treatment for hypertension and diabetes at recruitment.

Table 7.1 Baseline characteristics of study participants by age at menarche

	Age at menarche, years				
	<12 (mean 10.8)	**12**	**13**	**14**	**15+ (mean 15.5)**
Women, n	245,606	216,686	297,722	249,999	204,980
Age at recruitment, mean (SD), years	55.9 (4.6)	56.3 (4.7)	56.4 (4.8)	56.8 (4.9)	57.2 (4.9)
Lifestyle and anthropometric factors					
Current smoker, % (n)	21.6 (50,239)	18.4 (37,722)	18.2 (51,313)	20.9 (49,079)	23.4 (44,865)
Drinker of alcoholic beverages, % (n)	63.9 (155,990)	66.5 (143,160)	67.1 (198,537)	65.1 (161,559)	62.3 (126,507)
No strenuous physical activity, % (n)	60.7 (143,878)	59.4 (124,652)	59.2 (170,612)	61.0 (146,765)	62.1 (121,750)
Body mass index, mean (SD), kg/m^2	27.3 (5.1)	26.4 (4.6)	25.9 (4.4)	25.6 (4.3)	25.4 (4.3)
Height, mean (SD), cm	161.2 (6.5)	161.9 (6.5)	162.4 (6.5)	162.4 (6.6)	162.5 (6.8)
Socio-economic factors					
Most deprived fifth, % (n)	19.5 (47,457)	16.9 (36,440)	16.9 (50,066)	19.9 (49,392)	22.9 (46,558)
No educational qualifications, % (n)	42.3 (101,391)	37.7 (79,841)	38.7 (112,706)	46.4 (112,918)	51.3 (101,663)
Reproductive factors					
Age at first birth, mean (SD), years	23.6 (4.3)	24.0 (4.4)	24.1 (4.3)	23.9 (4.3)	23.9 (4.2)
Age at natural menopause, mean (SD), years	48.9 (4.4)	49.2 (4.2)	49.3 (4.1)	49.2 (4.2)	49.1 (4.5)
Pre/perimenopausal, % (n)	13.5 (32,983)	14.0 (30,255)	14.0 (41,365)	12.5 (31,085)	11.0 (22,383)
Parity, mean (SD)	2.1 (1.2)	2.1 (1.2)	2.1 (1.2)	2.2 (1.2)	2.3 (1.3)
Ever breastfed children in parous women, % (n)	66.8 (116,015)	68.5 (105,318)	69.2 (147,183)	67.8 (120,144)	67.2 (95,661)
Ever use of contraceptive pills, % (n)	61.1 (148,584)	59.7 (128,243)	60.4 (178,108)	59.1 (146,207)	58.9 (119,157)
Ever use of hormone replacement therapy, % (n)	52.3 (126,982)	49.1 (105,253)	49.3 (145,521)	50.0 (123,480)	51.0 (103,050)
Healthcare factors					
Self-reported treatment for hypertension, % (n)	16.2 (39,882)	14.9 (32,268)	14.0 (41,742)	13.8 (34,430)	13.5 (27,695)
Self-reported treatment for diabetes mellitus, % (n)	2.8 (6,796)	2.0 (4,393)	1.8 (5,259)	1.8 (4,553)	1.9 (3,923)
Self-reported treatment for high cholesterol, % (n)	2.9 (7,039)	2.5 (5,376)	2.4 (7,238)	2.6 (6,419)	2.7 (5,561)

7.3.1.2 Parity

Table 7.2 displays the baseline characteristics of study participants by parity. Parity refers to the number of children that each woman has given birth to, including stillbirths. The majority (89%) of women were parous. Parous women were similar in age at recruitment to nulliparous women, but were more likely to be current smokers, have no educational qualifications, and have ever used contraceptive pills and hormone replacement therapy. Among parous women, those with the highest parity (i.e. four or more children) were older and had a higher BMI at recruitment than women with 1-3 children, and were more likely to smoke, less likely to drink and engage in strenuous physical activity, and more likely to live in socioeconomically deprived areas. Women with the highest parity were also more likely to have ever breastfed children and to report being treated for hypertension, diabetes, and high blood cholesterol at recruitment compared to women with a parity of 1-3.

Table 7.2 Baseline characteristics of study participants by parity

	Parity				
	Nulliparous	**1**	**2**	**3**	**4+ (mean 4.5)**
Women, n	134,387	155,467	534,449	271,241	140,181
Age at recruitment, mean (SD), years	56.6 (5.0)	56.4 (4.9)	56.1 (4.7)	56.7 (4.7)	57.7 (4.8)
Lifestyle and anthropometric factors					
Current smoker, % (n)	17.3 (22,031)	21.7 (31,833)	18.3 (92,197)	21.6 (55,127)	27.7 (36,185)
Drinker of alcoholic beverages, % (n)	66.3 (88,531)	64.2 (99,095)	68.0 (361,329)	64.4 (173,257)	54.8 (75,789)
No strenuous physical activity, % (n)	59.7 (77,743)	62.0 (92,943)	58.7 (303,815)	60.7 (158,324)	65.6 (87,065)
Body mass index, mean (SD), kg/m^2	25.7 (4.8)	25.9 (4.6)	25.9 (4.4)	26.4 (4.6)	27.2 (5.0)
Height, mean (SD), cm	162.7 (6.9)	162.1 (6.7)	162.2 (6.5)	162.0 (6.5)	161.3 (6.6)
Socio-economic factors					
Most deprived fifth, % (n)	17.5 (23,376)	18.4 (28,337)	15.2 (80,721)	20.7 (55,627)	33.6 (46,692)
No educational qualifications, % (n)	33.3 (43,754)	43.5 (65,894)	39.9 (208,168)	45.9 (121,091)	58.5 (78,725)
Reproductive factors					
Age at menarche, mean (SD), years	12.9 (1.6)	13.0 (1.6)	13.0 (1.6)	13.0 (1.6)	13.1 (1.7)
Age at first birth, mean (SD), years	-	26.8 (5.5)	24.4 (3.9)	22.7 (3.4)	21.1 (3.0)
Age at natural menopause, mean (SD), years	48.7 (4.5)	48.8 (4.3)	49.2 (4.1)	49.3 (4.2)	49.2 (4.5)
Pre/perimenopausal, % (n)	13.0 (17,362)	13.0 (20,096)	14.6 (77,403)	12.0 (32,466)	9.4 (13,019)
Ever breastfed children, % (n)	-	52.8 (66,875)	66.9 (290,441)	74.6 (157,821)	78.2 (79,420)
Ever use of contraceptive pills, % (n)	45.6 (60,598)	58.5 (90,028)	64.1 (339,875)	61.5 (165,033)	55.6 (76,648)
Ever use of hormone replacement therapy, % (n)	46.0 (61,259)	50.1 (77,048)	51.5 (272,706)	51.5 (137,923)	48.7 (67,095)
Healthcare factors					
Self-reported treatment for hypertension, % (n)	14.0 (18,857)	15.1 (23,515)	13.7 (73,105)	14.6 (39,560)	17.2 (24,110)
Self-reported treatment for diabetes mellitus, % (n)	2.2 (2,958)	2.1 (3,238)	1.7 (8,815)	2.1 (5,721)	3.4 (4,743)
Self-reported treatment for high cholesterol, % (n)	2.8 (3,808)	2.8 (4,328)	2.4 (12,583)	2.6 (7,043)	3.2 (4,452)

7.3.1.3 Total breastfeeding duration

The baseline characteristics of parous women are shown overall and by total breastfeeding duration in Table 7.3. The mean age of childbearing women at recruitment was 56.5 (SD 4.8) years, and their mean age at first birth was 23.9 (SD 4.3) years. Mean parity among childbearing women was 2.4 (SD 1.0), and 68% of them reported ever breastfeeding their children. Among women who had ever breastfed their children, the mean total duration of breastfeeding was 8.1 (SD 9.4) months. Women who breastfed their children for longer were less likely to smoke and more likely to live in affluent areas, have any educational qualifications, drink alcohol, and engage in strenuous physical activity than those who breastfed for less than 6 months or not at all.

Table 7.3 Baseline characteristics of parous women, overall and by total breastfeeding duration

	Total breastfeeding duration, months					All parous women
	0	1	2 to 5	6 to 11	12+ (mean 20)	
Women, n	280,594	116,001	193,272	138,173	147,111	1,105,881
Age at recruitment, mean (SD), years	55.7 (4.4)	56.6 (4.6)	57.0 (4.8)	57.1 (5.0)	56.7 (5.2)	56.5 (4.8)
Lifestyle and anthropometric factors						
Current smoker, % (n)	24.2 (64,228)	20.7 (22,818)	19.8 (36,247)	16.5 (21,521)	12.8 (17,824)	20.8 (216,474)
Drinker of alcoholic beverages, % (n)	62.4 (174,100)	65.2 (75,260)	66.8 (128,465)	68.9 (94,686)	68.6 (100,392)	64.9 (711,885)
No strenuous physical activity, % (n)	65.4 (176,899)	62.6 (70,439)	59.6 (111,348)	55.5 (74,348)	52.4 (74,916)	60.5 (644,575)
Body mass index, mean (SD), kg/m^2	26.4 (4.7)	26.2 (4.6)	26.1 (4.5)	26.0 (4.4)	25.8 (4.4)	26.2 (4.6)
Height, mean (SD), cm	161.5 (6.5)	161.9 (6.4)	162.1 (6.4)	162.5 (6.5)	162.8 (6.5)	162.0 (6.6)
Socio-economic factors						
Most deprived fifth, % (n)	22.2 (61,827)	18.6 (21,377)	17.7 (34,037)	15.0 (20,634)	14.4 (21,050)	19.4 (212,655)
No educational qualifications, % (n)	50.6 (138,679)	43.7 (49,696)	40.9 (77,464)	33.7 (45,634)	26.9 (38,762)	44.3 (476,314)
Reproductive factors						
Age at menarche, mean (SD), years	13.0 (1.6)	12.9 (1.6)	13.0 (1.6)	13.0 (1.6)	13.0 (1.6)	13.0 (1.6)
Age at first birth, mean (SD), years	24.0 (4.2)	23.7 (4.2)	23.6 (4.2)	24.1 (4.4)	24.6 (4.5)	23.9 (4.3)
Age at natural menopause, mean (SD), years	48.8 (4.3)	49.2 (4.2)	49.3 (4.2)	49.5 (4.2)	49.7 (4.1)	49.2 (4.2)
Pre/perimenopausal, % (n)	13.7 (38,337)	11.4 (13,175)	11.1 (21,393)	12.7 (17,530)	17.2 (25,198)	13.0 (143,443)
Parity, mean (SD)	2.2 (0.9)	2.2 (1.0)	2.4 (0.9)	2.4 (1.0)	2.9 (1.1)	2.4 (1.0)
Ever use of contraceptive pills, % (n)	62.1 (172,519)	63.1 (72,688)	63.4 (121,560)	62.9 (86,337)	62.0 (90,556)	61.6 (673,465)
Ever use of hormone replacement therapy, % (n)	51.6 (143,377)	53.6 (61,655)	53.5 (102,490)	51.0 (69,865)	45.5 (66,418)	50.9 (556,709)
Healthcare factors						
Self-reported treatment for hypertension, % (n)	15.1 (42,448)	15.0 (17,433)	14.7 (28,433)	14.0 (19,349)	12.9 (18,945)	14.6 (160,985)
Self-reported treatment for diabetes mellitus, % (n)	2.3 (6,526)	1.8 (2,108)	1.8 (3,465)	1.7 (2,287)	1.7 (2,541)	2.1 (22,684)
Self-reported treatment for high cholesterol, % (n)	3.0 (8,358)	2.9 (3,394)	3.0 (5,799)	2.8 (3,843)	2.4 (3,532)	2.6 (28,515)

7.3.2 Risk of cardiovascular multimorbidity and complex cardiovascular multimorbidity in relation to reproductive factors

7.3.2.1 Age at menarche

Figure 7.1 shows minimally and multivariable adjusted HRs and 95% floated confidence intervals (FCIs) for CVM and complex CVM by age at menarche. In the minimally adjusted models, both early and late menarche appeared to be associated with a higher risk of CVM and complex CVM. Early menarche (<12 years vs. 13 years [referent]) was associated with a HR of 1.25 (95% CI 1.23-1.27) for CVM and 1.33 (1.29-1.38) for complex CVM. Late menarche (15+ years vs. 13 years [referent]) was associated with a HR of 1.11 (1.10-1.13) for CVM and 1.11 (1.07-1.15) for complex CVM.

However, multivariable adjustment substantially attenuated these results. The multivariable adjusted HRs associated with early menarche (<12 years) were 1.13 (1.10-1.15) and 1.18 (1.13-1.23) for CVM and complex CVM, respectively, and late menarche was no longer associated with a higher risk of both CVM outcomes. The attenuation in the LR statistics indicated that around 75% of the association with CVM and 65% of the association with complex CVM found in the minimally adjusted models could potentially be accounted for by multivariable adjustment.

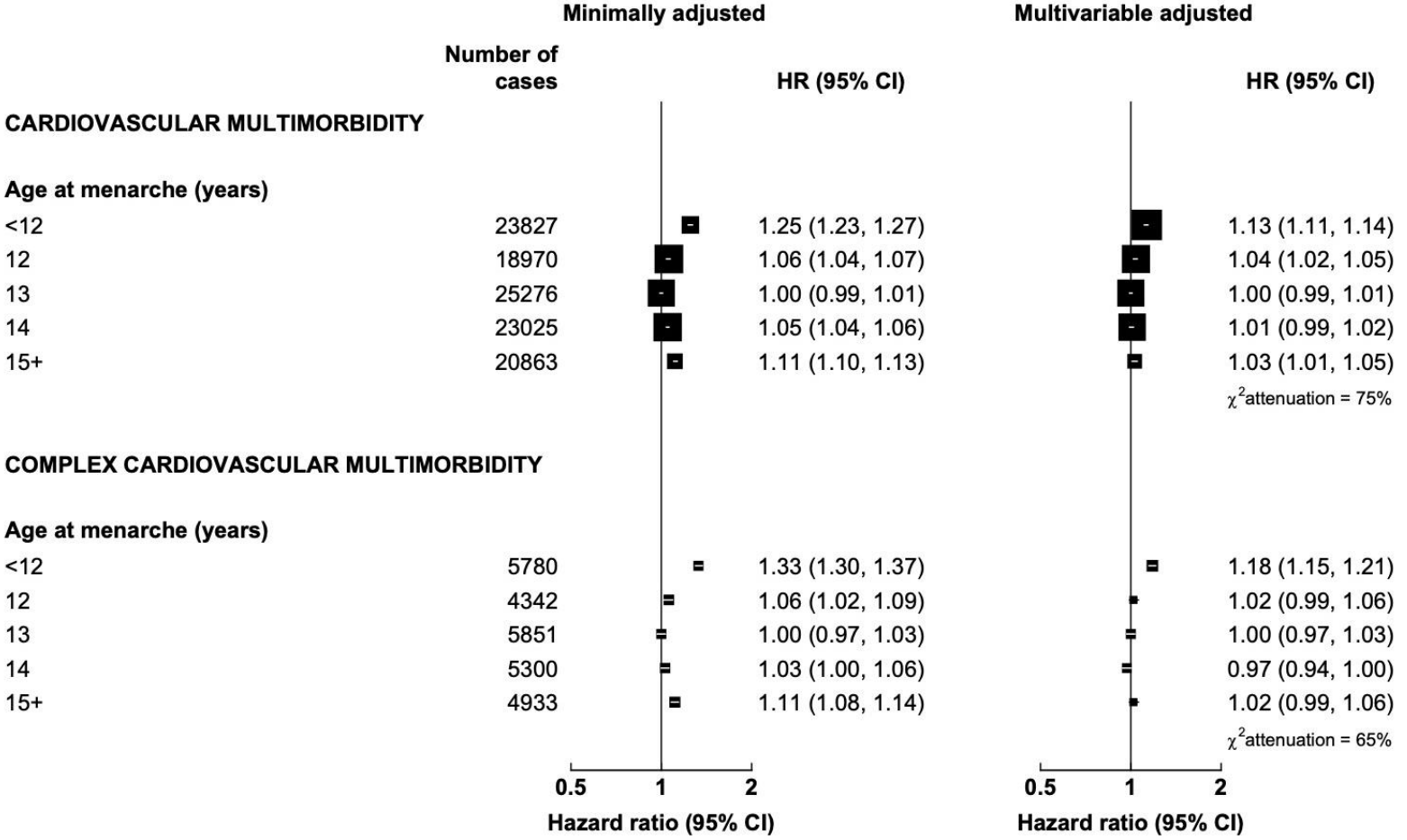


Figure 7.1 Relative risks of incident cardiovascular multimorbidity and complex cardiovascular multimorbidity in relation to age at menarche

Minimally adjusted were stratified by age at recruitment and area of residence only.
Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy.
$\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models.
The size of the square boxes was inversely proportional to the standard error of the effect estimate. Floated confidence intervals were displayed.

7.3.2.2 Parity

In general, parous women did not have a higher risk of CVM or complex CVM compared to nulliparous women. When stratifying by age and region of recruitment, being parous (vs. nulliparous) was associated with a HR of 1.05 (95% CI 1.03-1.07) for CVM. This was reduced to 0.99 (0.97-1.01) after multivariable adjustment ($\chi^2_{attenuation}$ =98%). For complex CVM, the minimally adjusted HR was 1.03 (0.99-1.07), while the multivariable adjusted HR was 0.94 (0.90-0.98) [$\chi^2_{attenuation}$ =-319%]; the LR statistics corresponding to the minimally and multivariable adjusted models were 1.7 and 7.1, respectively.

Figure 7.2 shows HRs and 95% FCIs for CVM and complex CVM by different levels of parity. In the minimally adjusted models, having two children (vs. none) was associated with a HR of 0.94 (95% CI 0.92-0.96) for CVM and 0.89 (0.86-0.93) for complex CVM. Having four or more children (vs. none) was associated with a HR of 1.38 (1.35-1.41) for CVM and 1.42 (1.35-1.49) for complex CVM. Multivariable adjustment attenuated some of these results: the HRs for both CVM and complex CVM associated with having 2 children were unchanged, while the HRs associated with having four or more children (vs. nulliparity) were reduced to 1.11 (1.08-1.14) for CVM ($\chi^2_{attenuation}$ =86%) and 1.05 (1.00-1.11) for complex CVM ($\chi^2_{attenuation}$ =88%).

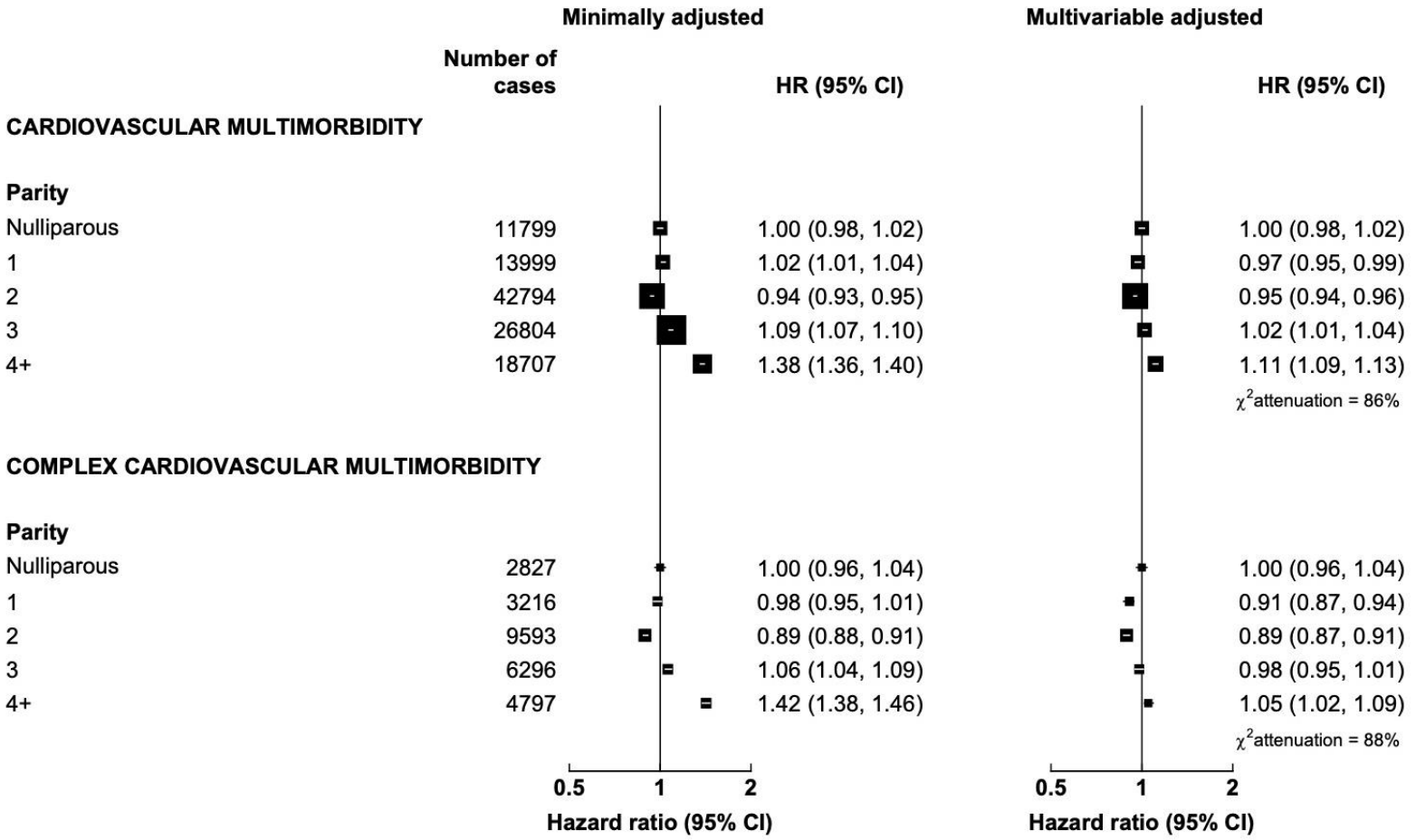


Figure 7.2 Relative risks of incident cardiovascular multimorbidity and complex cardiovascular multimorbidity in relation to parity

Minimally adjusted models were stratified by age at recruitment and area of residence only.
Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, and ever use of menopausal hormone therapy.
$\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models.
The size of the square boxes was inversely proportional to the standard error of the effect estimate. Floated confidence intervals were displayed.

7.3.2.3 Breastfeeding in parous women

Among parous women, ever breastfeeding was associated with a somewhat lower risk of CVM and complex CVM compared to never breastfeeding. The minimally adjusted HR associated with ever breastfeeding was 0.83 (0.81-0.84) for CVM and 0.76 (0.74-0.79) for complex CVM. Multivariable adjustment attenuated the HRs to 0.92 (0.90-0.94) for CVM ($\chi^2_{attenuation}$ =80%) and 0.87 (0.84-0.91) for complex CVM ($\chi^2_{attenuation}$ =85%).

Figure 7.3 displays the HRs and 95% FCIs for CVM and complex CVM in parous women, by total breastfeeding duration. Minimally adjusted models showed that with increasing total breastfeeding duration, the risk of both CVM and complex CVM decreased. However, multivariable adjustment substantially attenuated the magnitude of the HRs. For CVM, the HR associated with ≥12 months of breastfeeding (vs. never breastfeeding) was attenuated from 0.73 (95% CI 0.72-0.75) in the minimally adjusted model to 0.89 (0.86-0.91) in the multivariable adjusted model ($\chi^2_{attenuation}$ =88%). For complex CVM, the HR associated with ≥12 months of breastfeeding (vs. never breastfeeding) was attenuated from 0.68 (0.65-0.71) in the minimally adjusted model to 0.84 (0.79-0.88) in the multivariable adjusted model ($\chi^2_{attenuation}$ =81%).

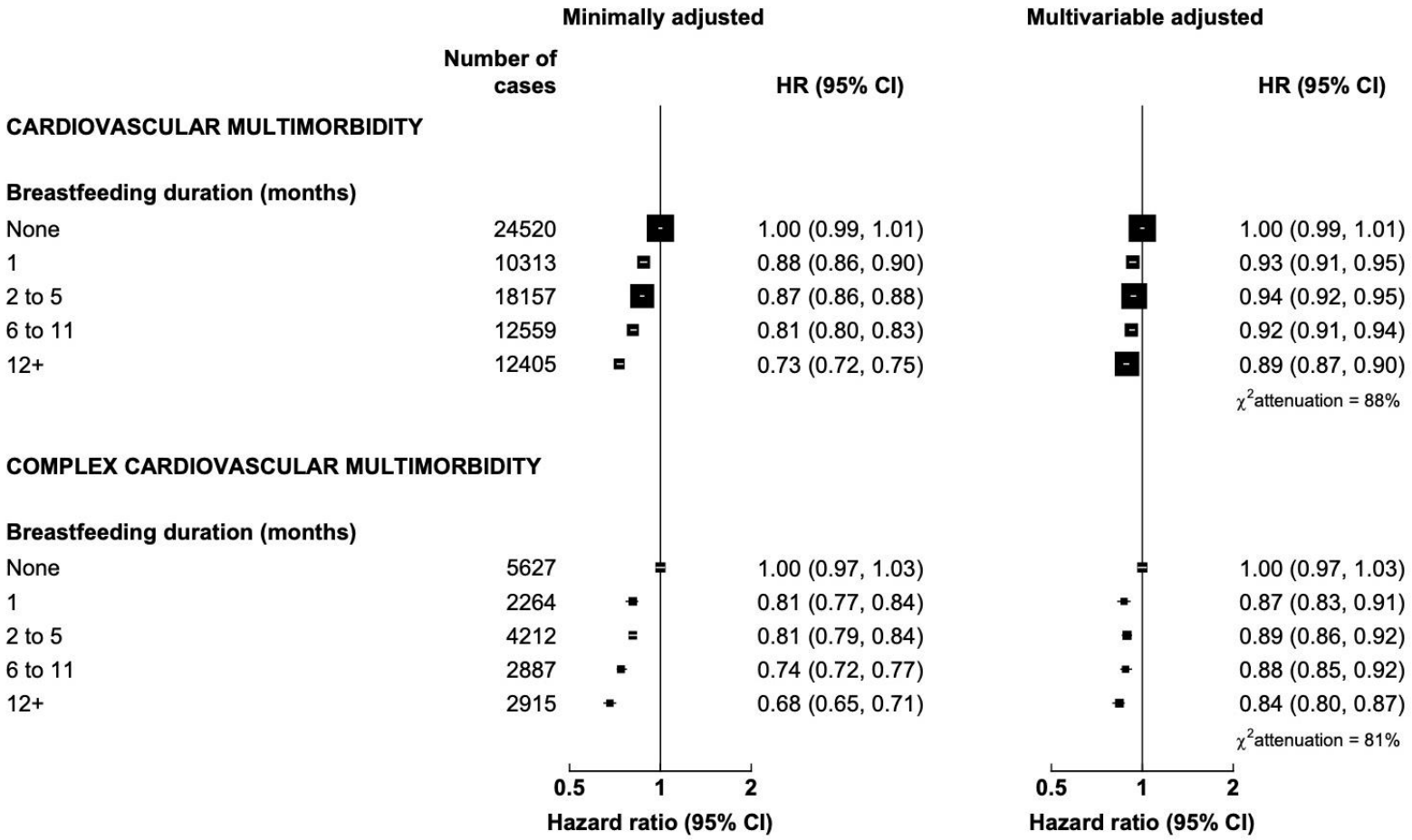


Figure 7.3 Relative risk of incident cardiovascular multimorbidity and complex cardiovascular multimorbidity in relation to total breastfeeding duration in parous women

Minimally adjusted models were stratified by age at recruitment and area of residence and adjusted for parity.
Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, and ever use of menopausal hormone therapy.
$\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models.
The size of the square boxes was inversely proportional to the standard error of the effect estimate. Floated confidence intervals were displayed.

7.3.2.4 Combined effect of parity and breastfeeding

Table 7.4 shows the risk of incident CVM and complex CVM compared across three groups: nulliparous women (referent), parous women who never breastfed, and parous women who ever breastfed. Nulliparous women will have never breastfed, which means these analyses show the relative risks associated with parity, with and without the effect of breastfeeding.

Parous women who had never breastfed children (vs. nulliparous women) had a somewhat higher risk of both CVM and complex CVM in the minimally adjusted models. In contrast, parous women who had ever breastfed children (vs. nulliparous women) had a slightly lower risk of both CVM outcomes.

There was substantial attenuation after multivariable adjustment, such that parous women who had never breastfed children had almost no difference in the risk of both CVM outcomes compared to nulliparous women. The HRs for both outcomes in parous women who had ever breastfed (vs. nulliparous women) remained virtually unchanged after multivariable adjustment.

Table 7.4 The combined effect of parity and breastfeeding on incident CVM and complex CVM

	Minimally adjusted	Multivariable adjusted	
CVM	**HR (95% CI)**	**HR (95% CI)**	
Nulliparous (Never breastfed)	1.00 (reference)	1.00 (reference)	
Parous - Never breastfed	1.13 (1.11-1.16)	1.03 (1.01-1.06)	
Parous - Ever breastfed	0.97 (0.95-0.99)	0.97 (0.95-0.99)	$\chi^2_{attenuation}$ = 86%
Complex CVM	**HR (95% CI)**	**HR (95% CI)**	
Nulliparous (Never breastfed)	1.00 (reference)	1.00 (reference)	
Parous - Never breastfed	1.15 (1.10-1.21)	1.02 (0.97-1.07)	
Parous - Ever breastfed	0.92 (0.88-0.96)	0.91 (0.87-0.95)	$\chi^2_{attenuation}$ = 75%

7.3.3 Risk of selected cardiovascular multimorbidity pairs in relation to reproductive factors

Figure 7.4 and Figure 7.5 show the multivariable adjusted HRs and FCIs for the six selected pairs of CVDs (i.e. CVM pairs) in relation to reproductive factors. The overall shape and magnitude of the associations for CVM pairs were similar to those for CVM (see Figure 7.1 - Figure 7.3, page 235-9). The exceptions were the associations of parity with IHD-ARR, AF-HF, and AF-STR. For IHD-ARR, any level of parity (including having 2 children) was associated with a higher risk than nulliparity, although having 2 children (vs. none) was associated with a lower risk of all other CVM outcomes (i.e. CVM, complex CVM, and other CVM pairs). For AF-HF and AF-STR, any level of parity (including having ≥4 children) was associated with a lower relative risk compared to nulliparity, while in contrast, having ≥4 children (vs. none) was always associated with a higher risk of all other CVM outcomes. Regardless of these variations, multivariable adjustment most often accounted for more than 70% of the relationships between the six selected CVM pairs and reproductive factors ($\chi^2_{attenuation}$ =59-91%) as with CVM. The level of attenuation in the LR statistics indicated that residual confounding was likely to be substantial.

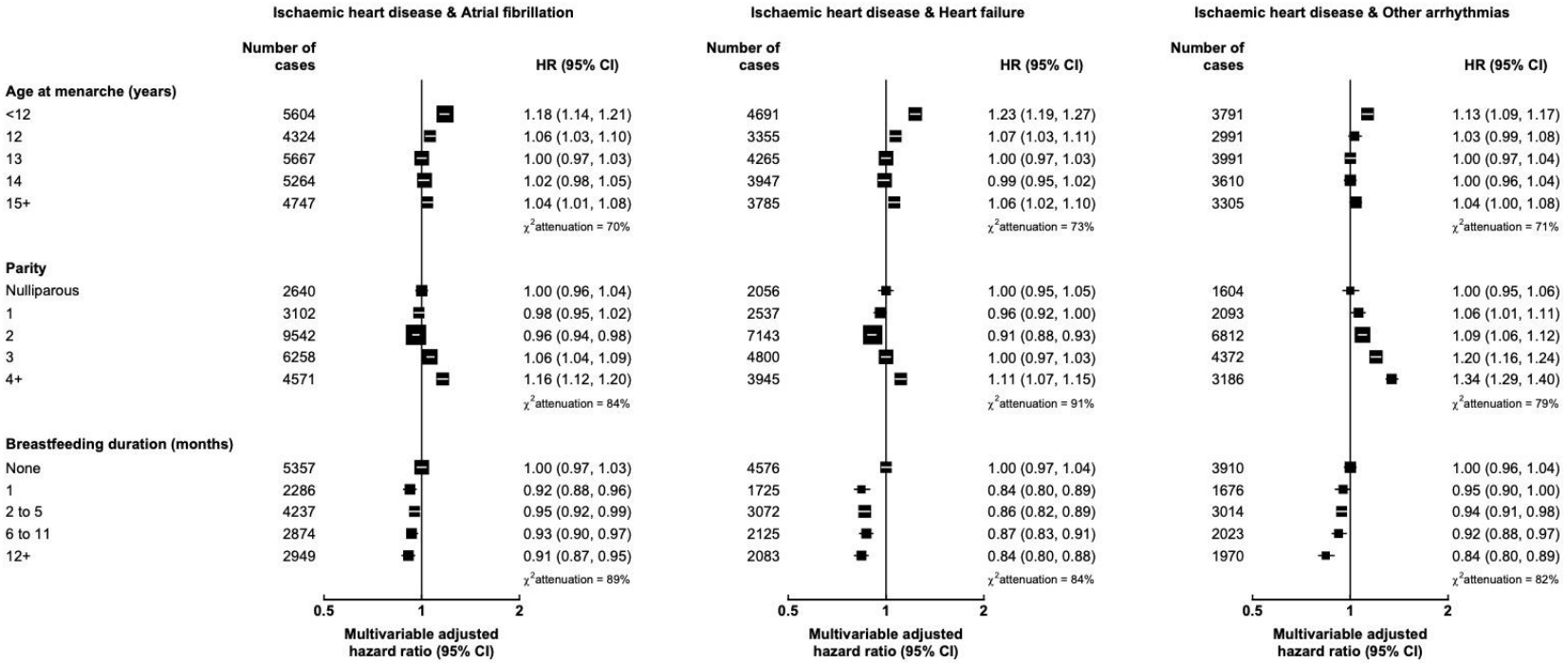


Figure 7.4 Relative risks of CVM pairs including ischaemic heart disease in relation to reproductive factors

Minimally adjusted models for age at menarche and parity were stratified by age at recruitment and area of residence only. Minimally adjusted models for breastfeeding duration were also adjusted for parity.
Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy.
$\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models.
The size of the square boxes was inversely proportional to the standard error of the effect estimate. Floated confidence intervals were displayed.

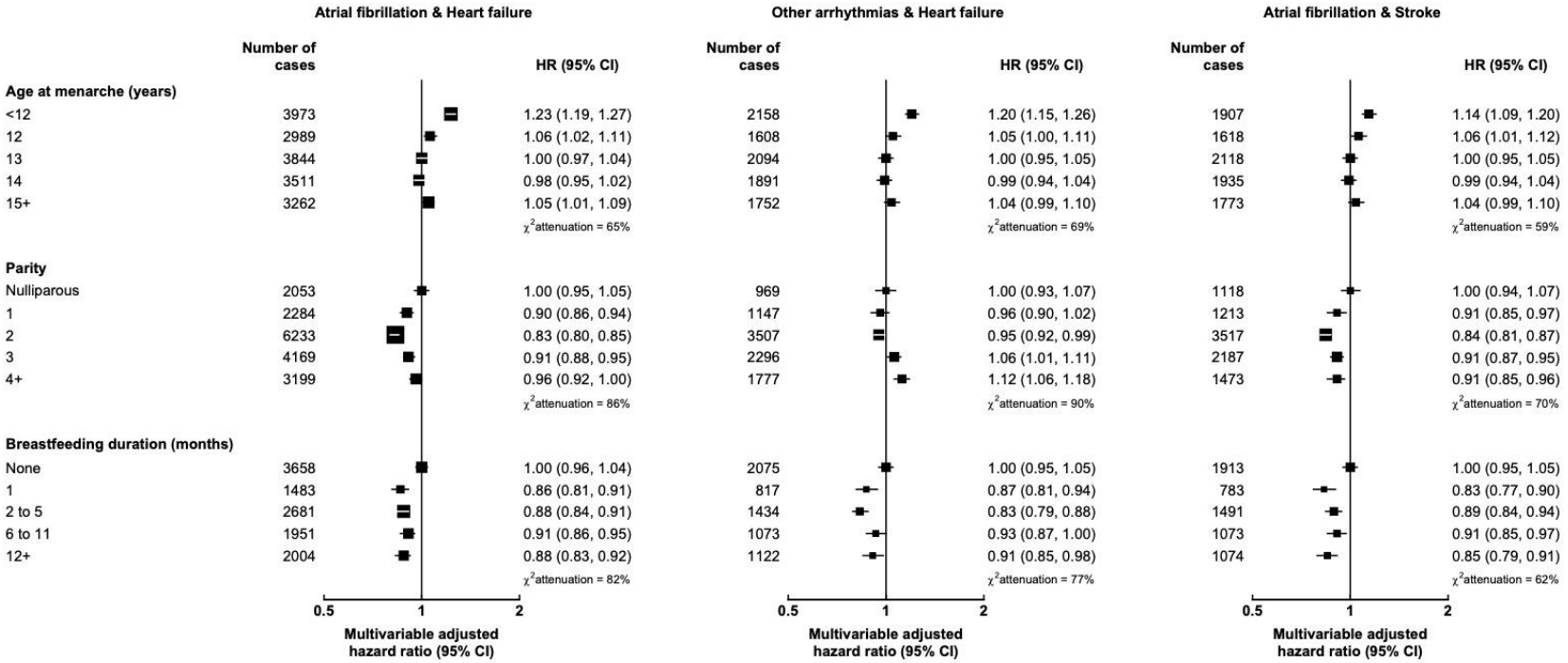


Figure 7.5 Relative risks of other CVM pairs including cardiac arrhythmias in relation to reproductive factors

Minimally adjusted models for age at menarche and parity were stratified by age at recruitment and area of residence only. Minimally adjusted models for breastfeeding duration were also adjusted for parity.
Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy.
$\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models.
The size of the square boxes was inversely proportional to the standard error of the effect estimate. Floated confidence intervals were displayed.

7.3.4 The role of covariates in accounting for reproductive factor associations with cardiovascular multimorbidity

As reported in the previous sections, the associations between reproductive factors and all CVM outcomes were greatly attenuated after multivariable adjustment ($\chi^2_{attenuation}$ =59-91%, Figure 7.1-Figure 7.5). The role of other factors (particularly, known cardiovascular risk factors) in the associations between the selected reproductive factors and CVM and complex CVM was explored in more detail by assessing the changes in the LR statistics after adjusting for covariates, both individually and in different combinations.

7.3.4.1 Age at menarche

Table 7.5 shows the relationship of age at menarche with the risk of CVM after adjusting for socioeconomic factors, known cardiovascular risk factors, and reproductive factors in addition to minimal adjustment.

Compared to the minimally adjusted model (stratified by age and region of recruitment), additionally adjusting for deprivation or education alone attenuated the LR statistic by 10% and 9%, respectively. Adjusting for BMI alone led to the largest decline in the LR statistic of 52%, followed by adjusting for smoking alone, which resulted in a 24% decline in the LR statistic. Adjusting individually for height, parity, or use of menopausal hormone therapy had virtually no effect on the relationship between age at menarche and CVM.

Simultaneously adjusting for four behavioural factors (i.e. smoking, alcohol consumption, physical activity, and BMI) resulted in a 79% decline in the LR statistic, and adding the two socioeconomic factors (i.e. area deprivation and education) led to an 80% decline in the LR statistic.

The role of these covariates in the relationship between age at menarche and complex CVM were very similar (Appendix F, Table F1, page 355): adjusting for BMI also led to

the largest reduction of 52% in the LR statistic among individual factors, and simultaneously adjusting for the four health behaviours and socioeconomic factors resulted in a 71% decline in the LR statistic.

Table 7.5 Association of age at menarche with incident cardiovascular multimorbidity after adjustment for various covariates

	LR[1]	attenuation in LR[1] (%)	Age at menarche (years)				
			<12	12	13	14	15+
Number of incident CVM events			23,827	18,970	25,276	23,025	20,863
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	678	-	1.25 (1.23 - 1.27)	1.06 (1.04 - 1.08)	1.00 (ref.)	1.05 (1.03 - 1.07)	1.11 (1.09 - 1.13)
Age, region, deprivation	610	10	1.23 (1.21 - 1.26)	1.06 (1.04 - 1.08)	1.00 (ref.)	1.03 (1.01 - 1.05)	1.08 (1.06 - 1.10)
Age, region, education	615	9	1.23 (1.21 - 1.26)	1.06 (1.04 - 1.08)	1.00 (ref.)	1.02 (1.00 - 1.04)	1.06 (1.04 - 1.08)
Age, region, smoking	512	24	1.22 (1.20 - 1.24)	1.06 (1.04 - 1.08)	1.00 (ref.)	1.03 (1.01 - 1.05)	1.07 (1.05 - 1.09)
Age, region, alcohol	614	9	1.24 (1.22 - 1.26)	1.06 (1.04 - 1.08)	1.00 (ref.)	1.04 (1.02 - 1.06)	1.10 (1.08 - 1.12)
Age, region, physical activity	603	11	1.24 (1.22 - 1.26)	1.06 (1.04 - 1.08)	1.00 (ref.)	1.03 (1.02 - 1.05)	1.09 (1.07 - 1.11)
Age, region, body mass index	325	52	1.15 (1.13 - 1.18)	1.03 (1.01 - 1.05)	1.00 (ref.)	1.06 (1.04 - 1.08)	1.13 (1.11 - 1.15)
Age, region, height	683	-1	1.26 (1.23 - 1.28)	1.06 (1.04 - 1.08)	1.00 (ref.)	1.05 (1.03 - 1.07)	1.11 (1.09 - 1.13)
Age, region, parity	672	1	1.25 (1.23 - 1.27)	1.06 (1.04 - 1.08)	1.00 (ref.)	1.05 (1.03 - 1.06)	1.11 (1.08 - 1.13)
Age, region, menopausal hormone therapy	656	3	1.25 (1.23 - 1.27)	1.06 (1.04 - 1.08)	1.00 (ref.)	1.05 (1.03 - 1.07)	1.11 (1.09 - 1.13)
Age, region, health behaviours[2]	145	79	1.12 (1.10 - 1.14)	1.03 (1.01 - 1.05)	1.00 (ref.)	1.03 (1.01 - 1.05)	1.07 (1.05 - 1.09)
Age, region, health behaviours[2], plus socioeconomic factors	138	80	1.12 (1.09 - 1.14)	1.03 (1.01 - 1.05)	1.00 (ref.)	1.01 (0.99 - 1.03)	1.04 (1.02 - 1.06)
Fully adjusted[3]	169	75	1.13 (1.10 - 1.15)	1.04 (1.02 - 1.06)	1.00 (ref.)	1.01 (0.99 - 1.03)	1.03 (1.01 - 1.05)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, education, deprivation, height, parity, and ever use of menopausal hormone therapy;
CVM = cardiovascular multimorbidity; ref. means reference category

7.3.4.2 Parity

Table 7.6 shows the attenuation in the LR statistics of the minimally adjusted models for the relationship between parity with CVM, after additionally adjusting for various covariates (both separately and simultaneously). Compared to the minimally adjusted model, adjusting additionally only for deprivation or education attenuated the LR statistic by 41% and 30%, respectively. Of the remaining individual factors, adjusting for BMI alone led to the largest decline of 40% in the LR statistic, followed by adjusting for smoking alone, resulting in a 38% decline in the LR statistic. Adjusting for alcohol consumption and physical activity individually each led to an 18% decline in the LR statistic. Adjusting individually for height or use of menopausal hormone therapy had no effect on the relationship between parity and CVM. Simultaneously adjusting for the four behavioural factors resulted in a 76% decline in the LR statistic and adding the two socioeconomic factors led to an 86% decline in the LR statistic. The findings were nearly identical for complex CVM (Appendix F, Table F2, page 356).

Table 7.6 Association of parity with incident cardiovascular multimorbidity after adjustment for various covariates

	LR[1]	attenuation in LR[1] (%)	Parity				
			Nulliparous	**1**	**2**	**3**	**4+**
Number of incident CVM events			11,799	13,999	42,794	26,804	18,707
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	1889	-	1.00 (ref.)	1.02 (1.00 - 1.05)	0.94 (0.92 - 0.96)	1.09 (1.07 - 1.11)	1.38 (1.35 - 1.41)
Age, region, deprivation	1109	41	1.00 (ref.)	1.02 (1.00 - 1.05)	0.95 (0.93 - 0.97)	1.08 (1.05 - 1.10)	1.28 (1.25 - 1.31)
Age, region, education	1326	30	1.00 (ref.)	0.98 (0.95 - 1.00)	0.91 (0.89 - 0.93)	1.04 (1.02 - 1.06)	1.27 (1.24 - 1.30)
Age, region, smoking	1163	38	1.00 (ref.)	0.99 (0.97 - 1.02)	0.94 (0.92 - 0.96)	1.06 (1.04 - 1.09)	1.28 (1.25 - 1.31)
Age, region, alcohol	1557	18	1.00 (ref.)	1.02 (0.99 - 1.04)	0.94 (0.92 - 0.96)	1.09 (1.06 - 1.11)	1.34 (1.31 - 1.37)
Age, region, physical activity	1551	18	1.00 (ref.)	1.01 (0.99 - 1.04)	0.94 (0.92 - 0.96)	1.09 (1.07 - 1.12)	1.35 (1.32 - 1.38)
Age, region, body mass index	1130	40	1.00 (ref.)	1.02 (1.00 - 1.05)	0.95 (0.93 - 0.97)	1.07 (1.05 - 1.10)	1.29 (1.26 - 1.32)
Age, region, height	1842	3	1.00 (ref.)	1.03 (1.00 - 1.05)	0.94 (0.92 - 0.96)	1.09 (1.07 - 1.12)	1.39 (1.36 - 1.42)
Age, region, menopausal hormone therapy	1847	2	1.00 (ref.)	1.02 (1.00 - 1.05)	0.94 (0.92 - 0.96)	1.09 (1.06 - 1.11)	1.38 (1.35 - 1.41)
Age, region, health behaviours[2]	450	76	1.00 (ref.)	0.98 (0.95 - 1.01)	0.95 (0.93 - 0.97)	1.04 (1.02 - 1.07)	1.16 (1.13 - 1.19)
Age, region, health behaviours[2], plus socioeconomic factors	261	86	1.00 (ref.)	0.96 (0.94 - 0.99)	0.94 (0.92 - 0.96)	1.02 (1.00 - 1.05)	1.10 (1.07 - 1.13)
Fully adjusted[3]	261	86	1.00 (ref.)	0.97 (0.94 - 0.99)	0.95 (0.93 - 0.97)	1.02 (1.00 - 1.05)	1.11 (1.08 - 1.14)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, education, deprivation, height, and ever use of menopausal hormone therapy;
CVM = cardiovascular multimorbidity; ref. means reference category

7.3.4.3 Total breastfeeding duration

Table 7.6 shows the attenuation in the LR statistics of the minimally adjusted models for the relationship between total breastfeeding duration with CVM in parous women, after additionally adjusting for various covariates.

Compared to the minimally adjusted model (adjusted for age, region, and parity), adjusting additionally only for deprivation or education attenuated the LR statistic by 27% and 52%, respectively. Of the remaining individual factors, adjusting for smoking alone led to the largest decline of 47% in the LR statistic, for physical activity alone resulted in a 35% reduction, and for BMI individually led to a 31% decline. Adjusting individually for height or use of menopausal hormone therapy had no effect on the relationship between parity and CVM.

Simultaneously adjusting for the four behavioural factors resulted in an 81% decline in the LR statistic and adding the two socioeconomic factors led to an 89% decline in the LR statistic. Again, the findings were almost the same in the relationship between breastfeeding duration and complex CVM (Appendix F, Table F3, page 357).

Table 7.7 Association of total breastfeeding duration with incident cardiovascular multimorbidity in parous women after adjustment for various covariates

In parous women	LR[1]	attenuation in LR[1] (%)	Total breastfeeding duration (months)				
			None	1	2 to 5	6 to 11	12+
Number of incident CVM events			24,520	10,313	18,157	12,559	12,405
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region, parity (minimally adjusted)	823	-	1.00 (ref.)	0.88 (0.86 - 0.90)	0.87 (0.85 - 0.88)	0.81 (0.80 - 0.83)	0.73 (0.72 - 0.75)
Age, region, parity, deprivation	598	27	1.00 (ref.)	0.89 (0.87 - 0.91)	0.88 (0.87 - 0.90)	0.84 (0.82 - 0.86)	0.76 (0.75 - 0.78)
Age, region, parity, education	393	52	1.00 (ref.)	0.90 (0.88 - 0.92)	0.90 (0.88 - 0.91)	0.86 (0.84 - 0.88)	0.80 (0.78 - 0.82)
Age, region, parity, smoking	435	47	1.00 (ref.)	0.90 (0.87 - 0.92)	0.88 (0.87 - 0.90)	0.85 (0.83 - 0.87)	0.80 (0.78 - 0.81)
Age, region, parity, alcohol	692	16	1.00 (ref.)	0.89 (0.87 - 0.91)	0.88 (0.86 - 0.90)	0.83 (0.81 - 0.85)	0.75 (0.73 - 0.77)
Age, region, parity, physical activity	534	35	1.00 (ref.)	0.89 (0.87 - 0.92)	0.89 (0.87 - 0.91)	0.85 (0.83 - 0.87)	0.77 (0.75 - 0.79)
Age, region, parity, body mass index	565	31	1.00 (ref.)	0.90 (0.87 - 0.92)	0.89 (0.87 - 0.91)	0.84 (0.82 - 0.86)	0.76 (0.75 - 0.78)
Age, region, parity, height	829	-1	1.00 (ref.)	0.88 (0.86 - 0.90)	0.87 (0.85 - 0.89)	0.81 (0.79 - 0.83)	0.73 (0.71 - 0.75)
Age, region, parity, menopausal hormone therapy	797	3	1.00 (ref.)	0.88 (0.86 - 0.90)	0.87 (0.85 - 0.88)	0.81 (0.80 - 0.83)	0.73 (0.72 - 0.75)
Age, region, parity, health behaviours	157	81	1.00 (ref.)	0.92 (0.90 - 0.95)	0.93 (0.91 - 0.95)	0.91 (0.89 - 0.93)	0.86 (0.84 - 0.88)
Age, region, parity, health behaviours, socioeconomic factors	93	89	1.00 (ref.)	0.93 (0.91 - 0.96)	0.94 (0.92 - 0.96)	0.93 (0.91 - 0.95)	0.89 (0.87 - 0.91)
Fully adjusted[3]	103	88	1.00 (ref.)	0.93 (0.90 - 0.95)	0.93 (0.91 - 0.95)	0.92 (0.90 - 0.95)	0.89 (0.86 - 0.91)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, education, deprivation, height, and ever use of menopausal hormone therapy;
CVM = cardiovascular multimorbidity; ref. means reference category

7.3.5 Sensitivity analyses

To assess the potential role of common individual CVDs in driving the association between reproductive factors and CVM, sensitivity analyses were conducted, each leaving out one of the six selected common individual CVDs from the definition of CVM (Appendix F, Figure F1-F2, page 358-9). This did not materially change any of the associations between reproductive factors and CVM, and the level of LR attenuation remained equally high ($\chi^2_{attenuation}$ =71-89%). The associations examined in these analyses were of small magnitude, and likely to be explained by residual mediation or confounding, but overall there was no evidence that any of the six individual CVDs substantially drove the relationships between reproductive factors and CVM.

Associations between common individual CVDs and the selected reproductive factors were also examined (Appendix F, Figure F3-F4, page 360-1). Most associations were similar to those for CVM (Figure 7.1 - Figure 7.3, page 235-9). For age at menarche, one exception was the relative risk of heart failure associated with early age at menarche (age <12 vs. 13 years, HR 1.21, 95% CI 1.17-1.25; Figure F3, page 360), which was somewhat greater in magnitude than that for CVM (1.13, 1.10-1.15). In the case of parity, the relative risks associated with highest parity (vs. nulliparity) for ischaemic heart disease (1.27, 1.24-1.31; Figure F3, page 360) and other arrhythmias (1.21, 1.16-1.26; Figure F4, page 361) were somewhat greater in magnitude than those for CVM (1.11, 1.08-1.14). And for breastfeeding duration, the relative risks of individual CVDs were similar to those for CVM. LR attenuation indicated that substantial proportions of all reproductive factor associations with individual CVDs could be accounted for by multivariable adjustment ($\chi^2_{attenuation}$ =59-93%).

7.4 Discussion

7.4.1 Key findings and interpretation of results

7.4.1.1 Overall summary

In this cohort of 1.2 million UK women, early menarche (<12 years vs. 13 years), high parity (≥4 children vs. none), and not having breastfed their children (vs. any duration of breastfeeding) were each associated with a somewhat higher risk of both incident CVM and complex CVM after full adjustment. Attenuation in the LR statistics suggested that the majority of these associations could be accounted for by multivariable adjustment for potential confounders and/or mediators ($\chi^2_{attenuation}$ =65-88%). In addition, applying the absolute scale difference method (as discussed in Section 5.4.4 of Chapter 5, page 175) [13, 14] to those relationships yielded estimates of 42-63%. This further corroborates the absence of substantial independent associations between reproductive factors and CVM.

The associations between these reproductive factors and CVM pairs were similar to those for CVM, except for parity. Any level of parity was associated with a somewhat lower risk of AF-HF or AF-STR compared to nulliparity after multivariable adjustment. Nevertheless, reproductive factor associations with CVM pairs were also substantially attenuated by multivariable adjustment ($\chi^2_{attenuation}$ =59-91%). For these associations, applying the absolute scale difference method produced estimates of 54-78%.

Analyses of individual covariates suggested that BMI and smoking were particularly important confounders and/or mediators in the observed associations between reproductive factors with CVM and complex CVM. Overall, these results did not offer any evidence of strong independent associations between reproductive factors and CVM outcomes. The magnitude of the associations was small, and the level of attenuation in the LR statistics

indicated that substantial parts of any remaining association may still be explained by residual confounding or mediation.

7.4.1.2 Age at menarche

At the time of writing, no other study has assessed the relationship between age at menarche and any CVM outcomes. It was therefore difficult to appraise whether the associations between age at menarche and the CVM outcomes in this chapter were consistent with evidence in the scientific literature. The most relevant evidence was a UK Biobank study of 267,440 women aged 40-69 years followed-up for a median of 7 years, which reported that the age-adjusted HR associated with early menarche (<12 years vs. ≥12 years [referent]) for composite CVD (defined as the first occurrence of either coronary heart disease or stroke) was 1.18 (95% CI 1.09-1.29), while the HR additionally adjusted for area deprivation, smoking, blood pressure, BMI, and diabetes was 1.10 (1.01-1.30). [15] Although direct quantitative comparisons could not be made to the findings of this chapter due to the differences in the definition of the outcome (CVD as opposed to CVM) and the reference group (≥12 years as opposed to 13 years as the referent), the qualitative finding that controlling for the effects of socioeconomic and behavioural factors removed most of the observed associations was consistent with this chapter.

The composite CVD outcome in the UK Biobank study was different from CVM because it included people with only one CVD, who may have different associations with age at menarche compared with people with multiple CVDs. In terms of relevant findings on individual CVDs, previous observational studies have consistently found early menarche to be associated with a somewhat higher risk of coronary heart disease and stroke [5, 11, 15, 16], including in the Million Women Study [5]. One study found that there was no statistically significant association with heart failure after multivariable adjustment for potential confounders and mediators [17].

Previous observational studies have found that higher BMI (both before and after menarche) was associated with early menarche and have postulated that BMI is an important factor in the association between age at menarche and CVD risk. [11, 16, 18, 19] Furthermore, genome-wide association studies have found substantial overlap between the genes associated with age at menarche and those associated with BMI [20, 21], and a recent Mendelian randomisation study suggested that adiposity may mediate most of the relationship between age at menarche and adverse metabolic profile in adulthood [22]. These findings were consistent with this chapter, where BMI at study recruitment alone could account for more than half of the relationship between age at menarche and incident CVM.

7.4.1.3 Parity

There were no CVM-specific studies in relation to parity in the literature to date, so it was not possible to directly compare previous findings with the results of this chapter. The most comparable studies were prospective studies on composite CVD outcomes. [15, 23, 24] The UK Biobank study described in the previous section (Section 7.4.1.2) also found that women with children did not have a statistically significantly higher risk of CVD (i.e. coronary heart disease or stroke) compared to women without children (multivariable adjusted HR 1.08, 95% CI 0.98-1.20). They additionally found that those with children had a higher risk of coronary heart disease but not of stroke. Although the composite CVD outcome included having just one CVD (as opposed to multiple CVDs), these results were consistent with the findings of this chapter that there was no statistically significant association between parity (vs. nulliparity) and CVM, and that associations with parity varied across individual CVDs.

The other prospective studies which examined different levels of parity found that increasing parity was generally associated with a somewhat higher risk of composite CVD.

[23, 24] A study in the USA found that ≥5 pregnancies (vs. none) was associated with a HR of 1.52 (95% CI 1.20-1.93) for CVD (i.e. myocardial infarction, coronary artery vascularisation, or ischaemic stroke), after adjusting for age, ethnicity, income, education, hypertension, hypercholesterolemia, diabetes, smoking status, physical activity, BMI, and age at menarche. [23] Although this study had adjusted for a comprehensive range of covariates, it only had 1,211 cases of CVD in total. Furthermore, Parikh et al. [24] found that in a cohort of 1.3 million Swedish women with 65,204 CVD events (i.e. coronary heart disease, stroke, or heart failure), there was a J-shaped association between parity and CVD, with 2 births representing the lowest level of risk. After adjusting for age, birth year, income, education, and country of birth, having ≥5 pregnancies was associated with a HR of 1.57 (1.52-1.64) compared to having 2 pregnancies. However, this study had not adjusted for any health-related behavioural factors such as BMI and smoking, which were able to account for most of the relationship between parity and CVM reported in this chapter.

The associations of parity with AF-HF and AF-STR were different from those of all other CVM outcomes in this chapter. Parous women had a lower risk of these CVM pairs compared to women with no children. There was no evidence on specific CVM pairs in the scientific literature. Findings from the few studies on the component individual CVDs were often inconsistent with the associations for the CVM pairs in this chapter. For example, a study on the relationship between parity and atrial fibrillation in US women found that compared to nulliparous women, women with increasing numbers of pregnancies had higher risks of atrial fibrillation after adjusting for sociodemographic, behavioural, and cardiovascular risk factors. [25] Furthermore, the Swedish study by Parikh et al. [24] found that the risk of heart failure was lowest in women who had given birth twice (referent), with the highest relative risk in women who had given birth ≥5

times. Studies on stroke often reported findings which conflicted with each other. [9, 15, 24, 26, 27] Many of these studies had small sample sizes, and the magnitude of the relative risks were materially reduced after adjustment for potential confounders such as socioeconomic deprivation and smoking. Taken together with the large LR statistic attenuation observed in this chapter, there was no evidence to suggest that the modest inverse associations of parity with AF-HF and AF-STR were indicative of any real protective effects.

Multiple physiological mechanisms potentially linking parity in mothers with higher CVD risk have been hypothesised, including progressive insulin resistance, gestational weight gain, higher circulating lipid levels, and pregnancy complications like pre-eclampsia. [6, 7, 28, 29] This chapter found that jointly controlling for socioeconomic (i.e. education and area deprivation) and behavioural factors (i.e. smoking, BMI, physical activity, and alcohol consumption) could account for most of the observed associations between parity and CVM. In addition, previous studies found that the associations between the number of children and CVD risk were not significantly different between men and women. [15, 30] Fathers do not undergo the physiological changes of pregnancy but are similarly affected by changes in socioeconomic environments and health-related behaviours as mothers. Hence, the fact that there was no effect modification by sex implied that social or behavioural factors related to childbearing probably contributes more to any association between parity and CVD risk than direct physiological changes due to pregnancy.

7.4.1.4 Total breastfeeding duration

No previous study investigated the relationship between breastfeeding and CVM risk in the literature. However, studies have consistently found that in childbearing women, breastfeeding was associated with a lower risk of having hypertension, diabetes, obesity, ischaemic heart disease, and cerebrovascular disease later in life. [10, 31-37] Some studies

found a dose-response relationship between breastfeeding duration and cardiovascular risk factors such as hypertension and lipid levels [10, 31, 37], but not with incident CVD. In this chapter, breastfeeding children (compared to not breastfeeding) was associated with a somewhat lower risk of CVM, complex CVM, and CVM pairs. There was no dose-response relationship, at least up to 12 months of total breastfeeding duration.

The "reset hypothesis" proposes that breastfeeding resets the adverse metabolic changes women undergo during a normal pregnancy, such that women who do not breastfeed after a pregnancy remain in a dysmetabolic state. [8] Breastfeeding may protect against cardiometabolic diseases via calorie expenditure and the release of cholesterol from the mother's body in breastmilk. [38] Consistent with this hypothesis, a study in a European cohort of 14,917 women with 5,138 cases of coronary heart disease found that although parous women who had never breastfed had a higher risk of coronary heart disease compared to nulliparous women (HR 1.58, 95% CI 1.09-2.30), parous women who had ever breastfed did not have such an elevated risk (1.19, 0.99-1.43) after adjusting for age, study centre, education, smoking status, high blood pressure, cholesterol levels, BMI, and history of diabetes. [39]

In contrast, in this chapter, the combined effects of breastfeeding history and parity did not suggest that breastfeeding protects against the potential adverse cardiometabolic effects of pregnancies. After multivariable adjustment, parous women who never breastfed only had a 3% higher risk of CVM compared to nulliparous women, while parous women who ever breastfed only had a 3% lower risk of CVM compared to nulliparous women. Combined with the finding that adjusting simultaneously for socioeconomic factors and behavioural factors (i.e. smoking, alcohol consumption, physical activity, and BMI) attenuated the HRs for CVM and led to a 89% decline in the LR statistic, it appeared unlikely that the

somewhat inverse relationship between breastfeeding duration and CVM risk was causal rather than due to residual confounding.

7.4.2 Strengths and limitations

This study contributes to an important evidence gap in women's cardiovascular health. There are currently only a few studies in the literature on the relationship between reproductive factors and different types of individual CVDs, and none which studied CVM specifically. Other strengths of this study include the uniquely large number of cases of a range of CVM outcomes and the long, virtually complete follow-up. Another key strength was the ability to adjust for many different types of covariates including known cardiovascular risk factors and assess their relative importance in the association between reproductive factors and CVM.

There were some limitations specific to each reproductive factor that was investigated in this chapter. The validity of age at menarche recalled in middle-age has been questioned in previous studies. [40-42] The largest study was in 1,050 UK women, which compared age at menarche recorded during medical examination at age 14-15 years with self-reported age at menarche at age 48 years. [42] This study found that only about half of the women recalled the same age at menarche in middle-age as in adolescence, but that there was no systematic under- or over-estimation of age at menarche in middle-age. Therefore, if reporting error in age at menarche was present in this chapter, it is likely to have been non-differential, biasing the overall association with CVM outcomes to the null.

In terms of parity (number of children), self-reported information is likely to be reliable and is therefore unlikely to bias the effect estimates. However, this study did not investigate other pregnancy-related factors which may influence the development of CVDs, such as history of pre-eclampsia, gestational diabetes, or preterm birth. [1] There

was also no information on whether women decided not to have children or were unable to due to health problems. For example, it was not possible to distinguish between women who had never been pregnant and those who had a miscarriage, although there is some evidence in the literature that women who miscarried have an elevated CVD risk. [43]

It is important to consider that breastfeeding initiation and duration can be affected by factors such as BMI, smoking, health issues in pregnancy, employment conditions, and education, which are also associated with CVD risk. [44, 45] Self-reported breastfeeding duration can also be subject to recall error and misclassification, although previous studies have found that self-reported breastfeeding tends to be reliable. [46, 47]

7.4.3 Conclusions

In this large cohort of middle-aged UK women, early menarche and greater parity was associated with only a slightly higher risk of incident CVM, and breastfeeding in parous women was associated with a somewhat lower risk after full adjustment for anthropometric, lifestyle behavioural, socioeconomic, and other reproductive factors. Most of the associations between the three reproductive factors and all CVM outcomes (i.e. CVM, complex CVM, and CVM pairs) were accounted for by multivariable adjustment, and the little association that remained after multivariable adjustment could be explained by residual confounding or mediation. It was therefore unlikely that the selected reproductive factors are strong risk factors for incident CVM. Nevertheless, the shape and magnitude of the associations varied across some CVM pairs, some associations were less well accounted for by socioeconomic and behavioural factors, and potentially important factors such as pregnancy complications were not investigated in this study. Therefore, further research is needed to clarify whether the selected reproductive factors could be substantial risk factors for particular types of CVM.

7.5 References

[1] Okoth K, Chandan JS, Marshall T, Thangaratinam S, Thomas GN, Nirantharakumar K, et al. Association between the reproductive health of young women and cardiovascular disease in later life: umbrella review. BMJ. 2020 Oct;371:m3502. Available from: *https://doi.org/10.1136/bmj.m3502*.

[2] DiVall SA, Radovick S. Pubertal development and menarche. Annals of the New York Academy of Sciences. 2008 Jul;1135:19–28. Available from: *https://doi.org/-10.1196/annals.1429.026*.

[3] Remsberg KE, Demerath EW, Schubert CM, Chumlea WC, Sun SS, Siervogel RM. Early menarche and the development of cardiovascular disease risk factors in adolescent girls: the Fels Longitudinal Study. The Journal of clinical endocrinology and metabolism. 2005 May;90:2718–24. Available from: *https://doi.org/10.1210/jc.2004-1991*.

[4] Frontini MG, Srinivasan SR, Berenson GS. Longitudinal changes in risk variables underlying metabolic Syndrome X from childhood to young adulthood in female subjects with a history of early menarche: the Bogalusa Heart Study. International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity. 2003 Nov;27:1398–404. Available from: *https://doi.org/10.1038/-sj.ijo.0802422*.

[5] Canoy D, Beral V, Balkwill A, Wright FL, Kroll ME, Reeves GK, et al. Age at menarche and risks of coronary heart and other vascular diseases in a large UK cohort. Circulation. 2015 Jan;131(3):237–44. Available from: *https://doi.org/10.1161/-CIRCULATIONAHA.114.010070*.

[6] Gunderson EP, Lewis CE, Wei GS, Whitmer RA, Quesenberry CP, Sidney S. Lactation and changes in maternal metabolic risk factors. Obstetrics and gynecology. 2007 Mar;109:729–38. Available from: *https://doi.org/10.1097/01.aog.0000252831.06695.03*.

[7] Mankuta D, Elami-Suzin M, Elhayani A, Vinker S. Lipid profile in consecutive pregnancies. Lipids in health and disease. 2010 Jun;9:58. Available from: *https://doi.org/-10.1186/1476-511x-9-58*.

[8] Stuebe AM, Rich-Edwards JW. The reset hypothesis: lactation and maternal metabolism. American journal of perinatology. 2009 Jan;26:81–8. Available from: *https://-doi.org/10.1055/s-0028-1103034*.

[9] Li W, Ruan W, Lu Z, Wang D. Parity and risk of maternal cardiovascular disease: A dose-response meta-analysis of cohort studies. European journal of preventive cardiology. 2019 Apr;26:592–602. Available from: *https://doi.org/10.1177/-2047487318818265*.

[10] Nguyen B, Jin K, Ding D. Breastfeeding and maternal cardiovascular risk factors and outcomes: A systematic review. PloS one. 2017 Nov;12:e0187923. Available from: *https://doi.org/10.1371/journal.pone.0187923*.

[11] Prentice P, Viner RM. Pubertal timing and adult obesity and cardiometabolic risk in women and men: a systematic review and meta-analysis. International Journal of

Obesity. 2013 Nov;37(8):1036–1043. Available from: *https://doi.org/10.1038/-ijo.2012.177*.

[12] Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Series B Stat Methodol. 1995;57(1):289–300. Available from: *https://doi.org/10.1111/j.2517-6161.1995.tb02031.x*.

[13] Stringhini S, Sabia S, Shipley M, Brunner E, Nabi H, Kivimäki M, et al. Association of Socioeconomic Position With Health Behaviors and Mortality. JAMA. 2010 Mar;303(12):1159–1166. Available from: *https://doi.org/10.1001/jama.2010.297*.

[14] Petrovic D, de Mestral C, Bochud M, Bartley M, Kivimäki M, Vineis P, et al. The contribution of health behaviors to socioeconomic inequalities in health: A systematic review. Preventive Medicine. 2018 Aug;113:15–31. Available from: *https://doi.org/-10.1016/j.ypmed.2018.05.003*.

[15] Peters SAE, Woodward M. Women's reproductive factors and incident cardiovascular disease in the UK Biobank. Heart. 2018 Jul;104(13):1069. Available from: *https://doi.org/10.1136/heartjnl-2017-312289*.

[16] Yang L, Li L, Millwood IY, Peters SAE, Chen Y, Guo Y, et al. Age at menarche and risk of major cardiovascular diseases: Evidence of birth cohort effects from a prospective study of 300,000 Chinese women. International Journal of Cardiology. 2017 Jan;227:497–502. Available from: *https://doi.org/10.1016/j.ijcard.2016.10.115*.

[17] Plompen MVL, van der Schouw YT, Rutten FH, Verschuren WMM, Boer JMA, Asselbergs FW, et al. Age at menarche and heart failure risk: The EPIC-NL study. Maturitas. 2020 Jan;131:34–39. Available from: *https://doi.org/10.1016/-j.maturitas.2019.10.003*.

[18] Kivimäki M, Lawlor DA, Smith GD, Elovainio M, Jokela M, Keltikangas-Järvinen L, et al. Association of age at menarche with cardiovascular risk factors, vascular structure, and function in adulthood: the Cardiovascular Risk in Young Finns study. The American journal of clinical nutrition. 2008 Jun;87:1876–82. Available from: *https://doi.org/-10.1093/ajcn/87.6.1876*.

[19] Zheng Y, Wen TS, Shen Y, Hu H. Age at menarche and cardiovascular health: results from the NHANES 1999-2016. Menopause. 2021 Jan;28(1):18–24. Available from: *https://doi.org/10.1097/gme.0000000000001653*.

[20] Elks CE, Perry JRB, Sulem P, Chasman DI, Franceschini N, He C, et al. Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. Nature genetics. 2010 Dec;42:1077–85. Available from: *https://doi.org/10.1038/-ng.714*.

[21] Fernández-Rhodes L, Demerath EW, Cousminer DL, Tao R, Dreyfus JG, Esko T, et al. Association of adiposity genetic variants with menarche timing in 92,105 women of European descent. American journal of epidemiology. 2013 Aug;178:451–60. Available from: *https://doi.org/10.1093/aje/kws473*.

[22] Cao M, Cui B. Negative Effects of Age at Menarche on Risk of Cardiometabolic Diseases in Adulthood: A Mendelian Randomization Study. The Journal of clinical endocrinology and metabolism. 2020 Feb;105:515–522. Available from: *https://doi.org/-10.1210/clinem/dgz071*.

[23] Durazo EM, Cabeza de Baca T, Slopen N, Parikh NI, Buring JE, Glynn RJ, et al. Parity, Job Strain, and Cardiovascular Risk in the Women's Health Study. Current Cardiovascular Risk Reports. 2018 Mar;12(3):8. Available from: *https://doi.org/10.1007/-s12170-018-0571-z*.

[24] Parikh NI, Cnattingius S, Dickman PW, Mittleman MA, Ludvigsson JF, Ingelsson E. Parity and risk of later-life maternal cardiovascular disease. American Heart Journal. 2010 Feb;159(2):215–221e6. Available from: *https://doi.org/10.1016/j.ahj.2009.11.017*.

[25] Wong JA, Rexrode KM, Sandhu RK, Conen D, Albert CM. Number of Pregnancies and Atrial Fibrillation Risk: The Women's Health Study. Circulation. 2017 Feb;135(28154001):622–624. Available from: *https://doi.org/10.1161/-circulationaha.116.026629*.

[26] Oliver-Williams C, Vladutiu CJ, Loehr LR, Rosamond WD, Stuebe AM. The Association Between Parity and Subsequent Cardiovascular Disease in Women: The Atherosclerosis Risk in Communities Study. Journal of women's health. 2019 May;28(30481103):721–727. Available from: *https://dx.doi.org/10.1089/jwh.2018.7161*.

[27] Klingberg S, Brekke HK, Winkvist A, Engström G, Hedblad B, Drake I. Parity, weight change, and maternal risk of cardiovascular events. American Journal of Obstetrics & Gynecology. 2017 Feb;216(2):172e1–172e15. Available from: *https://doi.org/10.1016/-j.ajog.2016.09.105*.

[28] Germain AM, Romanik MC, Guerra I, Solari S, Reyes MS, Johnson RJ, et al. Endothelial dysfunction: a link among preeclampsia, recurrent pregnancy loss, and future cardiovascular events? Hypertension (Dallas, Tex : 1979). 2007 Jan;49:90–5. Available from: *https://doi.org/10.1161/01.hyp.0000251522.18094.d4*.

[29] Barbour LA, McCurdy CE, Hernandez TL, Kirwan JP, Catalano PM, Friedman JE. Cellular mechanisms for insulin resistance in normal pregnancy and gestational diabetes. Diabetes care. 2007 Jul;30 Suppl 2:S112–9. Available from: *https://doi.org/10.2337/dc07-s202*.

[30] Lawlor DA, Emberson JR, Ebrahim S, Whincup PH, Wannamethee SG, Walker M, et al. Is the Association Between Parity and Coronary Heart Disease Due to Biological Effects of Pregnancy or Adverse Lifestyle Risk Factors Associated With Child-Rearing? Circulation. 2003 Mar;107(9):1260–1264. Available from: *https://doi.org/10.1161/-01.CIR.0000053441.43495.1A*.

[31] Natland ST, Nilsen TIL, Midthjell K, Andersen LF, Forsmo S. Lactation and cardiovascular risk factors in mothers in a population-based study: the HUNT-study. International breastfeeding journal. 2012 Jun;7:8. Available from: *https://doi.org/10.1186/-1746-4358-7-8*.

[32] Kirkegaard H, Bliddal M, Støvring H, Rasmussen KM, Gunderson EP, Køber L, et al. Breastfeeding and later maternal risk of hypertension and cardiovascular disease - The role of overall and abdominal obesity. Preventive medicine. 2018 Sep;114:140–148. Available from: *https://doi.org/10.1016/j.ypmed.2018.06.014*.

[33] Jacobson LT, Hade EM, Collins TC, Margolis KL, Waring ME, Van Horn LV, et al. Breastfeeding History and Risk of Stroke Among Parous Postmenopausal Women in the Women's Health Initiative. Journal of the American Heart Association. 2018 Sep;7:e008739. Available from: *https://dx.doi.org/10.1161/JAHA.118.008739*.

[34] Nguyen B, Gale J, Nassar N, Bauman A, Joshy G, Ding D. Breastfeeding and Cardiovascular Disease Hospitalization and Mortality in Parous Women: Evidence From a Large Australian Cohort Study. Journal of the American Heart Association. 2019 Mar;8:e011056. Available from: *https://doi.org/10.1161/jaha.118.011056*.

[35] Peters SAE, Yang L, Guo Y, Chen Y, Bian Z, Du J, et al. Breastfeeding and the Risk of Maternal Cardiovascular Disease: A Prospective Study of 300,000 Chinese Women. Journal of the American Heart Association. 2017 Jun;6:e006081. Available from: *https://doi.org/10.1161/jaha.117.006081*.

[36] Pathirana MM, Ali A, Lassi ZS, Arstall MA, Roberts CT, Andraweera PH. Protective Influence of Breastfeeding on Cardiovascular Risk Factors in Women With Previous Gestational Diabetes Mellitus and Their Children: A Systematic Review and Meta-Analysis. Journal of human lactation. 2021 Oct:8903344211034779. Available from: *https://doi.org/10.1177/08903344211034779*.

[37] Rameez RM, Sadana D, Kaur S, Ahmed T, Patel J, Khan MS, et al. Association of Maternal Lactation With Diabetes and Hypertension: A Systematic Review and Meta-analysis. JAMA network open. 2019 Oct;2:e1913401. Available from: *https://doi.org/-10.1001/jamanetworkopen.2019.13401*.

[38] Perrine CG, Nelson JM, Corbelli J, Scanlon KS. Lactation and Maternal Cardio-Metabolic Health. Annual review of nutrition. 2016 Jul;36:627–45. Available from: *https://doi.org/10.1146/annurev-nutr-071715-051213*.

[39] Peters SA, van der Schouw YT, Wood AM, Sweeting MJ, Moons KG, Weiderpass E, et al. Parity, breastfeeding and risk of coronary heart disease: A pan-European case-cohort study. European journal of preventive cardiology. 2016 Nov;23(27378766):1755–1765. Available from: *https://doi.org/10.1177/2047487316658571*.

[40] Koprowski C, Coates RJ, Bernstein L. Ability of young women to recall past body size and age at menarche. Obesity research. 2001 Aug;9:478–85. Available from: *https://-doi.org/10.1038/oby.2001.62*.

[41] Must A, Phillips SM, Naumova EN, Blum M, Harris S, Dawson-Hughes B, et al. Recall of early menstrual history and menarcheal body size: after 30 years, how well do women remember? American journal of epidemiology. 2002 Apr;155:672–9. Available from: *https://doi.org/10.1093/aje/155.7.672*.

[42] Cooper R, Blell M, Hardy R, Black S, Pollard TM, Wadsworth MEJ, et al. Validity of age at menarche self-reported in adulthood. J Epidemiol Community Health. 2006 Nov;60(11):993. Available from: *https://doi.org/10.1136/jech.2005.043182*.

[43] Oliver-Williams CT, Heydon EE, Smith GCS, Wood AM. Miscarriage and future maternal cardiovascular disease: a systematic review and meta-analysis. Heart. 2013 Nov;99(22):1636. Available from: *https://doi.org/10.1136/heartjnl-2012-303237*.

[44] Cohen SS, Alexander DD, Krebs NF, Young BE, Cabana MD, Erdmann P, et al. Factors Associated with Breastfeeding Initiation and Continuation: A Meta-Analysis. The Journal of Pediatrics. 2018 Dec;203:190–196e21. Available from: *https://doi.org/-10.1016/j.jpeds.2018.08.008*.

[45] Amir LH, Donath S. A systematic review of maternal obesity and breastfeeding intention, initiation and duration. BMC Pregnancy and Childbirth. 2007 Jul;7(1):9. Available from: *https://doi.org/10.1186/1471-2393-7-9*.

[46] Natland ST, Andersen LF, Nilsen TIL, Forsmo S, Jacobsen GW. Maternal recall of breastfeeding duration twenty years after delivery. BMC medical research methodology. 2012 Nov;12:179. Available from: *https://doi.org/10.1186/1471-2288-12-179*.

[47] Promislow JHE, Gladen BC, Sandler DP. Maternal recall of breastfeeding duration by elderly women. American journal of epidemiology. 2005 Feb;161:289–96. Available from: *https://doi.org/10.1093/aje/kwi044*.

8 Summary and implications

8.1 Introduction

Multimorbidity is commonly defined as the existence of two or more chronic conditions in one person. [1] Multimorbidity increases with older age [2] and lower socioeconomic status [3, 4], and some studies found that it is more common in women than in men [4, 5]. Multimorbidity is a serious public health concern in the UK due to its high prevalence and association with negative outcomes such as higher mortality, lower quality of life, and greater use of health services (see Chapter 1).

Cardiovascular disease (CVD) has consistently been identified as one of the most common components of multimorbidity. [5-7] Evidence suggests that co-occurring CVDs (or cardiovascular multimorbidity [CVM] as defined in this thesis) may have an important role in multimorbidity onset and progression in older adults. [8-11] However, few prospective studies have investigated the risk factors for CVM (see Chapter 2). Furthermore, consistent with the overall underrepresentation of women in studies of CVD, there is little research on CVM specific to women.

The aim of this thesis was to conduct a broad overview of a range of possible risk factors for CVM in the Million Women Study. Briefly, the Million Women Study was a prospective cohort study of 1.3 million UK women aged 50 to 64 years, recruited through breast screening centres in 1996-2001 and followed-up for hospital admissions, cancer registrations, and deaths via linkage to NHS databases (see Chapter 3). This thesis selected 19 different CVDs (see Table 3.3 of Chapter 3) to comprise CVM, and investigated CVM outcomes in three main ways: CVM (≥2 CVDs), complex CVM (≥4 CVDs), and common CVM pairs (specific pairs of CVDs) in this cohort.

In the context of the low number and narrow scope of previous studies on risk factors for CVM, the investigation of the prospective associations between known cardiovascular risk

factors, socioeconomic status, and reproductive factors with the different CVM outcomes investigated in this thesis is a novel contribution to the current evidence base in the scientific literature.

8.2 Summary of main results

8.2.1 Descriptive study of cardiovascular multimorbidity (Chapter 4)

Among 1.2 million UK women with no CVD or cancer at recruitment (mean age 57 years) followed-up for 16 years on average, 9% (n=114,684) developed CVM (≥2 CVD subtypes) and 2% (n=26,896) developed complex CVM (≥4 CVD subtypes), as defined in this thesis. The probability of having CVM or complex CVM approximately doubled every five years between age 60 and 80 years.

By the end of the study period, CVM was recorded in 42% of participants who had any incident CVD during follow-up. The occurrence of CVM varied across different CVD subtypes. For example, while only about 50% of women who had a record of venous thromboembolism during follow-up also had a record of another incident CVD, around 90% of women with a record of rheumatic heart disease also had a record of at least one other CVD. The most common components of CVM were ischaemic heart disease, atrial fibrillation, other arrhythmias, and heart failure, which were each found in around a fourth to half of all women with CVM. These CVDs were the most common individual CVD subtypes in the cohort and were commonly found with one another.

CVM pairs were selected to investigate specific CVM outcomes at a higher resolution, based on the high number of cases in the cohort (for statistical power) and the distinctiveness of the component CVDs from each other. The four most common CVM pairs were, in descending order of frequency: ‘Ischaemic heart disease & Atrial fibrillation’ (IHD-AF; n=26,239), ‘Ischaemic heart disease & Heart failure’ (IHD-HF;

n=20,614), 'Ischaemic heart disease & Other Arrhythmias' (IHD-ARR; n=18,157), and 'Atrial fibrillation & Heart failure' (AF-HF; n=18,042), and these and 'Other arrhythmias & Heart failure' (ARR-HF; n=9,756), and 'Atrial fibrillation & Stroke' (AF-STR; n=9,562) were investigated as additional CVM outcomes. Although the pairs of 'Atrial fibrillation & Other arrhythmias' (n=15,622) and 'Stroke & Other cerebrovascular disease' (n=13,154) were more common than ARR-HF or AF-STR, they were not studied because the component conditions were not sufficiently distinct from each other.

8.2.2 Established cardiovascular risk factors and cardiovascular multimorbidity (Chapter 5)

After multivariable adjustment for potential confounders, severe obesity (BMI of ≥35 kg/m^2 vs. 20-24 kg/m^2) and smoking ≥15 cigarettes a day (vs. never smoking) were each associated with approximately 2.5 times the risk of CVM and 3 times the risk of complex CVM. For the six selected CVM pairs, the multivariable adjusted effect estimates associated with severe obesity ranged from 2.3 times the risk of IHD-ARR to 5.3 times the risk of AF-HF, while those associated with smoking ≥15 cigarettes a day ranged from 1.9 times the risk of AF-STR to 3.3 times the risk of IHD-HF. There was no evidence of a substantial association between any level of alcohol consumption or physical activity with any of the CVM outcomes after multivariable adjustment.

Self-reported treatment for diabetes (vs. not being treated) was generally associated with 3-4 times the risk of each CVM outcome after multivariable adjustment. Self-reported treatment for hypertension (vs. not being treated) was associated with about twice the risk of each CVM outcome, while self-reported treatment for high blood cholesterol (vs. not being treated) was associated with a 50-70% higher risk of each CVM outcome, all after multivariable adjustment.

The tallest women (≥170 cm) had a 30-40% higher risk of CVM and complex CVM compared to the shortest women (<155 cm) after multivariable adjustment. Taller height was particularly strongly associated with the risk of CVM pairs including atrial fibrillation (60-90% higher risk in the tallest vs. shortest women). Leaving out atrial fibrillation from the definition of CVM markedly attenuated the association between height and CVM, indicating that atrial fibrillation is likely to play an important role in this relationship. For all other risk factor associations, leaving out individual CVDs did not materially change the results.

8.2.3 Socioeconomic status and cardiovascular multimorbidity (Chapter 6)

When stratified by age and recruitment region (i.e. minimal adjustment), women who did not finish compulsory schooling (lowest level) had a 1.9 times higher risk of CVM, 2.3 times the risk of complex CVM, and 1.7-3 times the risk of each selected CVM pair compared to women with tertiary educational qualifications (highest level). Similarly, the minimally adjusted relative risks showed that women living in the most deprived areas had approximately 1.8 times the risk of CVM, 2.0 times the risk of complex CVM, and 1.6-2.4 times the risk of the selected CVM pairs compared to those living in the least deprived areas. After multivariable adjustment, however, the lowest versus the highest level of both SES factors was associated with only 1.2-1.6 times the risk of each CVM outcome.

There was evidence that most of the observed social gradient in CVM risk was mediated or otherwise accounted for by known cardiovascular risk factors. Adjusting for smoking behaviour alone led to 40-50% attenuation in the LR statistics of the associations between SES and each CVM outcome; smoking appeared to be the most important mediator/confounder of the SES-CVM relationships. Simultaneously adjusting for smoking, alcohol consumption, physical activity, and BMI resulted in 70-80% LR

attenuation for each CVM outcome. Since the LR attenuation was substantial and because the adjusted covariates were measured with error, it was possible that the association which remained after adjustment may be further accounted for by residual confounding. Therefore, a combination of mediation via known cardiovascular risk factors and residual confounding may explain most, if not all, of the observed social gradients in the risk of each CVM outcome.

8.2.4 Reproductive factors and cardiovascular multimorbidity (Chapter 7)

When stratifying by age and region of recruitment, early menarche (<12 years vs. 13 years), high parity (four or more children vs. none), and never having breastfed their children (vs. ever) were each associated with a somewhat (~30%) higher risk of CVM. However, as it did for socioeconomic status, simultaneous adjustment for known behavioural cardiovascular risk factors (i.e. smoking, BMI, alcohol consumption, and physical activity) accounted for 70-80% of the associations with CVM, and only about a 10% higher risk of CVM remained for each reproductive factor after multivariable adjustment. Therefore, there was little evidence of a substantial association between these reproductive factors and CVM, and it is plausible that residual confounding could account for any remaining association The results were similar for complex CVM and the CVM pairs. Although the shape of the association was different for some CVM pairs in relation to parity, the attenuation in the LR statistics again indicated that most of the associations were accounted for by mediation and/or confounding via known cardiovascular risk factors.

8.3 Summary of thesis strengths and limitations

8.3.1 Key strengths

One of the main strengths of this thesis was the inclusion of a comprehensive list of CVDs in the definition of CVM, rather than restricting to a few common CVDs like coronary heart disease, stroke, or heart failure. Therefore, this thesis had a more detailed examination of the various subtypes of CVDs that co-occur to comprise CVM than previous studies. Moreover, the large sample size (at least an order of magnitude larger than any of the other studies included in the systematic review in Chapter 2), the large number of cases of not only individual CVDs, but also a range of different types of CVM (e.g. complex CVM, CVM pairs), and the long and substantially complete follow-up (<1% of participants lost to follow-up over mean 16 years) allowed enough statistical power to reliably estimate the associations between risk factors and the different CVM outcomes. In addition, the availability of information on a wide range of anthropometric variables, health-related behaviours, and reproductive factors allowed for their investigation as risk factors as well as inclusion as covariates in the regression models.

Furthermore, the Million Women Study was able to recruit one in four women in the target age range at the time of study recruitment and findings are more likely to be generalisable to the target population. The method of sampling through breast screening centres meant that the cohort was restricted to women who attended breast screening. Women who attended breast screening were, on average, living in more affluent areas, ethnically white, and more likely to be receiving hormone replacement therapy than those who did not attend screening. [12] Nonetheless, they were comparable in age and recent prescriptions for different medications for chronic conditions (e.g. heart disease, diabetes) to women who did not attend breast screening. There was, therefore, limited evidence that the differences between the participants of the Million Women Study and the general

population of women would have resulted in major disparities in the risk factor associations observed in this thesis. However, the results of this thesis may not be generalisable to the subpopulations that were underrepresented in the study cohort (e.g. ethnic minorities, the most socioeconomically deprived).

8.3.2 Limitations

There were also important limitations to this thesis. First, information on risk factors and any potential confounders were based on self-reported data at a single point in time (i.e. study baseline). Although validation studies (including some within the Million Women Study) showed that self-reported measures were generally reliable and reproducible [13-22], self-reports are generally less accurate than direct measurement and sometimes susceptible to differential reporting (e.g. self-reporting lower weight due to social desirability), which could in turn bias the direction of the effect estimates and reduce study precision. Although some risk factors studied in this thesis were not possible to measure directly (e.g. smoking), other risk factors were potentially more susceptible to error or bias due to self-report (e.g. weight). Additionally, self-reported treatment for diabetes, hypertension, and high blood cholesterol, were used as proxies for having these conditions. Although using linked primary care records or direct measurement would have been more reliable sources of information for these risk factors, these were not available for this thesis.

This thesis only used risk factor data from a single timepoint (i.e. at recruitment), which may be an issue if there had been reporting error or changes in behaviour (and therefore risk factor status) over time. In fact, studies have shown that over the follow-up period of the Million Women Study for this thesis (i.e. 1996-2017) population-level secular trends in cardiovascular risk factors have shifted in the UK. For example, cigarette consumption decreased in England by 25% between 2011-2018 [23], there was a significant reduction in

the proportion of people who reported smoking to their GP between 2001-2007 [24], and the prevalence of obesity increased by 25% between 1991-2010 [25]. Several validation studies within the Million Women Study have addressed the potential for regression dilution bias for some risk factors (weight, height, smoking, and physical activity), and the implications for the main results of this thesis have been discussed in Chapter 5. [19, 22, 26, 27] Briefly, misclassification or changes in height and BMI were unlikely to have substantially biased the effect estimates, whereas differences in smoking behaviour and physical activity were likely to have resulted in attenuated effect estimates in this thesis.

Another limitation was the lack of information on some potential confounders in the baseline questionnaire. For instance, there was no information on dietary patterns, mental health, or quality of care of the participants; these may represent aspects of the participants' health that are not captured by other measured confounders. Some related issues were unknown confounders and residual confounding by factors which had ostensibly been adjusted for, but which were measured with error. When a confounder is measured with error, adjustment will not completely account for confounding by that factor. Well-known confounders such as age, BMI, and smoking behaviour as well as educational attainment and area deprivation as measures of socioeconomic status were adjusted for in this thesis, and the use of LR attenuation analysis helped quantify the scope for measurement error and residual confounding to account for the observed associations.

As with other observational research, this study was also prone to reverse causation bias, where ill health affects risk factor status rather than risk factors causing ill health. Even risk factors such as education can be prone to reverse causation: ill health in childhood may affect educational attainment and can also be linked to heightened disease risk later in life. [28] Furthermore, it is difficult to estimate the effect that potential reverse causation bias may have on the risk factor associations, because ill health may have varying effects

even on a single risk factor. For example, ill health can lead to weight loss (and thus, lower BMI) but sometimes can also lead to weight gain. Ill health can also lead to an inability to exercise, while some diagnoses may lead to more motivation to exercise and keep healthy. In this thesis, sensitivity analyses excluding women who reported fair or poor health and excluding the first 3 years of follow-up indicated that there may be some reverse causation bias in the effect estimates for behavioural risk factors such as physical activity (Chapter 5).

Finally, the completeness of CVM outcomes was an important concern. The only information sources for the CVDs that comprised CVM in this study were day-case and inpatient hospital admissions and death registrations, which did not include any information from general practice or outpatient, emergency, non-ambulatory, or long-term care. Therefore, this thesis will not have included cases of CVDs that are less serious and do not require inpatient hospital care. Indeed, a study in the Million Women Study which compared the validity of vascular outcomes (i.e. ischaemic heart disease, cerebrovascular disease, and venous thromboembolism) in hospital admissions data compared to general practice records showed that among a random sample of 864 women with no hospital admission record for any vascular disease, 3% (n=26) had a vascular disease (of the types listed above). [29] Among the 3%, the conditions tended to be of less serious nature: 18 women had ischaemic heart disease other than myocardial infarction, three had transient ischaemic attack, and three had venous thrombosis without embolism. This can mean that a substantial number of cases of less serious vascular conditions are missed by hospital admissions and death records. With additional information sources on CVDs, it would be possible to have a more complete picture of the prevalence, incidence, and distribution of CVM and its component conditions.

8.4 Implications for clinical practice

While the findings of this thesis will mainly inform understanding of risk factors for CVM and future research in CVM epidemiology, there are some potential implications for clinical practice. Although it was understood that patients with a CVD diagnosis are likely to develop other CVDs before this thesis was undertaken, the co-occurrence of CVD subtypes had not been quantified across as broad a range of diagnoses as examined here. This thesis found that more than 40% of women who have a diagnosis of any incident CVD in middle to older age developed at least one other CVD diagnosis. This highlights the need for physicians to be vigilant about secondary CVD prevention in female patients – especially since women are understudied and underdiagnosed with CVD compared to men, and physicians often underestimate the risk of CVD in their female patients. [30-32]

The co-occurrence of CVDs has been addressed to some extent in current UK National Institute for Health and Care Excellence (NICE) clinical guidelines (published in 2021) albeit inconsistently for different conditions. For example, the NICE guidelines for atrial fibrillation recommend discussing anticoagulation therapy with patients with atrial fibrillation for stroke prevention – recognising stroke as a common and important comorbid condition – and have done so since the first guidelines were published in 2006. [33] (In the Million Women Study, 13% of women with atrial fibrillation also had a stroke.) Many other CVD subtypes were also commonly found with each other in this thesis. For instance, 60% of women with aortic stenosis and 50% of women with other aortic or mitral valve disorders also had a record of ischaemic heart disease. However, the 2021 guidelines for heart valve disease did not recommend lipid lowering therapy for patients with heart valve disease. [34] It is possible that many people with any type of heart disease could benefit from lipid lowering therapy for prevention of further CVD diagnoses. Therefore, the benefits of recommending lipid lowering therapy for women

with any CVD warrants further investigation, including establishing the temporal order of disease occurrence.

Finally, the findings from this thesis lends further support for the modification of established cardiovascular risk factors in the context of multimorbidity. Such risk factors include smoking cessation, weight loss, and pharmacological control of blood pressure, lipid levels, and glycaemic levels, as recommended by the 2021 European Society of Cardiology (ESC) Guidelines on cardiovascular disease prevention in clinical practice. [35] These interventions are likely to be beneficial not only for the prevention of individual incident CVDs, but also for the delay in the development of any subsequent CVD diagnoses and the development of CVM.

8.5 Future research

8.5.1 Additional analyses in the Million Women Study

There is potential for further research using data in the Million Women Study which were not included in this thesis (see Chapter 3). Resurvey questionnaires were posted to all living participants approximately 3, 8, 12, 15, and 18 years after recruitment, although there was some loss to follow-up (e.g. the 3-year questionnaire was completed by 67% of participants at recruitment). The resurveys have information on additional risk factors and potential confounders that were not reported in the recruitment questionnaire. For instance, detailed information on menopausal status, age at menopause, menopausal hormone therapy, and oral contraceptive usage could be used to better control for these factors as confounders and investigate them as potential risk factor in future studies.

Moreover, some risk factors investigated in this thesis have been included in the subsequent resurvey questionnaires. Based on findings of previous validation studies in the Million Women Study, real changes over time and/or reporting errors in BMI were

unlikely to have strongly biased the effect estimates [22], while differences in smoking behaviour [26] and physical activity [19] suggested that the effect estimates reported in this thesis are likely to be underestimates of the true associations (see Chapter 5, Section 5.4.2, page 171). In future studies using repeat measures from multiple questionnaires, it would be possible to assess the validity of risk factors over time and account for time-varying covariates in the regression models.

Linkage of primary care datasets to the Million Women Study could identify cases of CVM that may have not been captured in hospital admission or death records. Previously, primary care diagnoses and prescribing data on 100,000 women in England from the CPRD (available up to 2013) and prescribing data on 100,000 women in Scotland from the ISD (available 2009-2015) had been linked to the Million Women Study. With updated linkages with primary care records, it would be possible to assess whether identifying additional cases of CVM results in different risk factor associations to those observed in this thesis.

Finally, it would be possible to explore risk factors associations for other specific definitions of CVM – such as trios of co-occurring CVD subtypes – to move another step toward understanding how CVDs cluster together, taking advantage of the large number of incident cases of CVD subtypes in the Million Women Study.

8.5.2 Differences by sex

Future studies could be conducted in cohorts that include men with a broader age range, particularly since men tend to develop CVD about a decade earlier than women on average. [36] Although it was not possible in this thesis to investigate CVM in men, both the prevalence [36-38] and strength and magnitude of associations between risk factors

and CVDs [39-42] have been reported to differ between men and women. It is also likely, therefore, that the prevalence and risk factors for CVM differ by sex.

8.5.3 Different definitions of multimorbidity

Although this thesis has defined CVM using a simple count of CVDs, other definitions could be used to gain a more detailed understanding of risk factors for CVM. For example, previous studies have conceptualised multimorbidity development as transitions from healthy state to first disease and then to multimorbidity, and investigated how risk factors affect each transition differently. [43, 44] These studies found that different types of risk factors contributed to different transitions. A study in the British Whitehall II study found that while hypertension, high cholesterol, obesity, and family history of diabetes were key predictors of the transition from no disease to first cardiometabolic disease, physical activity, alcohol consumption, and smoking were then the determinants of the transition from first disease to multimorbidity. This type of analysis could be applied to help inform the prioritisation of risk factors to target for the secondary prevention of CVD (in other words, primary prevention of CVM).

8.5.4 Analysis of causal association via Mendelian randomisation

Although this thesis is a very large prospective study of robust design, it is still affected by the limitations of observational research including residual confounding and reverse causation. I have considered such limitations in each chapter, but it was not always possible to assess how much these issues may have biased the effect estimates and in what direction. Therefore, conclusions of causality cannot be made based on the associations observed in this thesis. One way to address this limitation is to conduct Mendelian randomisation (MR) analyses, which use genetic variants assigned at birth as a "natural experiment" to assess the potential causal relationship between modifiable risk factors and health outcomes. MR studies are believed to be less susceptible to confounding and

reverse causation bias than traditional observational studies, although they are not immune to bias; the validity of MR studies depends on several assumptions about the relationships between genetic variants, risk factors, the outcome, and other factors, which must be met. [45]

Previous MR studies have investigated the risk factors studied in this thesis (e.g. BMI, smoking, alcohol consumption, height, age at menarche) and their relationship with common individual CVDs (e.g. coronary heart disease, stroke, heart failure, atrial fibrillation). [46-54] For instance, there was evidence of a causal association between being taller and atrial fibrillation, but not with other CVDs [52, 53], which was consistent with the findings of this thesis. It is important to explore these relationships further, and the results of MR studies must be interpreted in the context of good quality observational evidence and randomised controlled trials where feasible. Although genetic data is only available for a small proportion (~5%) of participants in the Million Women Study, meaning it is not ideal for conducting MR analyses, there are existing datasets such as the UK Biobank which are suitable for MR studies and can offer opportunities to examine the potential causality of the observational associations reported in this thesis.

8.6 References

[1] Ho ISS, Azcoaga-Lorenzo A, Akbari A, Black C, Davies J, Hodgins P, et al. Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies. The Lancet Public Health. 2021 Aug;6(8):e587–e597. Available from: *https://doi.org/10.1016/S2468-2667(21)00107-9*.

[2] Ho ISS, Azcoaga-Lorenzo A, Akbari A, Davies J, Hodgins P, Khunti K, et al. Variation in the estimated prevalence of multimorbidity: systematic review and meta-analysis of 193 international studies. BMJ Open. 2022 Apr;12(4):e057017. Available from: *https://doi.org/10.1136/bmjopen-2021-057017*.

[3] Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet. 2012 Jul;380(9836):37–43. Available from: *https://doi.org/-10.1016/S0140-6736(12)60240-2*.

[4] Cassell A, Edwards D, Harshfield A, Rhodes K, Brimicombe J, Payne R, et al. The epidemiology of multimorbidity in primary care: a retrospective cohort study. Br J Gen Pract. 2018 Apr;68(669):e245. Available from: *http://bjgp.org/content/68/669/-e245.abstract*.

[5] Violan C, Foguet-Boreu Q, Flores-Mateo G, Salisbury C, Blom J, Freitag M, et al. Prevalence, Determinants and Patterns of Multimorbidity in Primary Care: A Systematic Review of Observational Studies. PLOS ONE. 2014 Jul;9(7):e102149. Available from: *https://doi.org/10.1371/journal.pone.0102149*.

[6] Weiss CO, Boyd CM, Yu Q, Wolff JL, Leff B. Patterns of Prevalent Major Chronic Disease Among Older Adults in the United States. JAMA. 2007 Sep;298(10):1158–1162. Available from: *https://doi.org/10.1001/jama.298.10.1160-b*.

[7] Rocca WA, Boyd CM, Grossardt BR, Bobo WV, Finney Rutten LJ, Roger VL, et al. Prevalence of multimorbidity in a geographically defined American population: patterns by age, sex, and race/ethnicity. Mayo Clinic proceedings. 2014 Oct;89:1336–49. Available from: *https://doi.org/10.1016/j.mayocp.2014.07.010*.

[8] Kudesia P, Salimarouny B, Stanley M, Fortin M, Stewart M, Terry A, et al. The incidence of multimorbidity and patterns in accumulation of chronic conditions: A systematic review. Journal of Multimorbidity and Comorbidity. 2021 Jan;11:26335565211032880. Available from: *https://doi.org/10.1177/-26335565211032880*.

[9] Marian AJ, Bhatnagar A, Bolli R, Izpisua Belmonte JC. Introduction to Cardiovascular Aging Compendium. Circulation Research. 2018 Sep;123(7):737–739. Available from: *https://doi.org/10.1161/CIRCRESAHA.118.313940*.

[10] Zhu Y, Edwards D, Mant J, Payne RA, Kiddle S. Characteristics, service use and mortality of clusters of multimorbid patients in England: a population-based study. BMC Medicine. 2020 Apr;18(1):78. Available from: *https://doi.org/10.1186/s12916-020-01543-8*.

[11] The Emerging Risk Factors Collaboration. Association of Cardiometabolic Multimorbidity With Mortality. JAMA. 2015 Jul;314(1):52–60. Available from: *https://-doi.org/10.1001/jama.2015.7008*.

[12] Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, et al. Comparison of various characteristics of women who do and do not attend for breast cancer screening. Breast Cancer Res. 2001 Dec;4:R1. Available from: *https://doi.org/10.1186/bcr418*.

[13] Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ. 2004 Jun;328:1519. Available from: *https://doi.org/10.1136/bmj.38142.554479.ae*.

[14] Patrick DL, Cheadle A, Thompson DC, Diehr P, Koepsell T, Kinne S. The validity of self-reported smoking: a review and meta-analysis. Am J Public Health. 1994 Jul;84(7):1086–1093. Available from: *https://doi.org/10.2105/ajph.84.7.1086*.

[15] Brigham J, Lessov-Schlaggar CN, Javitz HS, Krasnow RE, Tildesley E, Andrews J, et al. Validity of recall of tobacco use in two prospective cohorts. Am J Epidemiol. 2010 Oct;172(7):828–835. Available from: *https://dx.doi.org/10.1093/aje/kwq179*.

[16] Wareham NJ, Jakes RW, Rennie KL, Schuit J, Mitchell J, Hennings S, et al. Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. Public Health Nutr. 2003 Jun;6(4):407–413. Available from: *https://-doi.org/10.1079/phn2002439*.

[17] Bohlscheid-Thomas S, Hoting I, Boeing H, Wahrendorf J. Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German part of the EPIC project. European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol. 1997 Jan;26(suppl_1):S59–S70. Available from: *https://-doi.org/10.1093/ije/26.suppl_1.s59*.

[18] Ocke MC, Bueno-de Mesquita HB, Pols MA, Smit HA, van Staveren WA, Kromhout D. The Dutch EPIC food frequency questionnaire. II. Relative validity and reproducibility for nutrients. Int J Epidemiol. 1997 Jan;26(suppl_1):S49–S58. Available from: *https://doi.org/10.1093/ije/26.suppl_1.s49*.

[19] Armstrong MEG, Cairns BJ, Green J, Reeves GK, Beral V, the Million Women Study Collaborators. Reported frequency of physical activity in a large epidemiological study: relationship to specific activities and repeatability over time. BMC Med Res Methodol. 2011 Jun;11:97. Available from: *https://doi.org/10.1186/1471-2288-11-97*.

[20] Roddam AW, Spencer E, Banks E, Beral V, Reeves G, Appleby P, et al. Reproducibility of a short semi-quantitative food group questionnaire and its performance in estimating nutrient intake compared with a 7-day diet diary in the Million Women Study. Public Health Nutr. 2005 Apr;8(2):201–213. Available from: *https://doi.org/-10.1079/phn2004676*.

[21] Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, et al. Agreement between general practice prescription data and self-reported use of hormone replacement therapy and treatment for various illnesses. J Epidemiol Biostat. 2001 Feb;6:357–363. Available from: *https://doi.org/10.1080/13595220152601837*.

[22] Wright FL, Green J, Reeves G, Beral V, Cairns BJ, on behalf of the Million Women Study collaborators. Validity over time of self-reported anthropometric variables during follow-up of a large cohort of UK women. BMC Med Res Methodol. 2015 Oct;15:81. Available from: *https://doi.org/10.1186/s12874-015-0075-1*.

[23] Jackson SE, Beard E, Kujawski B, Sunyer E, Michie S, Shahab L, et al. Comparison of Trends in Self-reported Cigarette Consumption and Sales in England, 2011 to 2018. JAMA network open. 2019 Aug;2(31461148):e1910161–e1910161. Available from: *https://dx.doi.org/10.1001/jamanetworkopen.2019.10161*.

[24] Simpson CR, Hippisley-Cox J, Sheikh A. Trends in the epidemiology of smoking recorded in UK general practice. The British journal of general practice : the journal of the Royal College of General Practitioners. 2010 Mar;60(20202355):e121–e127. Available from: *https://doi.org/10.3399/bjgp10X483544*.

[25] Agha M, Agha R. The rising prevalence of obesity: part A: impact on public health. International journal of surgery Oncology. 2017 Aug;2(29177227):e17–e17. Available from: *https://dx.doi.org/10.1097/IJ9.0000000000000017*.

[26] Pirie K, Peto R, Reeves GK, Green J, Beral V. The 21st century hazards of smoking and benefits of stopping: a prospective study of one million women in the UK. The Lancet. 2013 Jan;381(9861):133–141. Available from: *https://doi.org/10.1016/S0140-6736(12)61720-6*.

[27] Cairns BJ, Liu B, Clennell S, Cooper R, Reeves GK, Beral V, et al. Lifetime body size and reproductive factors: comparisons of data recorded prospectively with self reports in middle age. BMC Medical Research Methodology. 2011 Jan;11(1):7. Available from: *https://doi.org/10.1186/1471-2288-11-7*.

[28] Berkman LF, Kawachi I, Glymour MM, editors. Social Epidemiology. 2nd ed. Oxford University Press; 2014. 17-62 p.

[29] Wright FL, Green J, Canoy D, Cairns BJ, Balkwill A, Beral V, et al. Vascular disease in women: comparison of diagnoses in hospital episode statistics and general practice records in England. BMC Med Res Methodol. 2012 Oct;12:161. Available from: *https://doi.org/10.1186/1471-2288-12-161*.

[30] Ravindran TKS, Teerawattananon Y, Tannenbaum C, Vijayasingham L. Making pharmaceutical research and regulation work for women. BMJ. 2020 Oct;371:m3808. Available from: *https://doi.org/10.1136/bmj.m3808*.

[31] Lehto HR, Lehto S, Havulinna AS, Jousilahti P, Salomaa V. Gender differences in the prevalence, causes and treatment of high cardiovascular risk: findings from the FINRISK Survey. Eur J Prev Cardiol. 2012 Oct;19(5):1153–1160. Available from: *https://doi.org/10.1177/1741826711422454*.

[32] Mosca L, Linfante AH, Benjamin EJ, Berra K, Hayes SN, Walsh BW, et al. National Study of Physician Awareness and Adherence to Cardiovascular Disease Prevention Guidelines. Circulation. 2005 Feb;111(4):499–510. Available from: *https://-doi.org/10.1161/01.CIR.0000154568.43333.82*.

[33] National Institute for Health and Care Excellence. Atrial fibrillation: diagnosis and management NICE guideline [NG196]. 2021. Available from: *https://www.nice.org.uk/-guidance/ng196*.

[34] National Institute for Health and Care Excellence. Heart valve disease presenting in adults: investigation and management NICE guideline [NG208]. 2021. Available from: *https://www.nice.org.uk/guidance/ng208/chapter/Recommendations*.

[35] Visseren FLJ, Mach F, Smulders YM, Carballo D, Koskinas KC, Bäck M, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice: Developed by the Task Force for cardiovascular disease prevention in clinical practice with representatives of the European Society of Cardiology and 12 medical societies With the special contribution of the European Association of Preventive Cardiology (EAPC). Eur Heart J. 2021 Sep;42(34):3227–3337. Available from: *https://doi.org/10.1093/eurheartj/-ehab484*.

[36] Maas AHEM, Appelman YEA. Gender differences in coronary heart disease. Netherlands Heart Journal. 2010 Dec;18(12):598–603. Available from: *https://doi.org/-10.1007/s12471-010-0841-y*.

[37] George J, Rapsomaniki E, Pujades-Rodriguez M, Shah AD, Denaxas S, Herrett E, et al. How Does Cardiovascular Disease First Present in Women and Men? Circulation. 2015 Oct;132(14):1320–1328. Available from: *https://doi.org/10.1161/-CIRCULATIONAHA.114.013797*.

[38] Garcia M, Mulvagh SL, Bairey Merz CN, Buring JE, Manson JE. Cardiovascular Disease in Women. Circulation Research. 2016 Apr;118(8):1273–1293. Available from: *https://doi.org/10.1161/CIRCRESAHA.116.307547*.

[39] Sillars A, Ho FK, Pell GP, Gill JMR, Sattar N, Gray S, et al. Sex differences in the association of risk factors for heart failure incidence and mortality. Heart. 2020 Feb;106(3):203. Available from: *http://dx.doi.org/10.1136/heartjnl-2019-314878*.

[40] Millett ERC, Peters SAE, Woodward M. Sex differences in risk factors for myocardial infarction: cohort study of UK Biobank participants. BMJ. 2018 Nov;363:k4247. Available from: *https://doi.org/10.1136/bmj.k4247*.

[41] Peters SAE, Muntner P, Woodward M. Sex Differences in the Prevalence of, and Trends in, Cardiovascular Risk Factors, Treatment, and Control in the United States, 2001 to 2016. Circulation. 2019 Feb;139(8):1025–1035. Available from: *https://doi.org/-10.1161/CIRCULATIONAHA.118.035550*.

[42] Bolego C, Poli A, Paoletti R. Smoking and gender. Cardiovasc Res. 2002 Feb;53(3):568–576. Available from: *https://doi.org/10.1016/S0008-6363(01)00520-X*.

[43] Dugravot A, Fayosse A, Dumurgier J, Bouillon K, Rayana TB, Schnitzler A, et al. Social inequalities in multimorbidity, frailty, disability, and transitions to mortality: a 24-year follow-up of the Whitehall II cohort study. The Lancet Public Health. 2020 Jan;5(1):e42–e50. Available from: *https://doi.org/10.1016/S2468-2667(19)30226-9*.

[44] Singh-Manoux A, Fayosse A, Sabia S, Tabak A, Shipley M, Dugravot A, et al. Clinical, socioeconomic, and behavioural factors at age 50 years and risk of cardiometabolic multimorbidity and mortality: A cohort study. PLoS Med. 2018 May;15(5):e1002571. Available from: *https://doi.org/10.1371/journal.pmed.1002571*.

[45] Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ. 2018 Jul;362:k601. Available from: *https://doi.org/10.1136/bmj.k601*.

[46] van de Luitgaarden IAT, van Oort S, Bouman EJ, Schoonmade LJ, Schrieks IC, Grobbee DE, et al. Alcohol consumption in relation to cardiovascular diseases and mortality: a systematic review of Mendelian randomization studies. European Journal of Epidemiology. 2021 Aug. Available from: *https://doi.org/10.1007/s10654-021-00799-5*.

[47] Jiang Q, Wang K, Shi J, Li M, Chen M. No association between alcohol consumption and risk of atrial fibrillation: A two-sample Mendelian randomization study. Nutrition, metabolism, and cardiovascular diseases : NMCD. 2020 Jul;30:1389–1396. Available from: *https://doi.org/10.1016/j.numecd.2020.04.014*.

[48] Christensen AI, Nordestgaard BG, Tolstrup JS. Alcohol Intake and Risk of Ischemic and Haemorrhagic Stroke: Results from a Mendelian Randomisation Study. Journal of stroke. 2018 May;20(29886720):218–227. Available from: *https://dx.doi.org/-10.5853/jos.2017.01466*.

[49] Holmes MV, Dale CE, Zuccolo L, Silverwood RJ, Guo Y, Ye Z, et al. Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data. BMJ. 2014 Jul;349:g4164. Available from: *https://doi.org/-10.1136/bmj.g4164*.

[50] van Oort S, Beulens JWJ, van Ballegooijen AJ, Handoko ML, Larsson SC. Modifiable lifestyle factors and heart failure: A Mendelian randomization study. American Heart Journal. 2020 Sep;227:64–73. Available from: *https://doi.org/10.1016/-j.ahj.2020.06.007*.

[51] Cao M, Cui B. Negative Effects of Age at Menarche on Risk of Cardiometabolic Diseases in Adulthood: A Mendelian Randomization Study. The Journal of clinical endocrinology and metabolism. 2020 Feb;105:515–522. Available from: *https://doi.org/-10.1210/clinem/dgz071*.

[52] Levin MG, Judy R, Gill D, Vujkovic M, Verma SS, Bradford Y, et al. Genetics of height and risk of atrial fibrillation: A Mendelian randomization study. PLOS Medicine. 2020 Oct;17(10):e1003288. Available from: *https://doi.org/10.1371/-journal.pmed.1003288*.

[53] Nüesch E, Dale C, Palmer TM, White J, Keating BJ, van Iperen EP, et al. Adult height, coronary heart disease and stroke: a multi-locus Mendelian randomization meta-analysis. International journal of epidemiology. 2016 Dec;45:1927–1937. Available from: *https://doi.org/10.1093/ije/dyv074*.

[54] Carter AR, Gill D, Davies NM, Taylor AE, Tillmann T, Vaucher J, et al. Understanding the consequences of education inequality on cardiovascular disease: mendelian randomisation study. BMJ. 2019 May;365:l1855. Available from: *https://-doi.org/10.1136/bmj.l1855*.

Appendix A: Systematic literature review

Appendix

Table A1 CINAHL search terms

#	Query
S18	S8 AND S17
S17	S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16
S16	TX (population* N3 (survey? or study or studies)) OR TX (hospital* N3 (survey? or study or studies)) OR TX surveillance N4 risk* OR TX interview?
S15	TX observational N1 (study or studies)
S14	TX cohort N1 (study or studies)
S13	MH Cross Sectional Studies
S12	MH Nonconcurrent Prospective Studies
S11	MH Correlational Studies
S10	MH Case Control Studies
S9	MH Prospective Studies
S8	S1 AND S2 AND S7
S7	S3 OR S4 OR S5 OR S6
S6	TX (coronary heart disease or myocardial infarction or angina or isch?emic heart disease or coronary artery disease) AND TX (diabetes or stroke or hypertension or heart failure)
S5	TX ((longterm or long-term) N2 (disease? or condition? or disorder? or illness*)) OR TX chronic* OR TX cardiovascular* OR TX cardiometabolic*
S4	MH Cardiovascular Diseases
S3	MH Chronic Disease
S2	TX determinant? or relation* or factor? or risk? or distribution? or characteristic? or predictor? or prevalence or epidemiolog*
S1	TX multimorbid* or multi-morbid* or multipatholog* or multi-patholog* or multidiseas* or multi-diseas* or multidiagnos* or multi-diagnos* or multicondition* or multi-condition* or multidisor* or multi-disor* or "multiple chronic conditions" or "multiple chronic diseases" or "multiple chronic disease" or "multiple chronic morbidities" or "multiple disease" or "multiple diseases" or "multiple morbidity" or "multiple morbidities" or "multiple disorders" or "multiple long term conditions" or polymorbid* or poly-morbid* or polypatholog* or poly-patholog* or pluripatholog* or pluri-patholog* or polydiagnos* or poly-diagnos*

Table A2 EMBASE search terms

#	Query
1	*Comorbidity/
2	(multimorbid* or multi-morbid* or multipatholog* or multi-patholog* or multidiseas* or multi-diseas* or multidiagnos* or multi-diagnos* or multicondition* or multi-condition* or multidisor* or multi-disor* or "multiple chronic conditions" or "multiple chronic diseases" or "multiple chronic disease" or "multiple chronic morbidities" or "multiple disease" or "multiple diseases" or "multiple morbidity" or "multiple morbidities" or "multiple disorders" or "multiple long term conditions" or polymorbid* or poly-morbid* or polypatholog* or poly-patholog* or pluripatholog* or pluri-patholog* or polydiagnos* or poly-diagnos*).mp.
3	1 or 2
4	(determinant? or relation* or factor? or risk? or distribution? or characteristic? or predictor? or epidemiolog* or prevalence).mp.
5	exp Chronic Disease/
6	Cardiovascular Diseases/
7	(((longterm or long-term) adj2 (disease? or condition? or disorder? or illness*)) or chronic* or cardiovascular* or cardiometabolic*).tw.
8	((coronary heart disease or myocardial infarction or angina or isch?emic heart disease or coronary artery disease) and (diabetes or stroke or hypertension or heart failure)).mp.
9	5 or 6 or 7 or 8
10	3 and 4 and 9
11	Clinical study/
12	Case control study/
13	Family study/
14	Longitudinal study/
15	Retrospective study/
16	Prospective study/
17	Randomized controlled trials/
18	16 not 17
19	Cohort analysis/
20	(Cohort adj (study or studies)).mp.
21	(Case control adj (study or studies)).tw.
22	(follow up adj (study or studies)).tw.
23	(observational adj (study or studies)).tw.
24	(epidemiologic$ adj (study or studies)).tw.
25	(cross sectional adj (study or studies)).tw.
26	(population* adj3 (survey? or study or studies or (hospital* adj3 (survey? or study or studies)) or (surveillance adj4 risk*) or interview?)).mp.
27	Cardiovascular Diseases/ep [Epidemiology]
28	Chronic Disease/ep [Epidemiology]
29	or/11-15,18-28
30	10 and 29

Table A3 MEDLINE search terms

#	Query
1	*Comorbidity/
2	(multimorbid* or multi-morbid* or multipatholog* or multi-patholog* or multidiseas* or multi-diseas* or multidiagnos* or multi-diagnos* or multicondition* or multi-condition* or multidisor* or multi-disor* or "multiple chronic conditions" or "multiple chronic diseases" or "multiple chronic disease" or "multiple chronic morbidities" or "multiple disease" or "multiple diseases" or "multiple morbidity" or "multiple morbidities" or "multiple disorders" or "multiple long term conditions" or polymorbid* or poly-morbid* or polypatholog* or poly-patholog* or pluripatholog* or pluri-patholog* or polydiagnos* or poly-diagnos*).mp.
3	1 or 2
4	(determinant? or relation* or factor? or risk? or distribution? or characteristic? or predictor? or epidemiolog* or prevalence).mp.
5	exp Chronic Disease/
6	Cardiovascular Diseases/
7	(((longterm or long-term) adj2 (disease? or condition? or disorder? or illness*)) or chronic* or cardiovascular* or cardiometabolic*).tw.
8	((coronary heart disease or myocardial infarction or angina or isch?emic heart disease or coronary artery disease) and (diabetes or stroke or hypertension or heart failure)).mp.
9	5 or 6 or 7 or 8
10	3 and 4 and 9
11	Epidemiologic studies/
12	exp case control studies/
13	exp cohort studies/
14	Case control.tw.
15	(cohort adj (study or studies)).tw.
16	Cohort analy$.tw.
17	(Follow up adj (study or studies)).tw.
18	(observational adj (study or studies)).tw.
19	Longitudinal.tw.
20	Retrospective.tw.
21	Cross sectional.tw.
22	Cross-sectional studies/
23	((population* adj3 (survey? or study or studies)) or (hospital* adj3 (survey? or study or studies)) or (surveillance adj4 risk*) or interview?).tw.
24	Cardiovascular Diseases/ep [Epidemiology]
25	Chronic Disease/ep [Epidemiology]
26	11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25
27	10 and 26

Table A4 Web of Science search terms

#	Query
1	TS=(multimorbid* OR multi-morbid* OR multipatholog* OR multi-patholog* OR multidiseas* OR multi-diseas* OR multidiagnos* OR multi-diagnos* OR multicondition* OR multi-condition* OR multidisor* OR multi-disor* OR "multiple chronic conditions" OR "multiple chronic diseases" OR "multiple chronic disease" OR "multiple chronic morbidities" OR "multiple disease" OR "multiple diseases" OR "multiple morbidity" OR "multiple morbidities" OR "multiple disorders" OR "multiple long term conditions" OR polymorbid* OR poly-morbid* OR polypatholog* OR poly-patholog* OR pluripatholog* OR pluri-patholog* OR polydiagnos* OR poly-diagnos*)
2	TS=(determinant? or relation* or factor? or risk? or distribution? or characteristic? or predictor? or prevalence or epidemiolog*)
3	TS=(((longterm or long-term) NEAR/2 (disease? or condition? or disorder? or illness*)) or chronic* or cardiovascular* or cardiometabolic*)
4	TS=((coronary heart disease or myocardial infarction or angina or isch?emic heart disease or coronary artery disease) and (diabetes or stroke or hypertension or heart failure))
5	#4 OR #3
6	#5 AND #2 AND #1
7	TS=(Case control)
8	TS=((cohort NEAR/0 (study or studies)))
9	TS=(Cohort analy*)
10	TS=((Follow up NEAR/0 (study or studies)))
11	TS=((observational NEAR/0 (study or studies)))
12	TS=(Longitudinal)
13	TS=(Retrospective)
14	TS=(Cross sectional)
15	TS=((population* NEAR/3 (survey? or study or studies)) or (hospital* NEAR/3 (survey? or study or studies)) or (surveillance NEAR/4 risk*) or interview?)
16	#15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7
17	#16 AND #6

Table A5 Full list of data extraction items

Item	**Method**
Covidence #	Covidence
Study ID	Covidence
Title	Covidence
Reviewer Name	Covidence
Study ID	Covidence
Title	Covidence
Publication year	Covidence
Cohort/Study name	Covidence
Country in which the study conducted	Covidence
Study design	Covidence
Setting (e.g. general population, primary care)	Covidence
Sampling method	Covidence
Possible conflicts of interest for study authors	Covidence
Sample size after relevant exclusions	Covidence
Percentage of women	Covidence
Age, years	Covidence
Recruitment year(s)	Covidence
Duration of follow-up in years	Covidence
Total person-years	Covidence
Follow-up/response rate	Covidence
Method of recruitment of participants	Covidence
Definition of the concept of multimorbidity	Covidence
Disease selection criteria	Covidence
Number of selected diseases	Covidence
List of selected diseases	Covidence
Number of selected cardiovascular diseases	Covidence
List of selected cardiovascular diseases	Covidence
Multimorbidity incidence at the end of follow-up	Covidence
List of risk factors investigated	Covidence
What kind of subgroup or sensitivity analysis was done? (if any)	Covidence
Direction of association between exposure and multimorbidity	Excel spreadsheet
Categories of exposure	Excel spreadsheet
Number of incident multimorbidity cases in each category of exposure	Excel spreadsheet
Type of effect estimate (e.g. Hazard Ratio, Relative Risk)	Excel spreadsheet
Unadjusted/minimally adjusted effect estimate, CI, p-value	Excel spreadsheet
Adjusted effect estimate, CI, p-value	Excel spreadsheet
Covariates adjusted in addition to the explanatory variable	Excel spreadsheet

NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE
COHORT STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

Selection

1) Representativeness of the exposed cohort
   a) truly representative of the average in the target population ✸(all subjects or random sampling)
   b) somewhat representative of the average in the target population✸(large, diverse sample)
   c) highly selected group of users
   d) no description of the derivation of the cohort

2) Ascertainment of exposure
   a) secure record (eg surgical records) ✸
   b) structured interview ✸
   c) direct measurement ✸
   d) validated self-report ✸
   e) non-validated self-report
   f) no description

3) Demonstration that outcome of interest was not present at start of study
   a) yes ✸
   b) no

Comparability

1) Comparability of cohorts on the basis of the design or analysis
   a) controls for age, sex, and smoking ✸
   b) additionally controls for and/or sensitivity analyses for one or more of the following factors: body mass index, blood pressure and/or hypertension, serum cholesterol level and/or hypercholesterolemia, blood glucose level and/or diabetes, socioeconomic status ✸

Outcome

1) Assessment of outcome
   a) independent blind assessment ✸
   b) record linkage ✸
   c) validated self-report ✸
   d) non-validated self-report
   e) no description

2) Was follow-up long enough for outcomes to occur
   a) yes (follow-up period of at least five years) ✸
   b) no

3) Adequacy of follow up of cohorts
   a) complete follow up - all subjects accounted for ✸
   b) subjects lost to follow up unlikely to introduce bias (> 90 % follow up, or description provided of those lost) ✸
   c) follow up rate < 90% and no description of those lost
   d) no statement

Figure A1 Revised Newcastle-Ottawa Quality Assessment Scale for cohort studies

Table A6 Statistical methods and data sources used in each study

		Source of outcome data		Source of exposure data		
Author, year	**Statistical method**	**Self-report**	**Data linkage/ health records**	**Self-report**	**Data linkage**	**Direct measure**
Dugravot, 2020	proportional hazards regression with Weibull distribution with inverse probability weighting and interval censoring	✓	✓	✓ Education	✓ Occupation	
Keats, 2020	logistic regression	✓		Physical activity		
Kivimäki, 2017	logistic regression and random-effects meta-analysis	✓	✓	✓		✓
Li, 2021	Cox proportional hazards regression		✓	✓		
Mounce, 2018	time-to-event complementary log-log models (equivalent to Cox proportional hazards models for continuous time)	✓		✓		BMI
Shang, 2020	Not clear (for BMI), Cox regression (for number of health factors)		✓	✓ BMI and lifestyle factors	✓ Indices for SES and self-rated health	
Singh-Manoux, 2018	Cox regression; unidirectional multi-state models with Markov proportional hazards		✓	✓ Behavioural factors		✓ Clinical profile
Wikstrom, 2015	Cox regression		✓	✓ Behavioural factors		✓ Clinical profile
Xu, 2018	logistic regression	✓ (validated)		✓		
Xu, 2019b	logistic regression	✓		✓		

Table A7 Main findings of each study

Author, year	**Participants (location)**	**Adjustment variables**	**Main findings (fully adjusted)**
Dugravot, 2020	1,694 cases/6,425 in cohort (UK)	age, sex, race, marital status, and birth cohort	Hazard ratios (HRs) for those in high, medium, and low groups of education: 1 (referent), 1.21 (1.05-1.40), and 1.26 (1.10-1.44), respectively; of occupation: 1 (referent), 1.36 (1.21-1.52), and 1.66 (1.40-1.97), respectively
Keats, 2020	2,638 cases/6,912 in cohort (Canada)	sex, age, rurality, socioeconomic status (both material/social deprivation scores quintile), alcohol consumption, daily servings of fruit and vegetables, smoking status, and physical activity/walkability where appropriate	Odds ratios (ORs) for those who engaged in high, moderate, and low levels of physical activity: 1 (referent), 0.94 (0.83-1.06), and 1.26 (1.11-1.42), respectively
Kivimäki, 2017	1,627 cases / 120,813 in cohort (USA & Europe)	age, sex, ethnic origin, lifestyle factors (smoking, physical activity, alcohol consumption)	ORs in participants with BMI of <20, 20-24.9, 25-29.9, 30-34.9, and ≥35 kg/m^2 were 0.9 (0.6-1.5), 1.0 (reference), 2.0 (1.7-2.4), 4.5 (3.5-5.8), and 14.5 (10.1-21.0), respectively
Li, 2021	3,151 (in women), 2,598 (in men)/29,838 (Canada)	stratified by sex, adjusted for age group, marital status, immigration status, annual household income, current smoking, alcohol use, physical activity, BMI, has a regular doctor, morbidity score, depression	In women, HRs for developing a second physical illness (among those who developed a first physical illness) associated with depression 1.16 (1.02-1.33) [vs. no depression], heavy drinking 1.03 (0.86-1.23) [vs. light drinking], obesity 1.37 (1.22-1.53) [vs. normal weight], physical activity 0.97 (0.90-1.04) [vs. no physical activity], current smoking 1.05 (0.97-1.15) [vs. non-smoking] In men, HRs associated with depression 1.03 (0.84-1.26) [vs. no depression], heavy drinking 1.08 (0.94-1.25) [vs. light drinking], obesity 1.62 (1.44-1.83) [vs. normal weight], physical activity 0.95 (0.87-1.03) [vs. no physical activity], current smoking 1.25 (1.14-1.36) [vs. non-smoking]
Mounce, 2018	377 cases /1,477 in cohort (UK)	age, sex, wealth quintiles, education, lives alone, BMI, smoking, physical activity, social	HRs for no educational qualifications 0.73 (0.53- 1.00 [vs. degree/higher ed], obesity 1.92 (1.43- 2.59) [vs. normal or underweight], low physical activity

		detachment, external locus of control	1.43 (1.02-2.00) [vs. high activity], current smoking 1.22 (0.87-1.70) [vs. never smoking]
Shang, 2020	7,098 (in women), 5,955 (in men) /53,867 (Australia)	Not clear	In men, relative risks of having ≥2 diseases in overweight and obese men were 1.44 (1.35-1.53) and 2.26 (2.10-2.43) compared to men of normal weight, respectively. Relative risks of having ≥3 diseases in overweight and obese men were 1.52 (1.38-1.67) and 2.64 (2.37-2.94), respectively. Relative risks of having ≥4 diseases in overweight and obese men were 1.65 (1.41-1.93) and 2.93 (2.47-3.48), respectively. The study does not report the relative risks in women, but instead reports that "similar results were seen in women".
Singh-Manoux, 2018	511 cases / 8,270 in cohort (UK)	age, sex, race, marital status, and birth cohort	HRs for hypertension 1.49 (1.23-1.80), high cholesterol 1.32 (0.98-1.77), overweight 1.79 (1.50-2.14), family history of cardiometabolic disease 1.07 (0.86-1.34), current smoking 2.01 (1.64-2.47), alcohol consumption 1.34 (1.13-1.61), poor diet 1.20 (1.01-1.43), physical inactivity 1.21 (1.01-1.45) Individual risk factors were dichotomous; reference group was composed of persons without the risk factor. Hypertension was defined as blood pressure of ≥140/90 mm Hg or use of antihypertensive medication. High cholesterol corresponded to total blood cholesterol of ≥5 mmol/L. Overweight corresponded to BMI ≥25kg/m^2. Alcohol consumption referred to abstainers (<1 unit/week) plus heavy drinkers (>14 units/week in women and > 21 units/week in men). Poor diet referred to fruit and vegetable consumption <once a day. Physical inactivity corresponded to <2.5h of moderate and vigorous physical activity per week.
Wikstrom, 2015	166 (in women), 311 (in men) / 32,972 in cohort (Finland)	age, study year, study area, BMI, physical activity, smoking, education, blood pressure, blood cholesterol level	Women: HR = 1.08 (1.05-1.11) per unit increase in BMI, kg/m^2; 1.62 (1.14-2.30) for low v. high physical activity; 2.55 (1.76-3.71) for current smokers v. non-smokers; 1.03 (0.95-1.12) per 10 unit increase in blood pressure, mm Hg; 1.07 (0.93-1.25) per unit increase in cholesterol, mmol/L; 1.37 (0.90-2.10) for low vs. high education; 1.10 (0.78-1.55) for low (less than 3-5 times a week) vs. high fruit and vegetable consumption Men: HR = 1.11 (1.08-1.14) per unit increase in BMI, kg/m^2; 1.34 (1.03-1.73) for low vs. high physical activity; 2.68 (2.10-3.41) for current smokers vs. non-smokers; 1.14 (1.07-1.20) per 10 unit increase in blood pressure, mm Hg; 1.09

			(0.99-1.21) per unit increase in cholesterol, mmol/L; 1.40 (1.02-1.91) for low vs. high education; 1.07 (0.83-1.39) for low (less than 3-5 times a week) vs. high fruit and vegetable consumption.
Xu, 2018	423 cases / 11,941 in cohort (Australia)	age, marital status, area of residence, education, country of birth, ability to manage on income, BMI, hypertension, physical activity, smoking status, other chronic conditions (depression/anxiety, COPD, asthma, cancer, arthritis, osteoporosis)	ORs for no educational qualifications 1.48 (0.99-2.20) [vs. university/higher degree], obesity 3.66 (2.76-4.86) [vs. normal weight], hypertension 1.94 (1.43-2.64) [vs. no hypertension], physical inactivity 1.31 (0.91-1.88) [vs. high activity], current smoking 2.00 (1.51-2.65) [vs. never smoking], depression/anxiety 1.18 (0.86-1.62) [vs. no other condition] The ORs were estimated using the incidence of multimorbidity during the 2-year follow up and associated with predictors at baseline, compared with women who developed 0 condition. Obesity corresponded to BMI ≥30 kg/m^2, and normal weight to BMI 18.5-24.9 kg/m^2. Low physical activity corresponded to 0-39 MET min/week, and high physical activity to ≥1200 MET min/week.
Xu, 2019b	4,442 cases/7,357 in cohort (Australia)	age, marital status, area of residence, education, country of birth, ability to manage on income, BMI, physical activity, alcohol intake, smoking status, menopausal status	ORs in participants with baseline BMI of <18.5, 18.5-24.9, 25-29.9, 30-34.9, and ≥35 kg/m^2 were 0.82 (0.63-1.05), 1.0 (reference), 1.29 (1.21-1.39), 1.66 (1.52-1.82), and 1.81 (1.60-2.04), respectively

Table A8 List of diseases to define multimorbidity and exposures investigated in each study

Author, year	**List of selected diseases to define multimorbidity**	**All investigated exposures**
Dugravot, 2020	diabetes, coronary heart disease, stroke, chronic obstructive pulmonary disease, depression, arthritis, cancer, dementia, Parkinson's disease	three indicators of socioeconomic status (education, occupational position, and literacy) at age 50 years
Keats, 2020	cancer (excluding skin cancer), cardiovascular disease (myocardial infarction and/or stroke), diabetes, chronic respiratory disease (COPD, chronic bronchitis, and/or emphysema), depression	physical activity, walkability quintile
Kivimäki, 2017	diabetes, coronary heart disease, stroke	body mass index (BMI)
Li, 2021	acute myocardial infarction, asthma, cancer, cardiac arrhythmia, chronic coronary syndrome, chronic obstructive pulmonary disease, congestive heart failure, dementia, diabetes, hypertension, osteoarthritis, osteoporosis, renal failure, rheumatoid arthritis, stroke	depression (main risk factor), age, marital status, immigration status, annual household income, smoking, alcohol use, physical activity, BMI, having a regular doctor
Mounce, 2018	hypertension, ischemic heart disease, congestive heart failure, heart arrhythmia, diabetes mellitus, stroke, chronic obstructive pulmonary disease, asthma, arthritis, osteoporosis, cancer, Parkinson's disease, 1 or more affective mental health conditions (depression, anxiety, or emotional problems), 1 or more psychotic mental health conditions (schizophrenia, psychosis, bipolar disorder, or hallucinations), dementia (including Alzheimer's disease).	age, sex, wealth quintiles, educational attainment, living alone (vs cohabiting), BMI, smoking behaviour, physical activity, social detachment (vs none), external locus of control (vs internal)
Shang, 2020	cancer (excluding non-melanoma skin cancer), heart disease, stroke, hypertension, dyslipidaemia, diabetes, asthma, mental disorders (depression and anxiety), degenerative disorders (dementia and Parkinson's disease), hip replacement, osteoarthritis	BMI, number of health factors (out of 10 derived from machine learning methods: BMI between 18.5 and 24.9 kg/m^2, fruit intake ≥2 servings/day, vegetable intake ≥3 servings/day, red meat intake ≤1 serving/week, chicken intake ≤1 serving/week, physical activity ≥5 sessions/week, alcohol consumption between 1 and 4 drinks/week, never smoking, none-passive

		smoking, and sleep duration between 7 and 9 h/day.)
Singh-Manoux, 2018	coronary heart disease, stroke, diabetes	Three sets of risk factors, each individually and on a 5-point scale: clinical profile (hypertension [≥140/90 mm Hg or use of antihypertensive medication], hypercholesterolemia [total cholesterol ≥5 mmol/L], overweight/obesity [BMI≥25kg/m^2], family history of cardiometabolic disease), occupational position, and behavioural factors (current smoking, abstainers/heavy alcohol consumption, poor diet [fruit and vegetable consumption <once a day], physical inactivity [<2.5h of moderate and vigorous physical activity per week]).
Wikstrom, 2015	diabetes, cardiovascular disease (coronary heart disease or stroke, excluding subarachnoid haemorrhage), asthma/chronic obstructive pulmonary disease, cancer, rheumatoid arthritis	BMI, physical activity, smoking, education, blood pressure, cholesterol, fruit and vegetable consumption
Xu, 2018	diabetes, heart disease, stroke	sociodemographic factors, lifestyle factors, and other prior chronic conditions
Xu, 2019b	cancer, cardiovascular disease, hypertension, mental health disorders (depression/anxiety), diabetes mellitus, asthma, chronic obstructive pulmonary disease, arthritis, osteoporosis	baseline BMI, time-varying BMI, and short-term weight change

Appendix B: Million Women Study recruitment questionnaire

THE MILLION WOMEN STUDY

A national survey of women invited for breast screening

We need one million women to help us in research that will benefit women all over the world.
Would you become one of these special women?

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

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

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

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

Even if you are not taking HRT it is just as important that you fill in the questionnaire.
Please bring this questionnaire to your breast screening appointment

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

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

OR put numbers in the appropriate box e.g. *23rd April 1946* 2 3 / 0 4 / 4 6 age 4 4 years

SAMPLE

GENERAL QUESTIONS ABOUT YOU

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

☐☐ / ☐☐ / ☐☐

2. How old are you? ☐☐ years

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

☐ feet ☐☐ inches

4. About how much do you weigh?

☐☐ stone ☐☐ lbs

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

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

6. What qualification(s) do you have from school, college or the equivalent?
(please put a cross in the most appropriate box(es))

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

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

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

8. Are you an ex-smoker? ☐ No ☐ Yes

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

Wine (glasses per week)	**Lager/Cider/Beer** (half pints per week)	**Spirits** (tots per week)
☐ none	☐ none	☐ none
☐ less than 1	☐ less than1	☐ less than1
☐ 1-3	☐ 1-3	☐ 1-3
☐ 4-6	☐ 4-6	☐ 4-6
☐ 7-10	☐ 7-10	☐ 7-10
☐ 11-15	☐ 11-15	☐ 11-15
☐ 16-20	☐ 16-20	☐ 16-20
☐ 21+	☐ 21+	☐ 21+

If you drink **wine** is it

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

10. How often do you do *any* exercise?

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

11. How often do you do *strenuous* exercise?
(that is, enough to cause sweating or a fast heart beat.)

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

QUESTIONS ABOUT YOU AND YOUR FAMILY

12. Have you ever had any children? ☐ No ☐ Yes
- if No, please go on to question 15

13. How many children have you had? ☐☐
(please include stillbirths; it is not necessary to include miscarriages)

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

	DATE OF BIRTH (If you had twins or triplets please repeat the same date for each child) day / month / year	BREASTFEEDING (Months that you breastfed each child; put "0" if you did not breastfeed that child a "1" if you breastfed for month or less)
1st child	☐☐ / ☐☐ / ☐☐	☐☐ months
2nd child	☐☐ / ☐☐ / ☐☐	☐☐ months
3rd child	☐☐ / ☐☐ / ☐☐	☐☐ months
4th child	☐☐ / ☐☐ / ☐☐	☐☐ months
5th child	☐☐ / ☐☐ / ☐☐	☐☐ months
6th child	☐☐ / ☐☐ / ☐☐	☐☐ months
7th child	☐☐ / ☐☐ / ☐☐	☐☐ months
8th child	☐☐ / ☐☐ / ☐☐	☐☐ months
9th child	☐☐ / ☐☐ / ☐☐	☐☐ months
10th child	☐☐ / ☐☐ / ☐☐	☐☐ months

15. Have you ever been for breast screening before?

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

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

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

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

17. Have you ever had breast cancer diagnosed?

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

18. Has your mother ever had breast cancer diagnosed?

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

19. How many sisters do you have? ☐☐ sisters
(put "0" if you do not have any sisters, please include any sisters who have died)

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

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

1st sister ☐☐ years **2nd sister** ☐☐ years

QUESTIONS ABOUT YOUR HEALTH

21. Have you ever had any other cancer?

☐ Yes ☐ No

Please describe

22. Have you EVER had:
(please cross "Yes" or "No" for each condition)

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

23. Are you NOW being treated for:

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

SAMPLE

24. Are you *NOW* being treated for any other *serious* illness?
☐ Yes ☐ No

Please describe this illness

Please describe the treatment

QUESTIONS ABOUT PAST OPERATIONS

25. Have you had a hysterectomy?
☐ No
☐ Yes- If **Yes**, how old were you? ☐☐ years

26. Have you had BOTH ovaries removed?
☐ No ☐ Not sure
☐ Yes- If **Yes**, how old were you? ☐☐ years

27. Have you been sterilised *(had your tubes tied)*?
☐ No
☐ Yes- If **Yes**, how old were you? ☐☐ years

QUESTIONS ABOUT YOUR USE OF THE PILL

28. Have you ever used the pill *(oral contraceptive)*?
☐ Yes
☐ No - *if No, please go to question 32*

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

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

31. For how many years *in total* did you take the pill?
☐☐ years
(Add together the years and months when you actually took the pill -do not count the years and months when you were not taking it. Please write "0" if you used the pill for less than a year in total)

QUESTIONS ABOUT YOUR USE OF HORMONE REPLACEMENT THERAPY (HRT)

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

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

34. Had your periods stopped before you started using HRT? *(Cross "Yes" if you had a hysterectomy before starting HRT)*
☐ No
☐ Yes - **if Yes**, how old were you when your periods stopped? ☐☐ years

35. For about how many years *in total* have you used HRT? ☐☐ years
(Add together the years and months when you used HRT - do not count the years and months when you were not using HRT. Please write "0" if you used HRT for less than a year in total)

36. Are you *now* using HRT?
☐ Yes
☐ No - **if No**, how old were you when you **last** used HRT? ☐☐ years

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

☐ Prempak C 0.625mg	☐ Premarin 0.625mg
☐ Prempak C 1.25mg	☐ Premarin 1.25mg
☐ Tridestra	☐ Evorel 25mcg/50mcg
☐ Trisequens	☐ Evorel 75mcg/100mcg
☐ Trisequens Forte	☐ Progynova 1mg
☐ Cycloprogynova 1mg	☐ Progynova 2mg
☐ Cycloprogynova 2mg	☐ Estraderm 25mcg
☐ Estrapak	☐ Estraderm 50mcg
☐ Estracombi	☐ Estraderm 100mcg
☐ Climaval 1mg	☐ Zumenon 1mg
☐ Climaval 2mg	☐ Zumenon 2mg
☐ Premique Cycle	☐ Ethinyloestradiol
☐ Premique	☐ Micronor
☐ Nuvelle	☐ Provera
☐ Kliofem	☐ Duphaston
☐ Livial	

☐ Do not know ☐ Implants ☐ Oestrogel

Other *(please write here)*

38. For how many years in total did you use the most *recent* type of HRT? ☐☐ years
(Please write "0" if you used this recent HRT for less than a year in total)

SAMPLE

QUESTIONS ABOUT YOUR PERIODS

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

40. Have your periods NOW stopped?

*Cross "**Yes**"-if you are not having periods now, either because of your menopause, after a hysterectomy or after stopping HRT.*
*Cross "**No**"- if you are still having regular periods now, even if they are because you are taking HRT.*
*Cross "**Irregular**"-if your periods have been irregular and you think it might be because of the menopause.*

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

☐ **No**

☐ **Irregular**

FINAL SECTION

41. So that we can find out about your health in the future we may need to contact you again or look at your screening or medical records. We would be grateful if you gave us permission to contact you again or to use information from those records. **We guarantee that all information obtained will be treated with absolute confidentiality and used for medical research only.** Of course, you do not have to give permission. Your response to this request will not affect your screening or the treatment you receive in any way.

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

Signature: ☐ today's date: ☐☐ ☐☐ ☐☐

Surname: ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

Given name(s): ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

House number and street: ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

District: ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

Town: ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

County: ☐☐☐☐☐☐☐☐☐☐ Postcode: ☐☐☐☐☐☐☐☐

Surname at birth: ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

Town of birth: ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

NHS (National Health Service)Number: ☐☐☐☐☐☐☐☐☐☐
(The number on your medical card)

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

SAMPLE

Please bring this questionnaire to your breast screening appointment

If you would like to post this questionnaire back to us, please send it to:
THE MILLION WOMEN STUDY CO-ORDINATING CENTRE
ICRF-CEU, GIBSON BUILDING,
RADCLIFFE INFIRMARY,
OXFORD OX2 6HE

FREEPHONE: 0800 262 872

Breast Screening Unit
Linton House
Thirlestaine Road
Cheltenham
GL53 7AS

LGL ☐

THANK YOU VERY MUCH FOR YOUR HELP

Appendix C: Descriptive study of cardiovascular multimorbidity

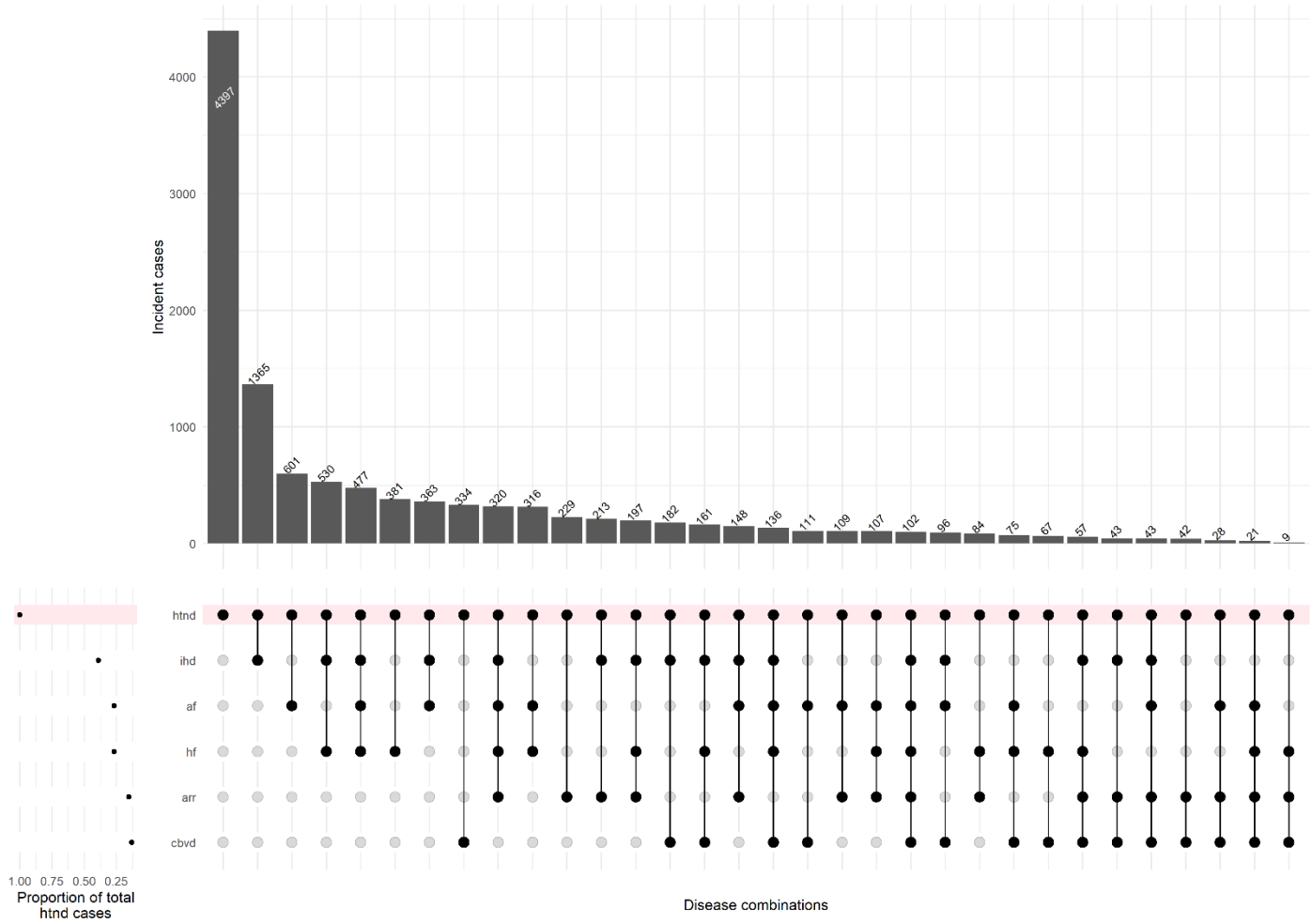


Figure C1 UpSet plot of the top five comorbidities of hypertensive heart and renal disease (n = 11,344)

hntd=hypertensive heart and renal disease, ihd=ischaemic heart disease, af=atrial fibrillation, hf=heart failure, arr=other arrhythmias, cbvd=other cerebrovascular disease

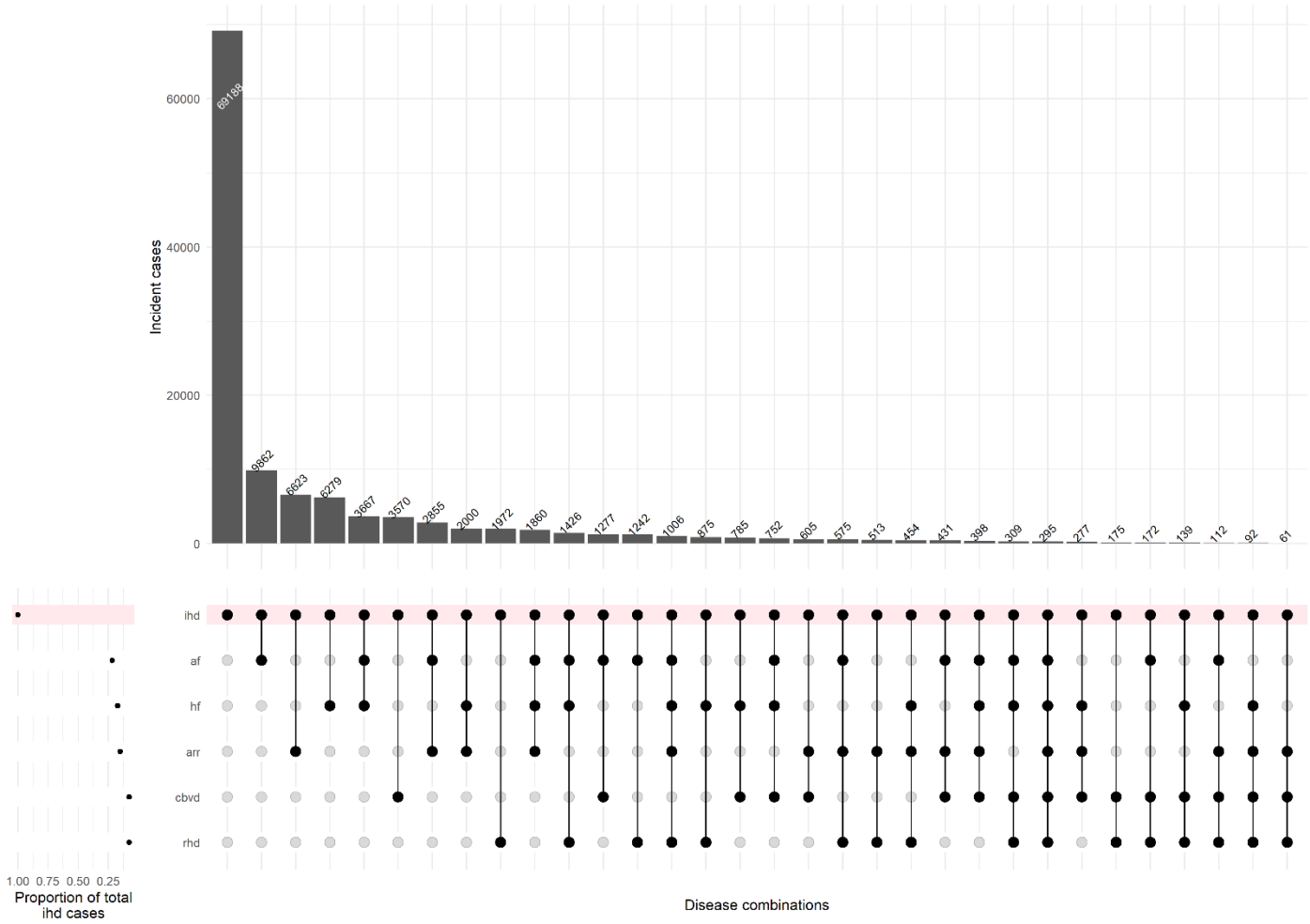


Figure C2 UpSet plot of the top five comorbidities of ischaemic heart disease (n = 119,847)

ihd=ischaemic heart disease, af=atrial fibrillation, hf=heart failure, arr=other arrhythmias, cbvd=other cerebrovascular disease, rhd=rheumatic heart disease

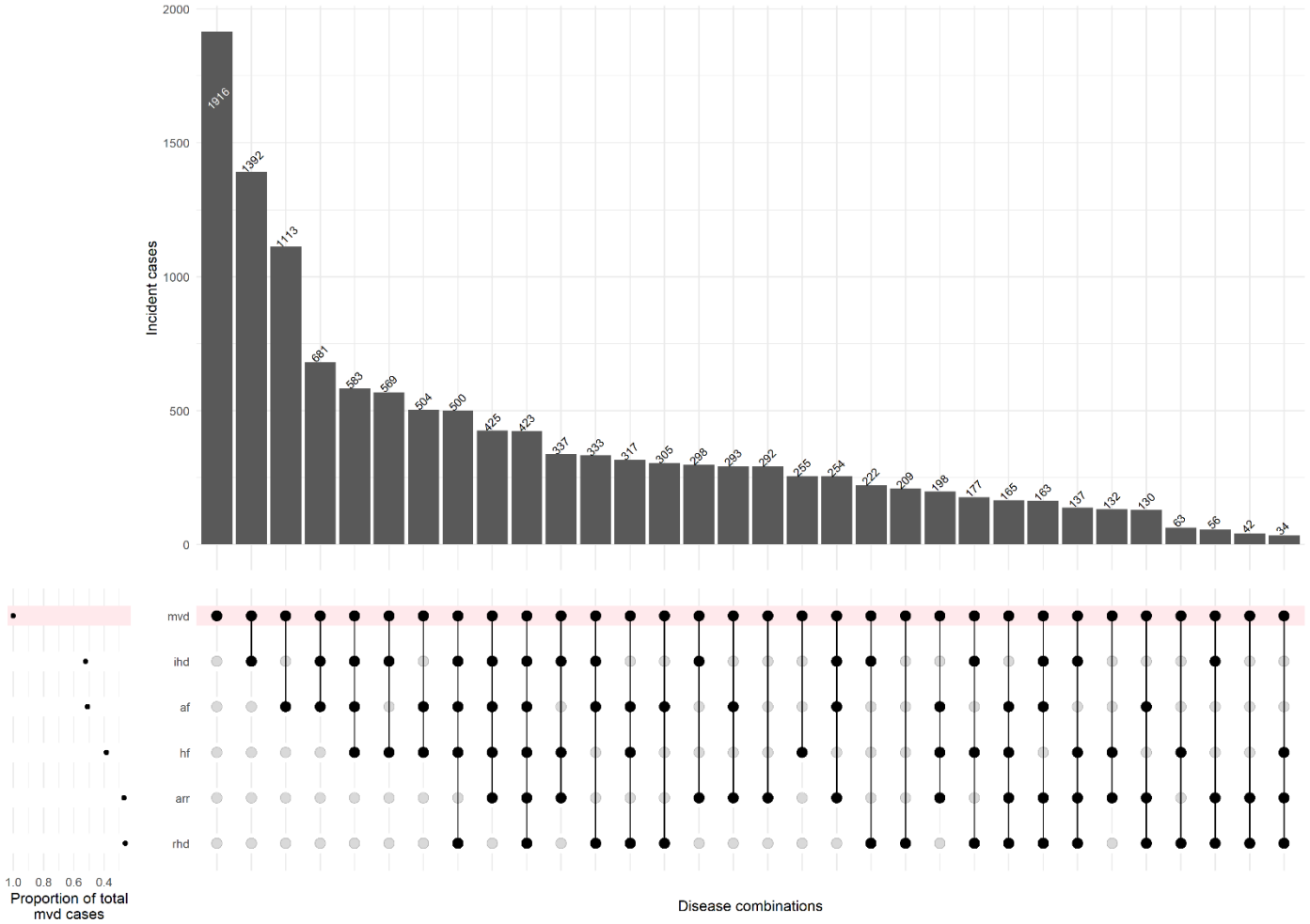
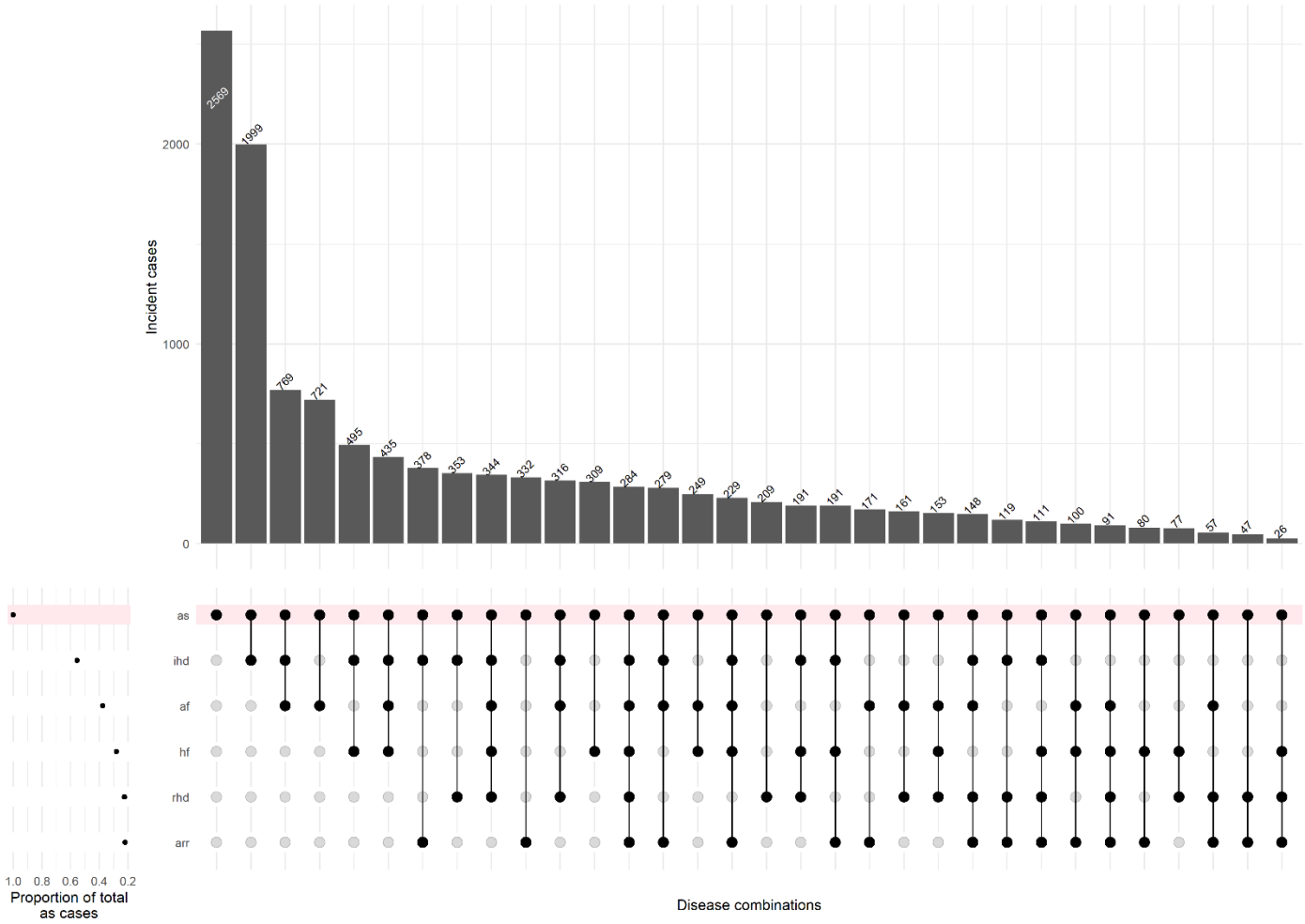


Figure C3 UpSet plot of the top five comorbidities of mitral valve disorders (n = 12,518)

mvd=mitral valve disorders, ihd=ischaemic heart disease, af=atrial fibrillation, hf=heart failure, arr=other arrhythmias, rhd=rheumatic heart disease



FigureC4 UpSet plot of the top five comorbidities of aortic stenosis (n = 11,993)

as=aortic stenosis, ihd=ischaemic heart disease, af=atrial fibrillation, hf=heart failure, rhd=rheumatic heart disease, arr=other arrhythmias

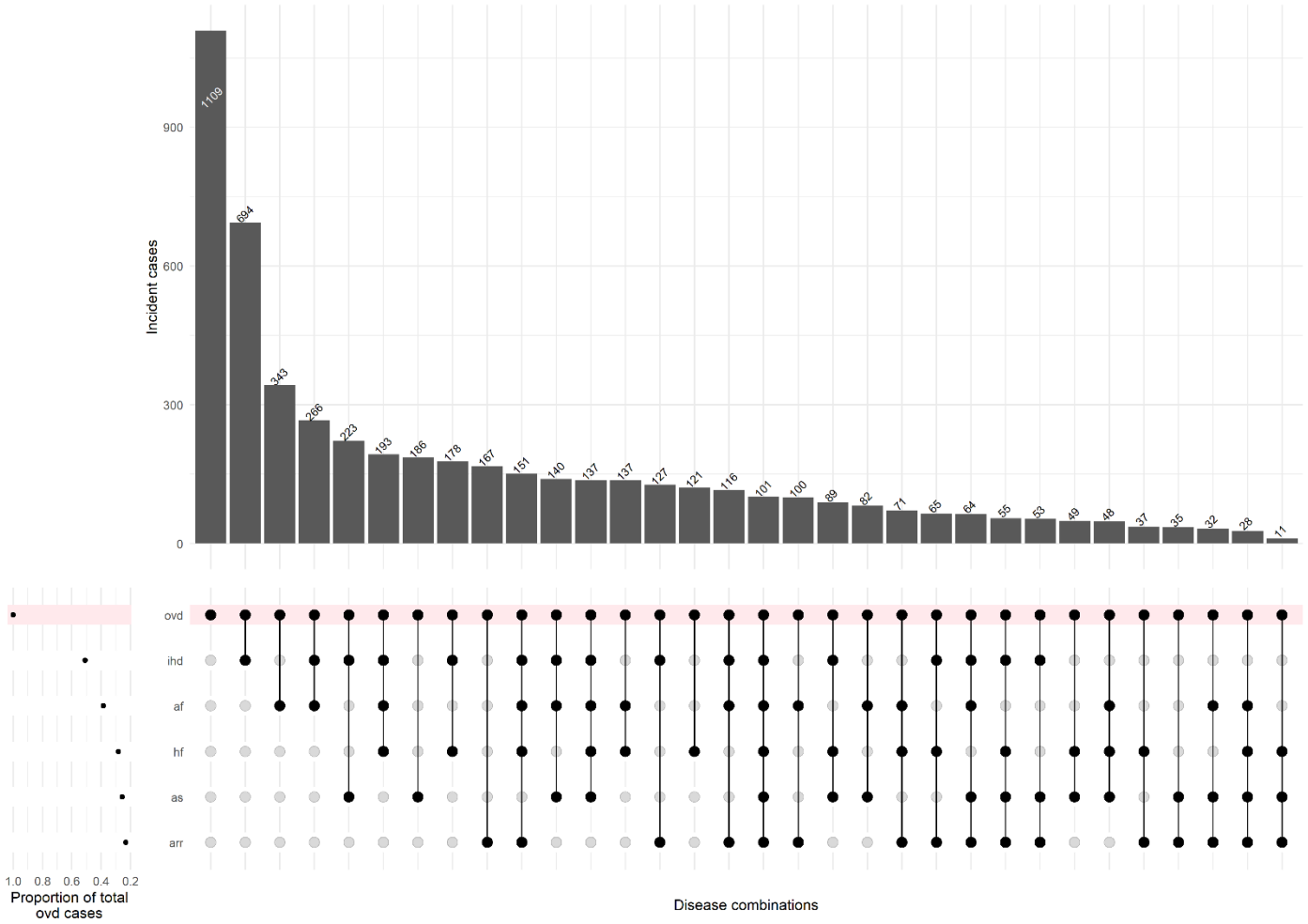


Figure C5 UpSet plot of the top five comorbidities of other aortic valve disorders (n = 5,208)

ovd=other aortic valve disorders, ihd=ischaemic heart disease, af=atrial fibrillation, hf=heart failure, as=aortic stenosis, arr=other arrhythmias

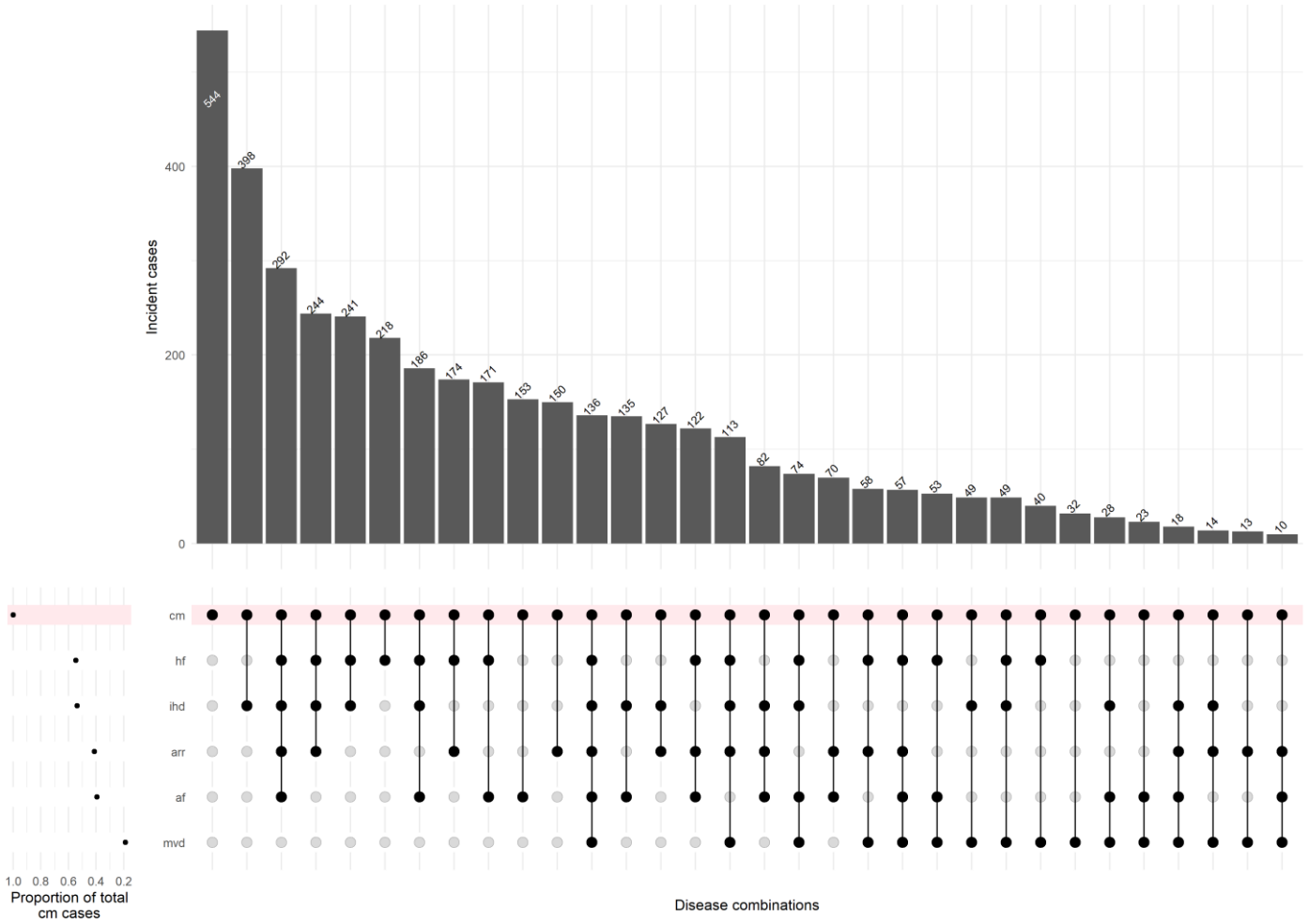


Figure C6 UpSet plot of the top five comorbidities of cardiomyopathy (n = 4,074)

cm=cardiomyopathy, hf=heart failure, ihd=ischaemic heart disease, arr=other arrhythmias, af=atrial fibrillation, mvd=mitral valve disorders

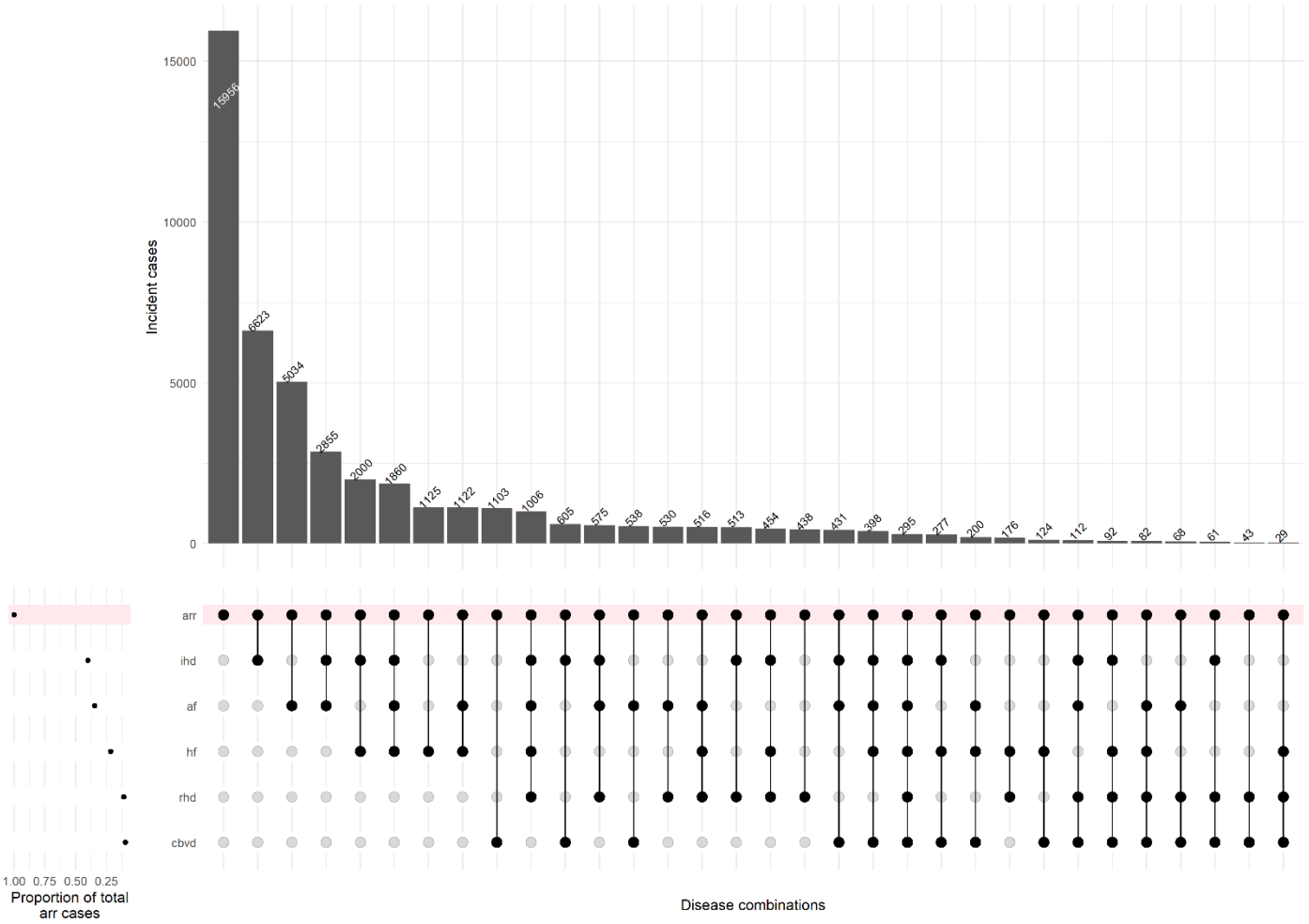


Figure C7 UpSet plot of the top five comorbidities of other arrhythmias (n = 45,241)

arr=other arrhythmias, ihd=ischaemic heart disease, af=atrial fibrillation, hf=heart failure, rhd=rheumatic heart disease, cbvd=other cerebrovascular disease

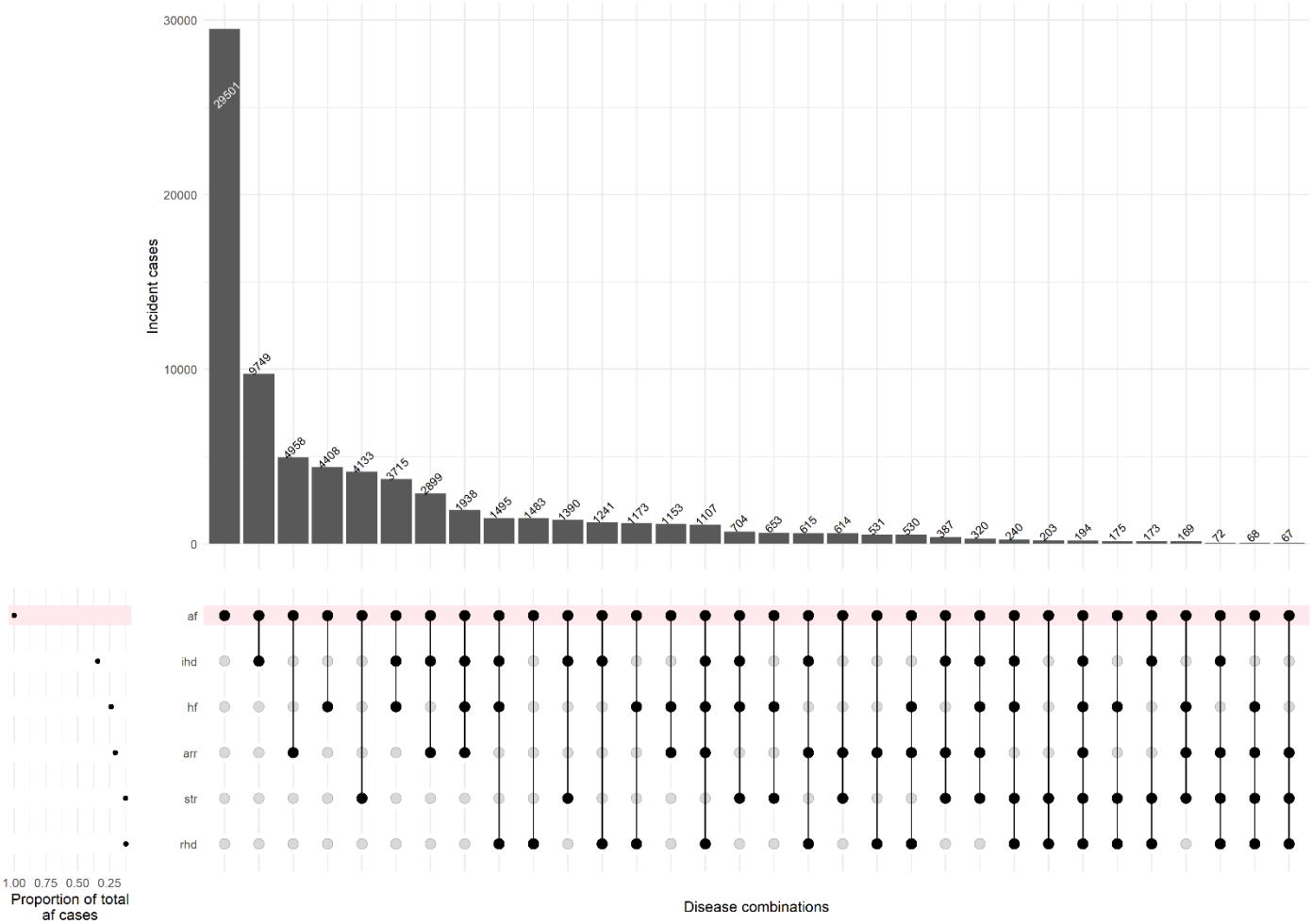


Figure C8 UpSet plot of the top five comorbidities of atrial fibrillation (n = 76,058)

af=atrial fibrillation, ihd=ishcaemic heart disease, hf=heart failure, arr=other arrhythmias, str=stroke, rhd=rheumatic heart disease

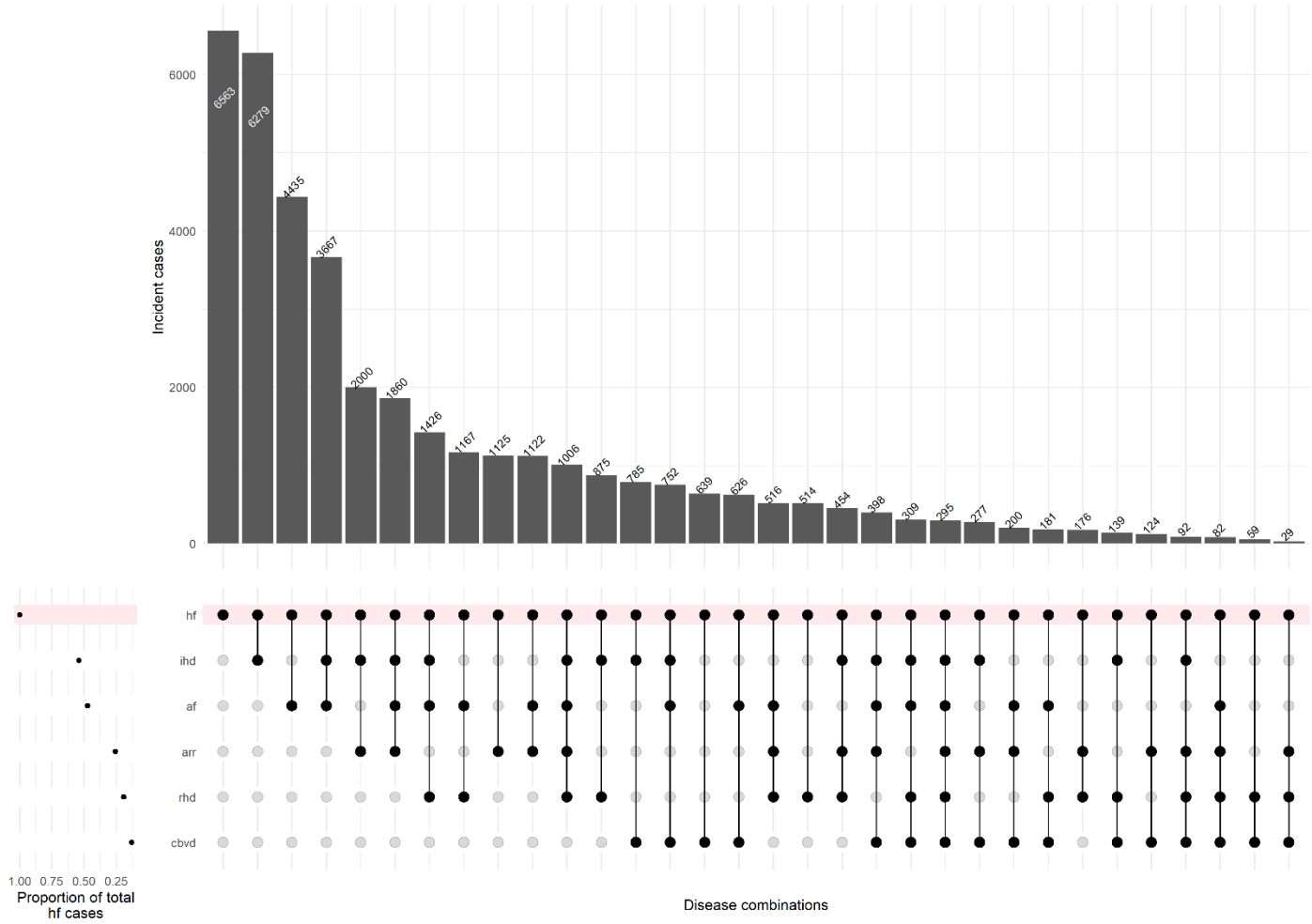


Figure C9 UpSet plot of the top five comorbidities of heart failure (n = 38,172)

Hf=heart failure, ihd=ischaemic heart disease, af=atrial fibrillation, other arrhythmias, rhd=rheumatic heart disease, cbvd=other cerebrovascular disease

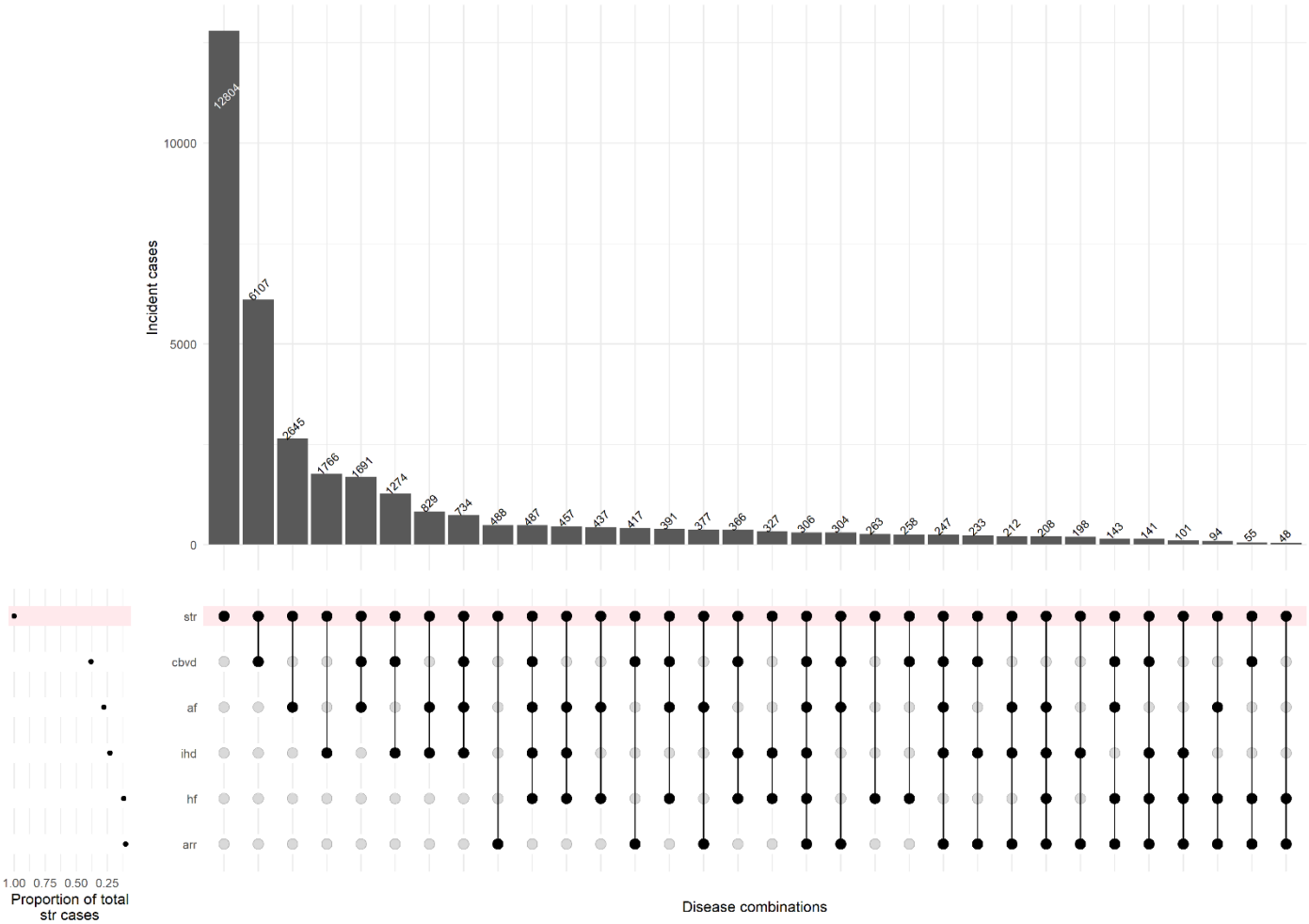


Figure C10 UpSet plot of the top five comorbidities of stroke (n = 34,408)

str=stroke, cbvd=other cerebrovascular disease, af=atrial fibrillation, ihd=ischaemic heart disease, hf=heart failure, arr=other arrhythmias

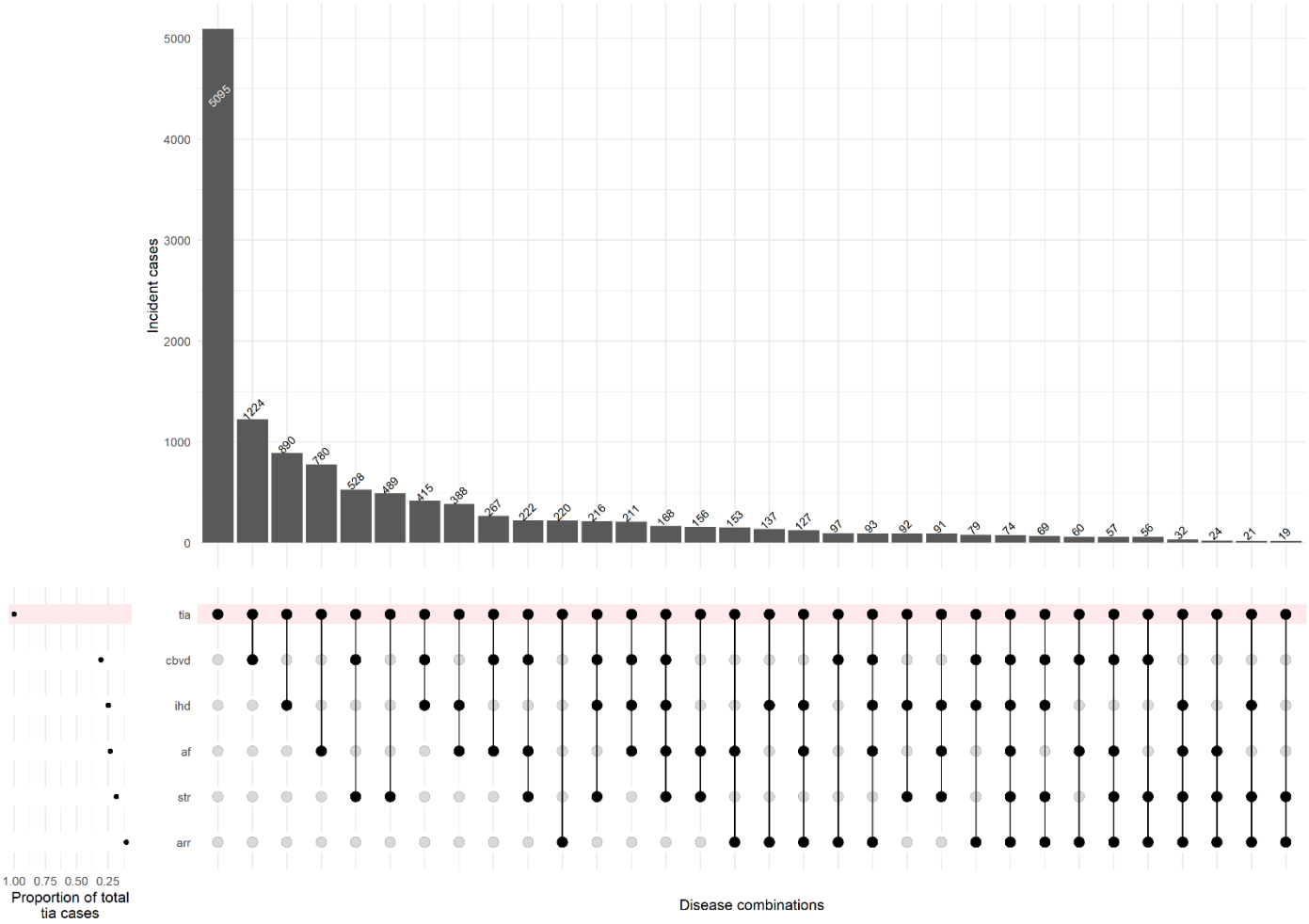


Figure C11 UpSet plot of the top five comorbidities of transient cerebral ischaemic attacks (n = 12,550)

tia=transient cerebral ischaemic attacks, cbvd=other cerebrovascular disease, ihd=ischaemic heart disease, af=atrial fibrillation, str=stroke, arr=other arrhythmia

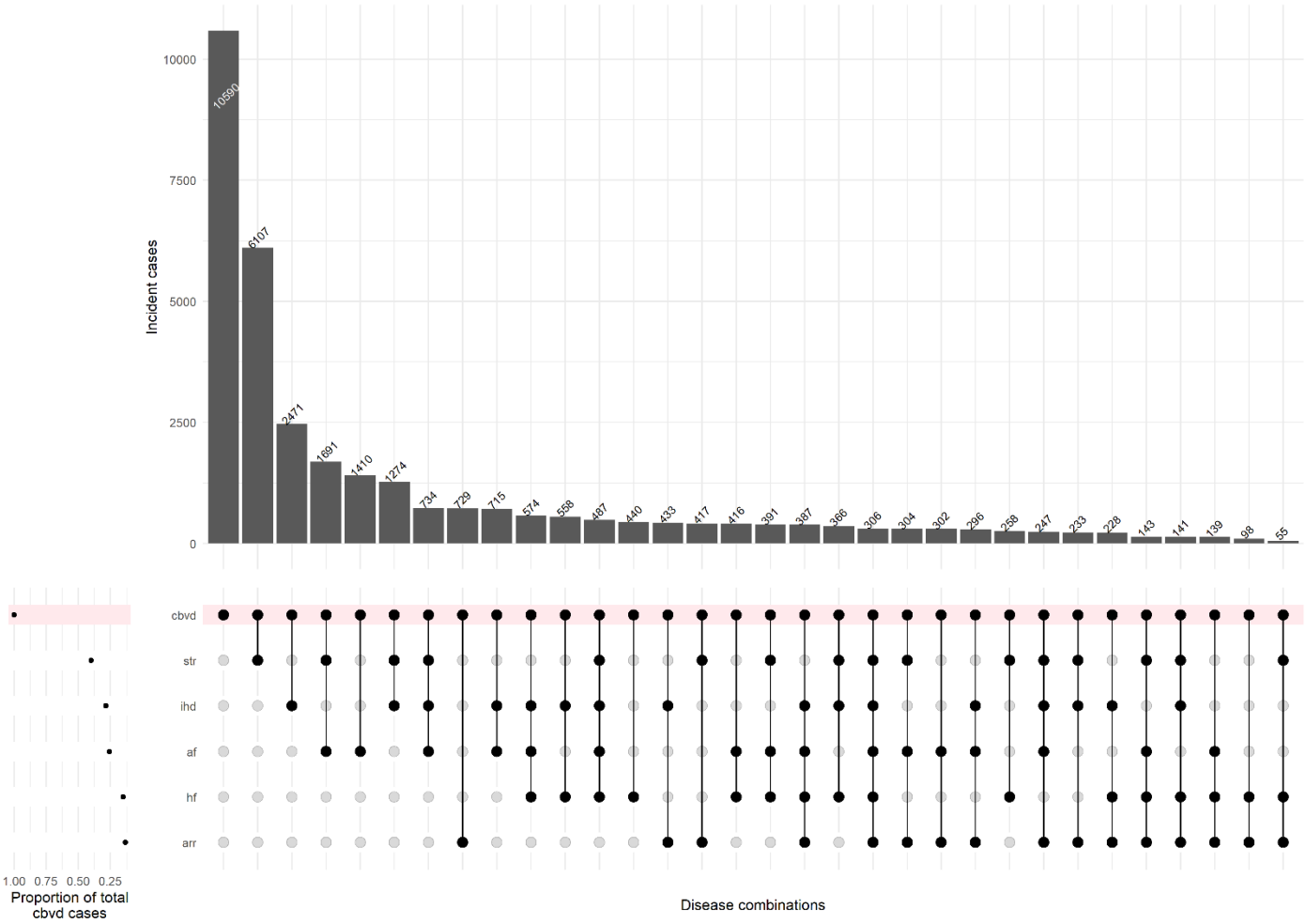


Figure C12 UpSet plot of the top five comorbidities of other cerebrovascular disease (n = 32,940)

cbvd=other cerebrovascular disease, str=stroke, ihd=ischaemic heart disease, af=atrial fibrillation, hf=heart failure, arr=other arrhythmias

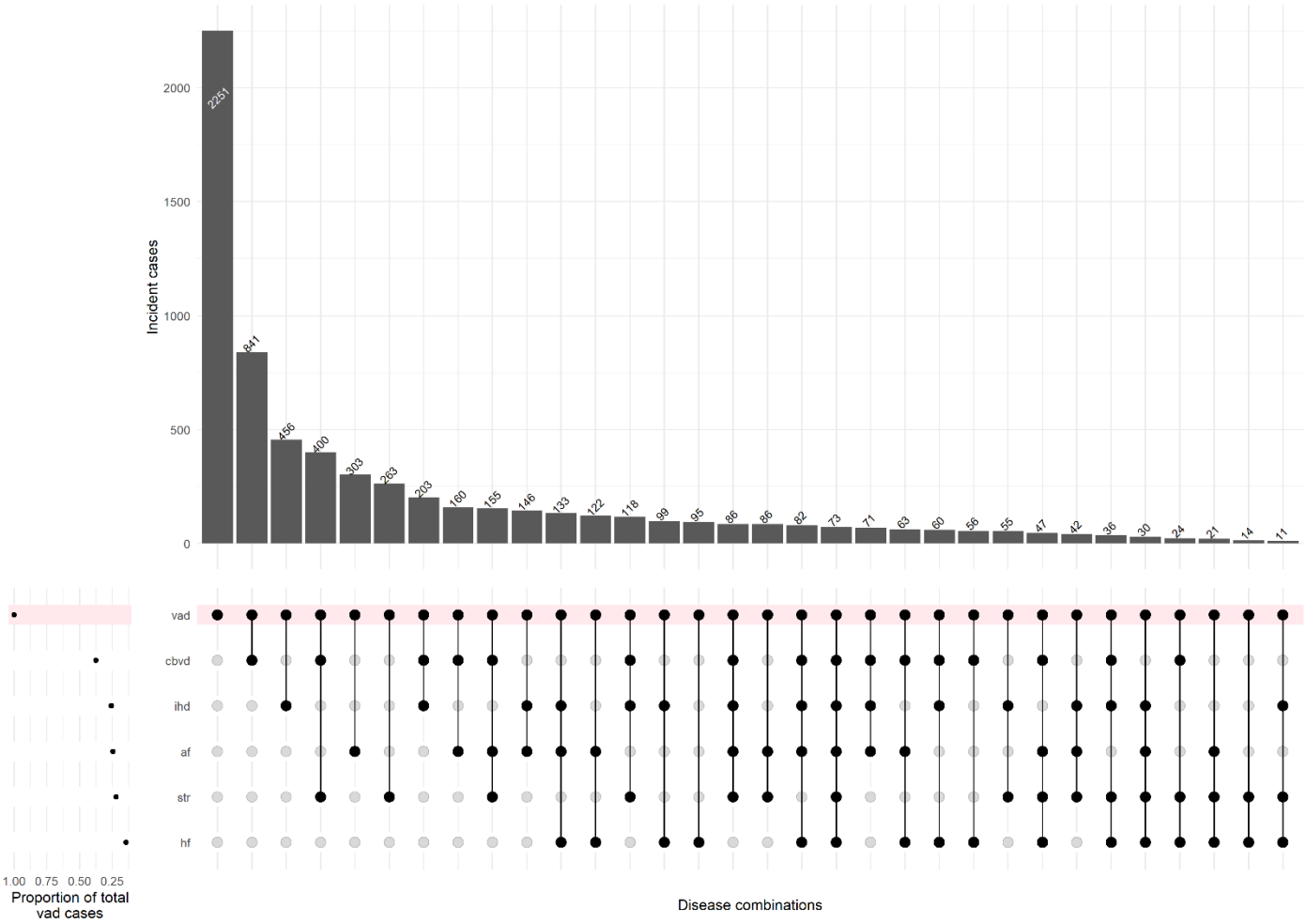


Figure C13 UpSet plot of the top five comorbidities of vascular dementia (n = 6,602)

vad=vascular dementia, cbvd=other cerebrovascular disease, ihd=ischaemic heart disease, af=atrial fibrillation, str=stroke, hf=heart failure

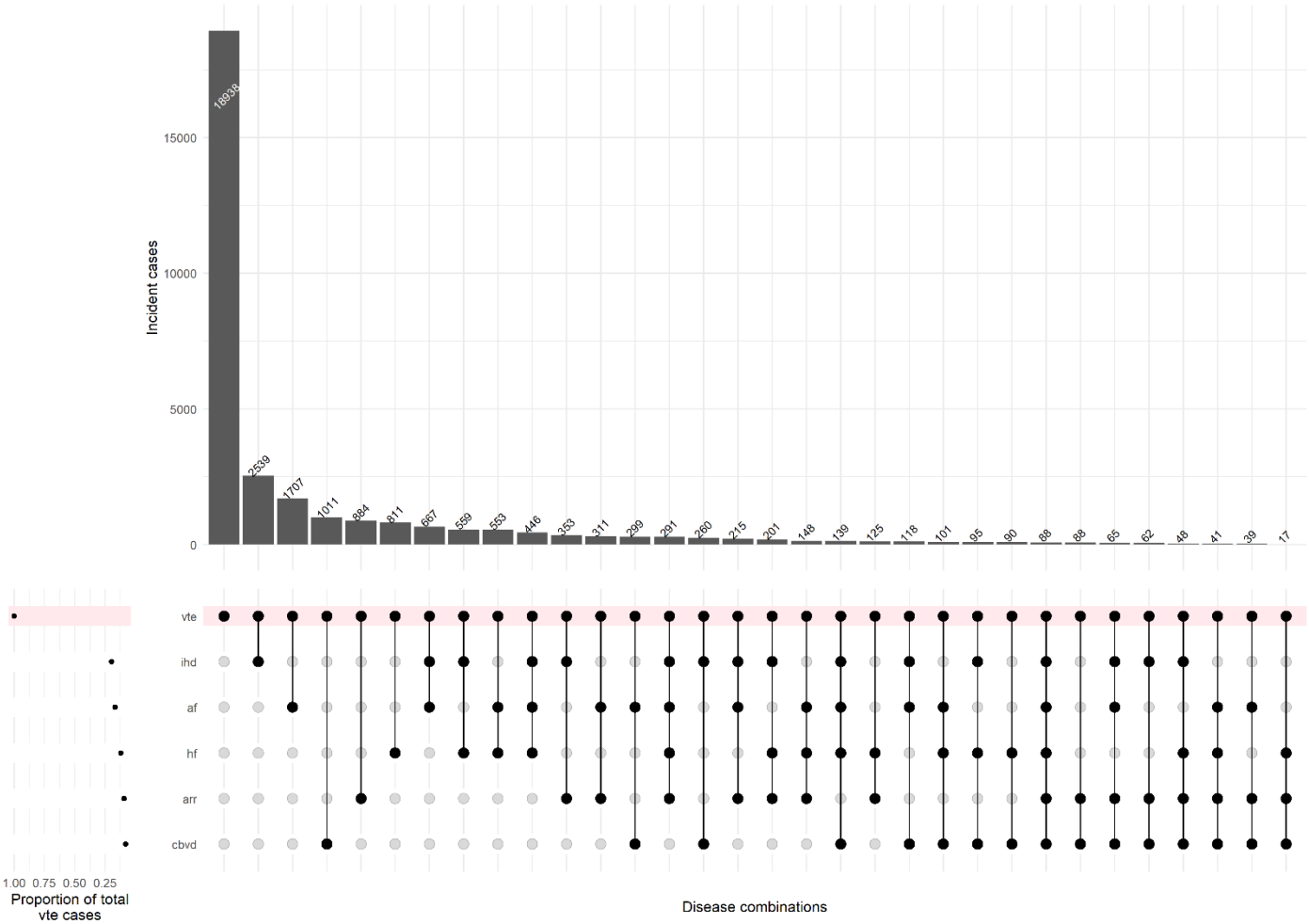


Figure C14 UpSet plot of the top five comorbidities of venous thromboembolism (n = 31,309)

vte=venous thromboembolism, ihd=ischaemic heart disease, af=atrial fibrillation, hf=heart failure, arr=other arrhythmias, cbvd=other cerebrovascular disease

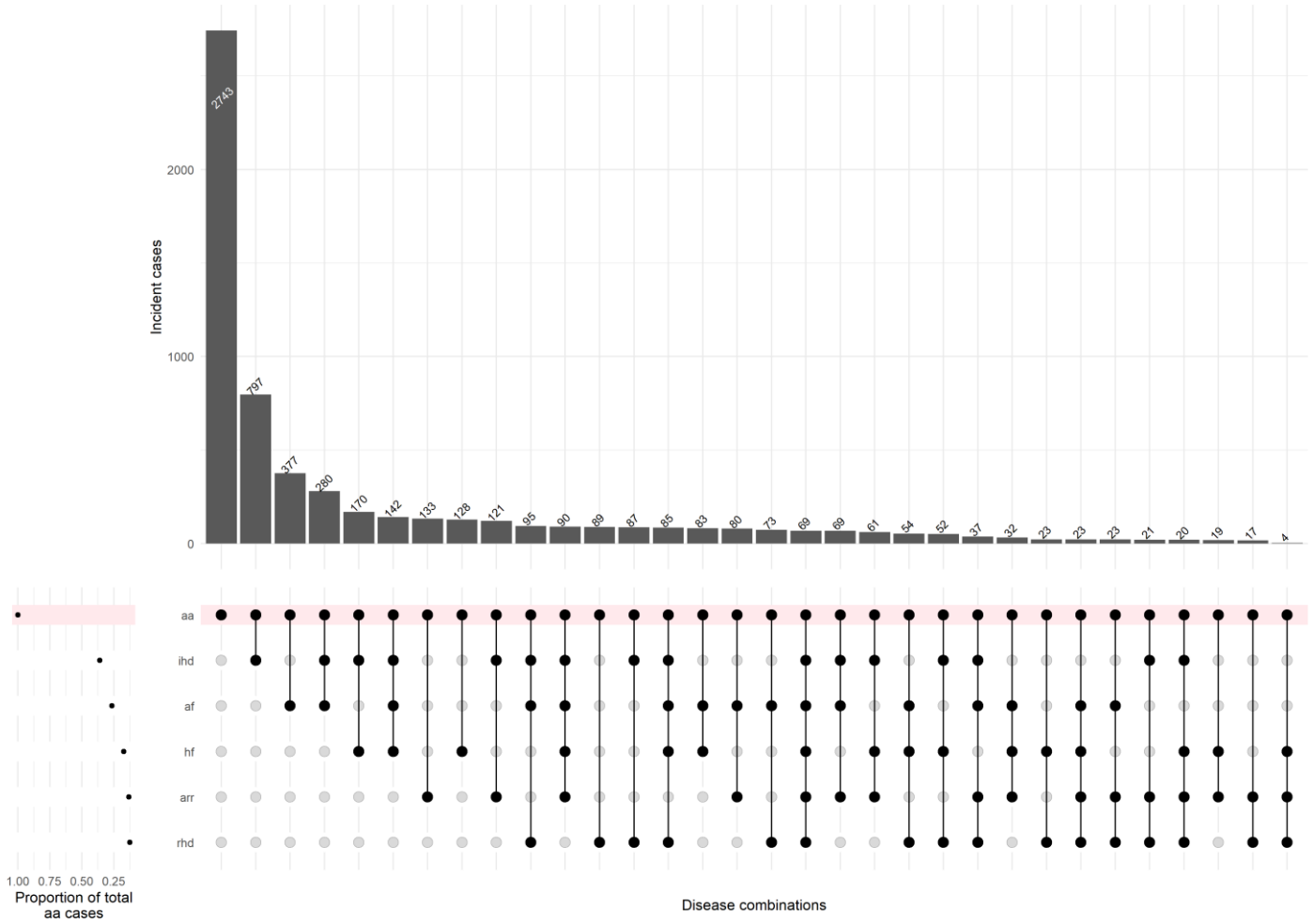


Figure C15 UpSet plot of the top five comorbidities of aortic aneurysm (n = 6,097)

aa=aortic aneurysm, ihd=ischaemic heart disease, af=atrial fibrillation, hf=heart failure, arr=other arrhythmias, rhd=rheumatic heart disease

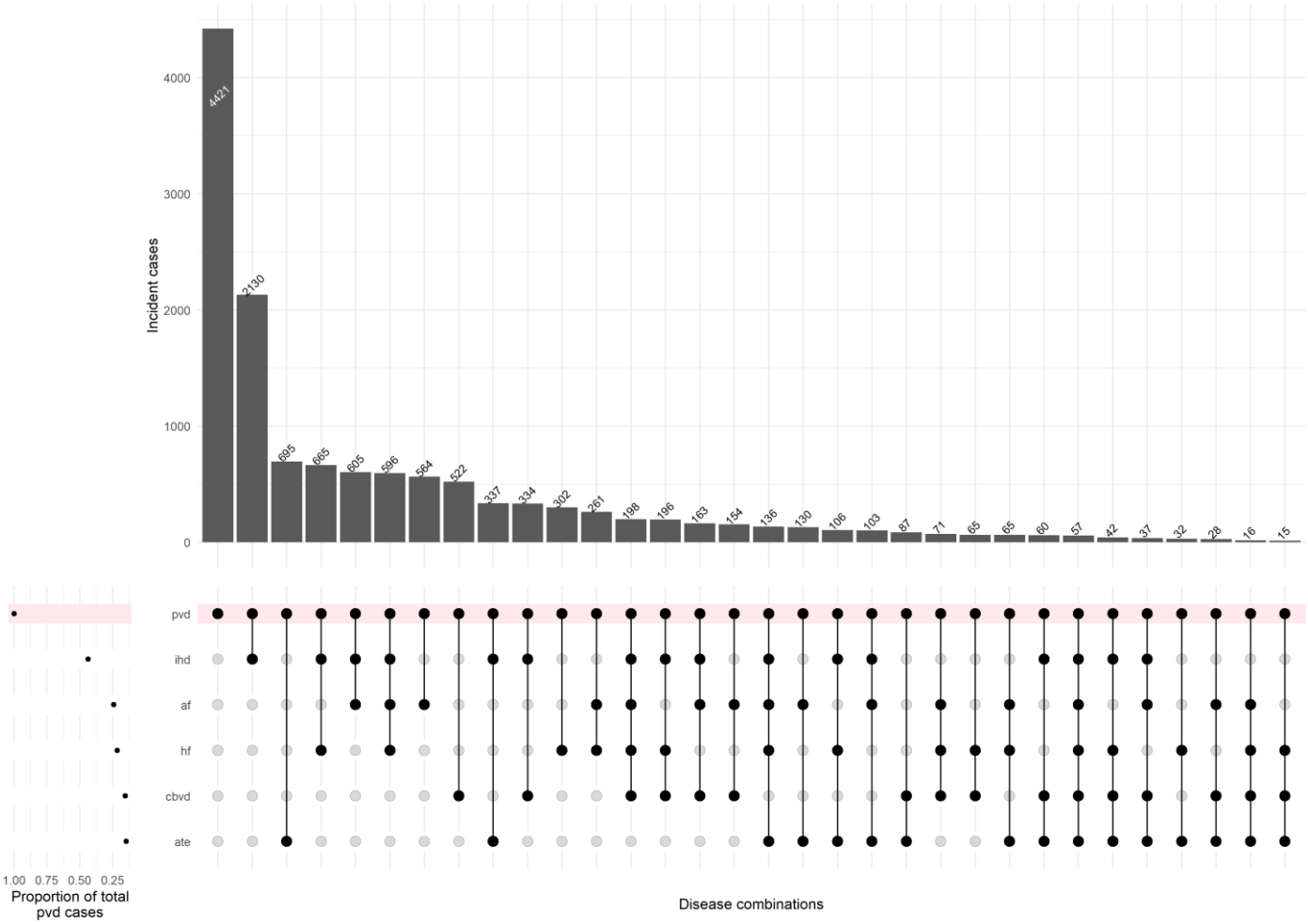


Figure C16 UpSet plot of the top five comorbidities of peripheral vascular disease (n = 13,193)

pvd=peripheral vascular disease, ihd=ischaemic heart disease, af=atrial fibrillation, hf=heart failure, cbvd=other cerebrovascular disease, ate=arterial embolism and thrombosis

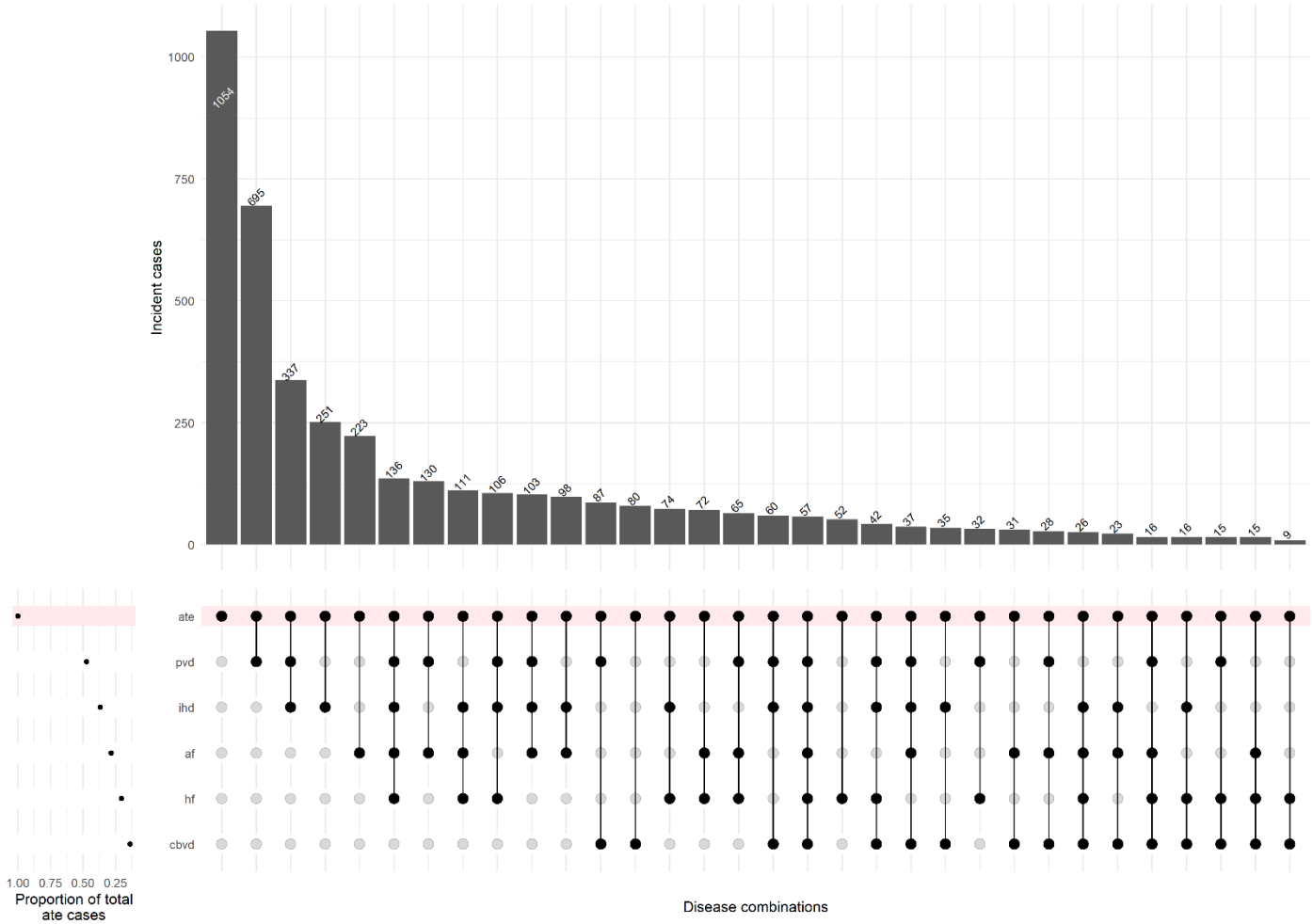


Figure C17 UpSet plot of the top five comorbidities of arterial embolism and thrombosis (n = 4,116)

ate=arterial embolism and thrombosis, pvd=peripheral vascular disease, ihd=ischaemic heart disease, af=atrial fibrillation, hf=heart failure, cbvd=other cerebrovascular disease

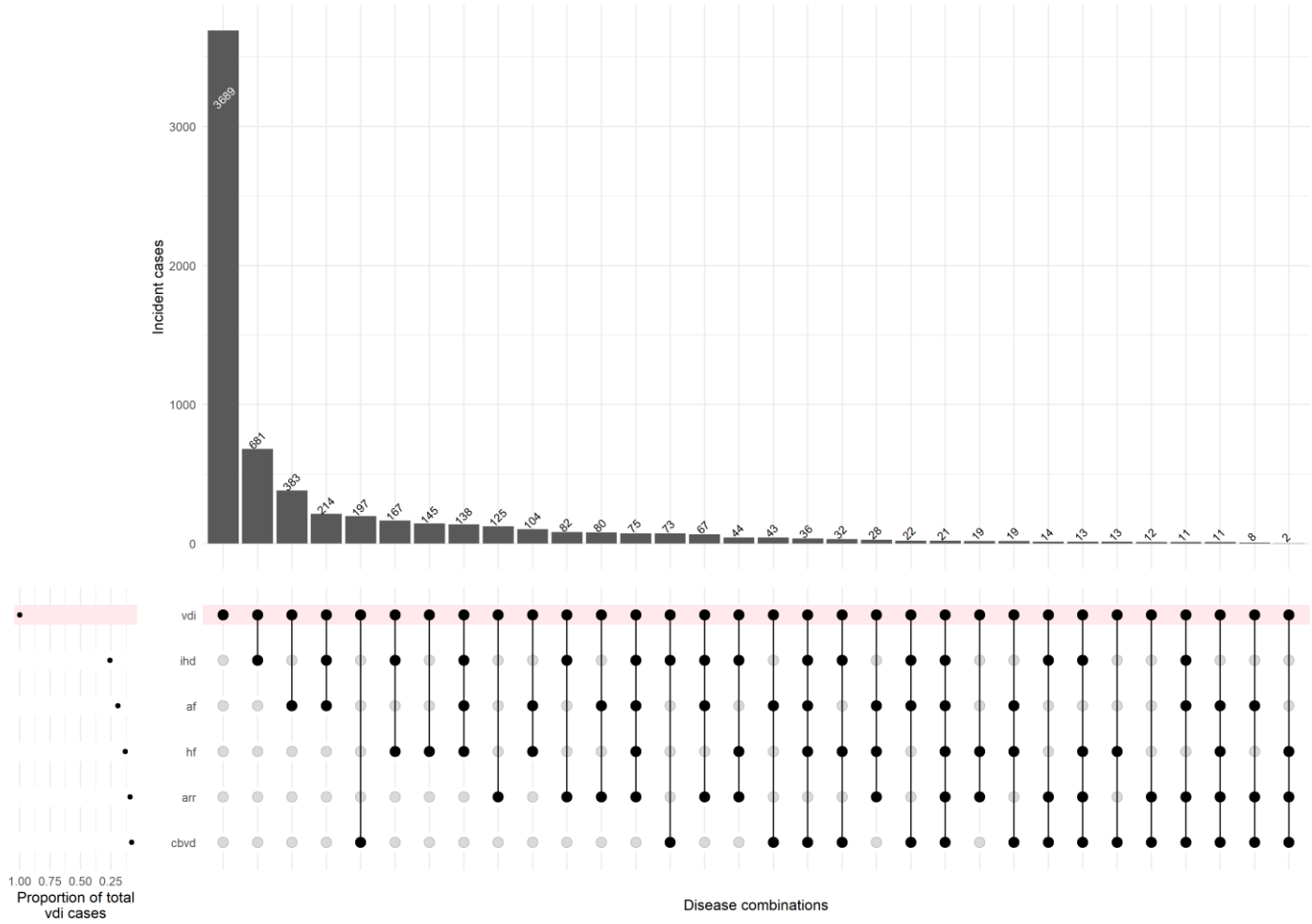


Figure C18 UpSet plot of the top five comorbidities of vascular disorders of the intestine (n = 6,568)

vdi=vascular disorders of the intestine, ihd=ischaemic heart disease, af=atrial fibrillation, hf=heart failure, arr=other arrhythmias, cbvd=other cerebrovascular disease

s

Appendix D: Established cardiovascular risk factors and cardiovascular multimorbidity

Table D1 Number and proportion of missing values for each risk factor and adjustment variable

All women, n	1,240,268
Established cardiovascular risk factors	% (n)
Smoking	
Never	48.7 (603,782)
Past	26.3 (326,064)
Current <15 per day	9.8 (121,936)
Current 15+ per day	9.4 (116,569)
Missing	5.8 (71,917)
Alcohol consumption (units/week)	
Rarely/Never	34.7 (430,441)
1-6	40.8 (505,430)
7-14	18.7 (232,001)
15+	5.1 (62,985)
Missing	0.8 (9,411)
Strenuous physical activity	
Rarely/Never	46.1 (571,571)
Up to once	29.8 (369,188)
2-3 times	14.4 (178,406)
4+ times	6.1 (75,972)
Missing	3.6 (45,131)
Body mass index (kg/m^2)	
<20	3.9 (47,864)
20-25	41.2 (511,097)
25-30	33.5 (416,089)
30-35	11.6 (143,864)
35+	4.6 (57,070)
Missing	5.2 (64,284)
Height (cm)	
<155	17.4 (216,011)
155-159	14.9 (184,383)
160-164	29.5 (366,466)
165-169	22.1 (274,369)
170+	14.5 (179,717)
Missing	1.6 (19,322)
Being treated for hypertension	
Yes	14.5 (179,842)
No	85.4 (1,059,757)
Missing	0.1 (669)
Being treated for diabetes mellitus	
Yes	2.1 (25,642)
No	97.9 (1,213,852)
Missing	0.1 (774)
Being treated for high blood cholesterol	
Yes	2.6 (32,323)

No	85.8 (1,064,276)
Not asked	11.5 (142,959)
Missing	0.1 (710)
Socio-economic factors	% (n)
Area deprivation	
Least deprived fifth	20.3 (252,383)
2	20.2 (250,544)
3	20.0 (247,443)
4	19.7 (244,832)
Most deprived fifth	19.0 (236,031)
Missing	0.7 (9,035)
Educational attainment	
1	13.2 (163,123)
2	25.9 (321,451)
3	16.4 (203,018)
4	39.7 (492,035)
5	2.3 (28,033)
Missing	2.6 (32,608)
Reproductive factors	% (n)
Age at menarche	
<12	19.8 (245,606)
12	17.5 (216,686)
13	24.0 (297,722)
14	20.2 (249,999)
15+	16.5 (204,980)
Missing	2.0 (25,275)
Parity	
Nulliparous	10.8 (134,387)
1	12.5 (155,467)
2	43.1 (534,449)
3	21.9 (271,241)
4+	11.3 (140,181)
Missing	0.4 (4,543)
Ever use of hormone replacement therapy	
Never	49.0 (608,103)
Ever	49.8 (617,968)
Missing	1.1 (14,197)
Breastfeeding duration (in parous women)	
None	25.4 (280,594)
1 month	10.5 (116,001)
2-5 months	17.5 (193,272)
6-11 months	12.5 (138,173)
12+ months	13.3 (147,111)
Missing	20.9 (230,730)

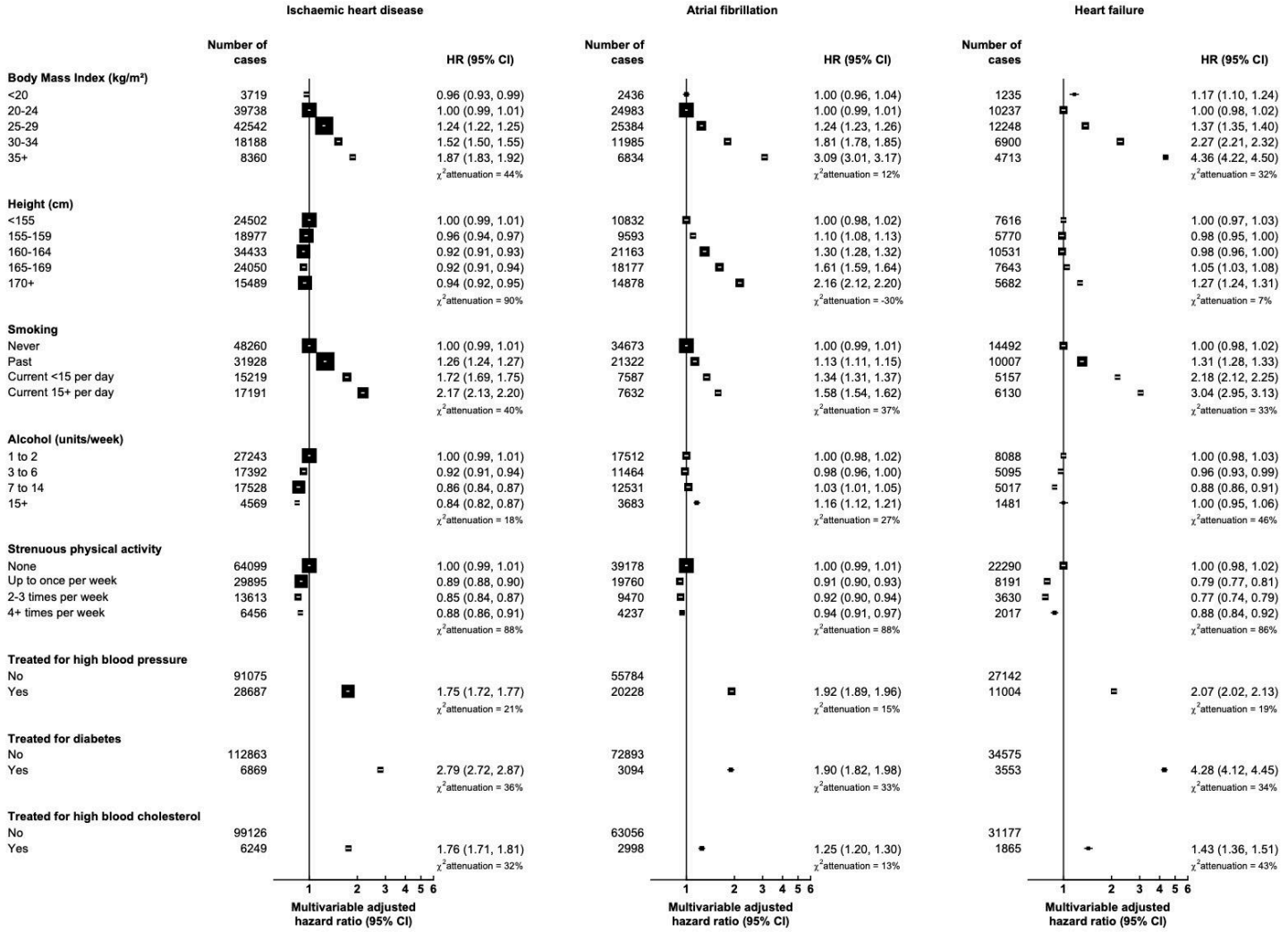


Figure D1 Relative risks of ischaemic heart disease, atrial fibrillation, and heart failure in relation to established cardiovascular risk factors

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. Models for treatment for high blood pressure, diabetes, and high blood cholesterol were not adjusted for body mass index. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed where more than two groups were compared.

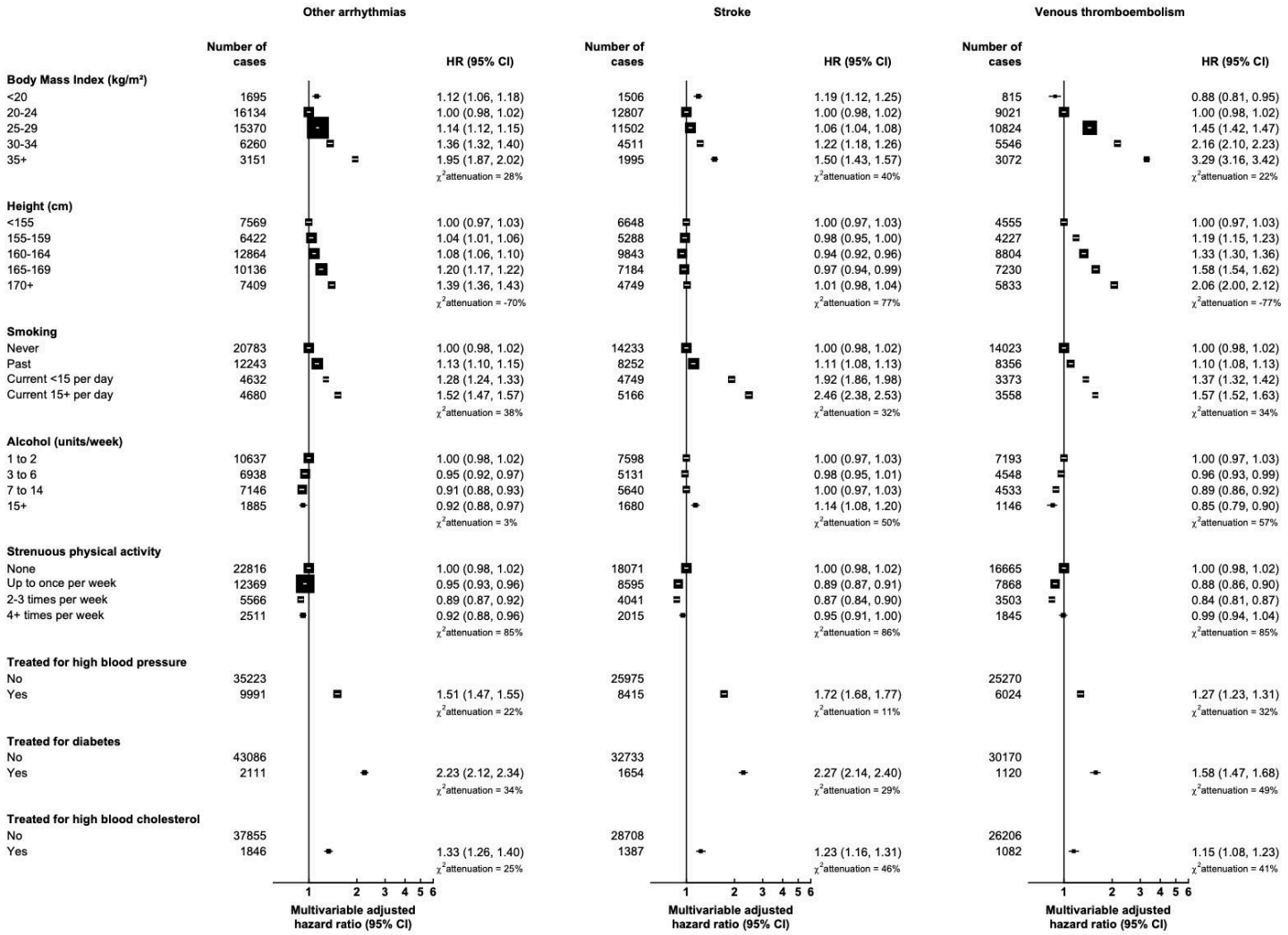


Figure D2 Relative risks of other arrhythmias, stroke, and venous thromboembolism in relation to established cardiovascular risk factors

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. Models for treatment for high blood pressure, diabetes, and high blood cholesterol were not adjusted for body mass index. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed where more than two groups were compared.

Table D2 Comparison of multivariable hazard ratios for incident IHD-AF to the averaged effects calculated from HRs for component individual CVDs

Risk factors	HR for IHD-AF	Simple average	Inverse variance weighted average
BMI (kg/m^2)			
<20	0.96 (0.89-1.05)	0.98 (0.94-1.03)	0.98 (0.94-1.02)
20-24	1.00 (ref)	1.00 (ref)	1.00 (ref)
25-29	1.31 (1.26-1.35)	1.24 (1.22-1.27)	1.24 (1.22-1.26)
30-34	1.90 (1.82-1.98)	1.67 (1.63-1.71)	1.66 (1.40-1.97)
35+	3.18 (3.02-3.34)	2.48 (2.41-2.56)	2.04 (1.47-3.93)
Height (cm)			
<155	1.00 (ref)	1.00 (ref)	1.00 (ref)
155-159	1.06 (1.01-1.11)	1.03 (1.01-1.06)	1.03 (0.90-1.17)
160-164	1.16 (1.11-1.21)	1.11 (1.09-1.14)	1.09 (0.78-1.54)
165-169	1.34 (1.28-1.40)	1.27 (1.24-1.30)	1.22 (0.70-2.11)
170+	1.63 (1.55-1.71)	1.55 (1.51-1.59)	1.43 (0.63-3.22)
Smoking			
Never	1.00 (ref)	1.00 (ref)	1.00 (ref)
Past	1.27 (1.23-1.32)	1.20 (1.18-1.22)	1.19 (1.07-1.33)
Current <15 per day	1.72 (1.65-1.80)	1.53 (1.49-1.57)	1.52 (1.19-1.94)
Current 15+ per day	2.18 (2.08-2.28)	1.88 (1.83-1.92)	1.85 (1.36-2.53)
Alcohol (units/week)			
1 to 2	1.00 (ref)	1.00 (ref)	1.00 (ref)
3 to 6	0.94 (0.90-0.98)	0.95 (0.93-0.98)	0.95 (0.89-1.01)
7 to 14	0.94 (0.89-0.98)	0.95 (0.93-0.97)	0.94 (0.79-1.12)
15+	1.03 (0.96-1.11)	1.00 (0.97-1.04)	0.99 (0.72-1.35)
Strenuous physical activity			
None	1.00 (ref)	1.00 (ref)	1.00 (ref)
Up to once per week	0.87 (0.84-0.90)	0.90 (0.89-0.92)	0.90 (0.88-0.92)
2-3 times per week	0.86 (0.83-0.90)	0.89 (0.87-0.91)	0.88 (0.82-0.96)
4+ times per week	0.95 (0.90-1.01)	0.91 (0.88-0.94)	0.91 (0.85-0.97)
Being treated for high blood pressure			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	2.22 (2.15-2.28)	1.84 (1.81-1.87)	1.83 (1.67-2.01)
Being treated for diabetes			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	2.89 (2.74-3.06)	2.35 (2.27-2.43)	2.30 (1.58-3.36)
Being treated for high blood cholesterol			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	1.67 (1.57-1.77)	1.51 (1.46-1.56)	1.48 (1.06-2.08)

IHD=ischaemic heart disease, AF=atrial fibrillation, HR=hazard ratio, CVD=cardiovascular disease, BMI=body mass index

Table D3 Comparison of multivariable hazard ratios for incident IHD-HF to the averaged effects calculated from HRs for component individual CVDs

Risk factors	HR for IHD-HF	Simple average	Inverse variance weighted average
BMI (kg/m^2)			
<20	1.03 (0.94-1.13)	1.07 (1.01-1.13)	1.06 (0.87-1.29)
20-24	1.00 (ref)	1.00 (ref)	1.00 (ref)
25-29	1.38 (1.33-1.43)	1.31 (1.28-1.34)	1.30 (1.18-1.44)
30-34	2.20 (2.10-2.30)	1.90 (1.84-1.95)	1.86 (1.25-2.75)
35+	3.81 (3.61-4.03)	3.12 (3.01-3.23)	2.89 (1.23-6.77)
Height (cm)			
<155	1.00 (ref)	1.00 (ref)	1.00 (ref)
155-159	0.96 (0.91-1.01)	0.97 (0.94-1.00)	0.97 (0.95-0.98)
160-164	0.93 (0.89-0.97)	0.95 (0.93-0.98)	0.95 (0.89-1.01)
165-169	0.97 (0.93-1.02)	0.99 (0.96-1.02)	0.98 (0.86-1.12)
170+	1.10 (1.04-1.16)	1.11 (1.07-1.14)	1.09 (0.81-1.47)
Smoking			
Never	1.00 (ref)	1.00 (ref)	1.00 (ref)
Past	1.38 (1.32-1.43)	1.29 (1.26-1.31)	1.28 (1.24-1.33)
Current <15 per day	2.47 (2.35-2.59)	1.95 (1.90-2.01)	1.94 (1.53-2.44)
Current 15+ per day	3.34 (3.19-3.50)	2.61 (2.53-2.68)	2.57 (1.85-3.57)
Alcohol (units/week)			
1 to 2	1.00 (ref)	1.00 (ref)	1.00 (ref)
3 to 6	0.94 (0.89-0.99)	0.94 (0.91-0.97)	0.94 (0.90-0.98)
7 to 14	0.82 (0.77-0.86)	0.87 (0.85-0.90)	0.86 (0.85-0.88)
15+	0.88 (0.80-0.96)	0.92 (0.88-0.97)	0.92 (0.77-1.09)
Strenuous physical activity			
None	1.00 (ref)	1.00 (ref)	1.00 (ref)
Up to once per week	0.80 (0.77-0.83)	0.84 (0.82-0.86)	0.84 (0.75-0.94)
2-3 times per week	0.76 (0.72-0.81)	0.81 (0.79-0.84)	0.81 (0.74-0.89)
4+ times per week	0.90 (0.84-0.96)	0.88 (0.85-0.92)	0.88 (0.86-0.90)
Being treated for high blood pressure			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	2.18 (2.11-2.26)	1.91 (1.87-1.95)	1.90 (1.61-2.24)
Being treated for diabetes			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	5.37 (5.12-5.64)	3.54 (3.42-3.66)	3.46 (2.27-5.25)
Being treated for high blood cholesterol			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	1.71 (1.60-1.83)	1.60 (1.54-1.66)	1.59 (1.30-1.95)

IHD=ischaemic heart disease, HF=heart failure, HR=hazard ratio, CVD=cardiovascular disease, BMI=body mass index

Table D4 Comparison of multivariable hazard ratios for incident IHD-ARR to the averaged effects calculated from HRs for component individual CVDs

Risk factors	HR for IHD-ARR	Simple average	Inverse variance weighted average
BMI (kg/m^2)			
<20	0.99 (0.90-1.09)	1.04 (0.99-1.09)	1.04 (0.89-1.21)
20-24	1.00 (ref)	1.00 (ref)	1.00 (ref)
25-29	1.23 (1.19-1.28)	1.19 (1.17-1.21)	1.19 (1.10-1.29)
30-34	1.57 (1.49-1.65)	1.44 (1.41-1.48)	1.44 (1.29-1.60)
35+	2.26 (2.12-2.41)	1.91 (1.85-1.98)	1.90 (1.83-1.98)
Height (cm)			
<155	1.00 (ref)	1.00 (ref)	1.00 (ref)
155-159	1.00 (0.94-1.05)	1.00 (0.97-1.03)	1.00 (0.92-1.08)
160-164	0.97 (0.93-1.02)	1.00 (0.98-1.03)	0.98 (0.85-1.17)
165-169	1.07 (1.01-1.13)	1.06 (1.03-1.09)	1.05 (0.81-1.36)
170+	1.20 (1.14-1.28)	1.17 (1.13-1.20)	1.14 (0.78-1.68)
Smoking			
Never	1.00 (ref)	1.00 (ref)	1.00 (ref)
Past	1.26 (1.21-1.31)	1.20 (1.17-1.22)	1.19 (1.07-1.33)
Current <15 per day	1.71 (1.62-1.80)	1.50 (1.46-1.54)	1.48 (1.11-1.98)
Current 15+ per day	2.16 (2.05-2.28)	1.85 (1.79-1.89)	1.82 (1.28-2.58)
Alcohol (units/week)			
1 to 2	1.00 (ref)	1.00 (ref)	1.00 (ref)
3 to 6	0.95 (0.90-1.00)	0.94 (0.91-0.96)	0.93 (0.90-0.96)
7 to 14	0.84 (0.80-0.89)	0.89 (0.86-0.91)	0.88 (0.84-0.93)
15+	0.83 (0.75-0.90)	0.88 (0.84-0.92)	0.88 (0.80-0.96)
Strenuous physical activity			
None	1.00 (ref)	1.00 (ref)	1.00 (ref)
Up to once per week	0.89 (0.86-0.93)	0.92 (0.90-0.94)	0.92 (0.86-0.98)
2-3 times per week	0.86 (0.81-0.90)	0.87 (0.86-0.90)	0.87 (0.83-0.91)
4+ times per week	0.91 (0.85-0.98)	0.90 (0.87-0.94)	0.90 (0.86-0.93)
Being treated for high blood pressure			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	1.87 (1.80-1.94)	1.63 (1.60-1.66)	1.63 (1.41-1.88)
Being treated for diabetes			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	3.29 (3.09-3.52)	2.51 (2.42-2.61)	2.50 (2.01-3.11)
Being treated for high blood cholesterol			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	1.69 (1.57-1.81)	1.55 (1.49-1.61)	1.53 (1.16-2.02)

IHD=ischaemic heart disease, ARR=other arrhythmias, HR=hazard ratio, CVD=cardiovascular disease, BMI=body mass index

Table D5 Comparison of multivariable hazard ratios for incident AF-HF to the averaged effects calculated from HRs for component individual CVDs

Risk factors	HR for AF-HF	Simple average	Inverse variance weighted average
BMI (kg/m^2)			
<20	1.05 (0.95-1.16)	1.09 (1.03-1.15)	1.08 (0.93-1.26)
20-24	1.00 (ref)	1.00 (ref)	1.00 (ref)
25-29	1.43 (1.37-1.49)	1.31 (1.28-1.34)	1.30 (1.18-1.44)
30-34	2.45 (2.33-2.57)	2.04 (1.98-2.10)	2.03 (1.62-2.53)
35+	5.24 (4.95-5.54)	3.73 (3.60-3.87)	3.67 (2.62-5.14)
Height (cm)			
<155	1.00 (ref)	1.00 (ref)	1.00 (ref)
155-159	1.05 (0.99-1.11)	1.04 (1.01-1.08)	1.04 (0.93-1.16)
160-164	1.19 (1.13-1.25)	1.14 (1.11-1.18)	1.13 (0.86-1.49)
165-169	1.43 (1.35-1.51)	1.33 (1.30-1.38)	1.30 (0.86-1.98)
170+	1.91 (1.81-2.02)	1.72 (1.66-1.77)	1.66 (0.98-2.79)
Smoking			
Never	1.00 (ref)	1.00 (ref)	1.00 (ref)
Past	1.24 (1.19-1.29)	1.22 (1.19-1.25)	1.22 (1.05-1.41)
Current <15 per day	1.96 (1.86-2.07)	1.76 (1.71-1.83)	1.71 (1.06-2.75)
Current 15+ per day	2.58 (2.45-2.72)	2.31 (2.24-2.39)	2.19 (1.15-4.16)
Alcohol (units/week)			
1 to 2	1.00 (ref)	1.00 (ref)	1.00 (ref)
3 to 6	0.94 (0.89-1.00)	0.97 (0.94-1.01)	0.98 (0.95-1.00)
7 to 14	0.92 (0.87-0.98)	0.96 (0.93-0.99)	0.95 (0.82-1.11)
15+	1.15 (1.06-1.25)	1.08 (1.03-1.14)	1.08 (0.93-1.25)
Strenuous physical activity			
None	1.00 (ref)	1.00 (ref)	1.00 (ref)
Up to once per week	0.82 (0.79-0.86)	0.85 (0.84-0.87)	0.85 (0.74-0.97)
2-3 times per week	0.80 (0.76-0.85)	0.85 (0.82-0.87)	0.84 (0.71-1.00)
4+ times per week	0.90 (0.83-0.97)	0.91 (0.87-0.95)	0.91 (0.86-0.97)
Being treated for high blood pressure			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	2.32 (2.24-2.40)	2.00 (1.96-2.05)	1.99 (1.85-2.15)
Being treated for diabetes			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	3.46 (3.26-3.69)	3.09 (2.97-3.22)	2.85 (1.29-6.32)
Being treated for high blood cholesterol			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	1.39 (1.28-1.50)	1.34 (1.28-1.41)	1.34 (1.17-1.52)

AF=atrial fibrillation, HF=heart failure, HR=hazard ratio, CVD=cardiovascular disease, BMI=body mass index

Table D6 Comparison of multivariable hazard ratios for incident ARR-HF to the averaged effects calculated from HRs for component individual CVDs

Risk factors	HR for ARR-HF	Simple average	Inverse variance weighted average
BMI (kg/m^2)			
<20	1.16 (1.02-1.32)	1.15 (1.08-1.22)	1.14 (1.09-1.19)
20-24	1.00 (ref)	1.00 (ref)	1.00 (ref)
25-29	1.38 (1.31-1.46)	1.26 (1.22-1.29)	1.25 (1.04-1.50)
30-34	2.08 (1.94-2.22)	1.82 (1.76-1.87)	1.76 (1.06-2.90)
35+	4.03 (3.72-4.36)	3.16 (3.03-3.29)	2.92 (1.33-6.42)
Height (cm)			
<155	1.00 (ref)	1.00 (ref)	1.00 (ref)
155-159	1.04 (0.96-1.13)	1.01 (0.97-1.05)	1.01 (0.95-1.07)
160-164	1.08 (1.01-1.16)	1.03 (1.00-1.06)	1.03 (0.94-1.13)
165-169	1.14 (1.06-1.22)	1.13 (1.09-1.17)	1.12 (0.99-1.28)
170+	1.44 (1.34-1.56)	1.33 (1.28-1.38)	1.33 (1.22-1.45)
Smoking			
Never	1.00 (ref)	1.00 (ref)	1.00 (ref)
Past	1.30 (1.23-1.37)	1.22 (1.19-1.25)	1.22 (1.05-1.41)
Current <15 per day	2.02 (1.88-2.17)	1.73 (1.68-1.80)	1.67 (0.99-2.82)
Current 15+ per day	2.69 (2.50-2.89)	2.28 (2.20-2.36)	2.15 (1.09-4.24)
Alcohol (units/week)			
1 to 2	1.00 (ref)	1.00 (ref)	1.00 (ref)
3 to 6	0.93 (0.87-1.01)	0.96 (0.92-0.99)	0.95 (0.93-0.98)
7 to 14	0.83 (0.77-0.90)	0.90 (0.87-0.93)	0.90 (087-0.93)
15+	0.86 (0.75-0.97)	0.96 (0.91-1.02)	0.96 (0.88-1.04)
Strenuous physical activity			
None	1.00 (ref)	1.00 (ref)	1.00 (ref)
Up to once per week	0.82 (0.77-0.86)	0.87 (0.85-0.89)	0.87 (0.72-1.04)
2-3 times per week	0.78 (0.73-0.84)	0.83 (0.81-0.86)	0.83 (0.72-0.96)
4+ times per week	0.89 (0.80-0.98)	0.90 (0.85-0.94)	0.90 (0.86-0.94)
Being treated for high blood pressure			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	2.03 (1.93-2.13)	1.79 (1.75-1.84)	1.77 (1.30-2.41)
Being treated for diabetes			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	4.38 (4.05-4.73)	3.26 (3.12-3.40)	3.09 (1.63-5.85)
Being treated for high blood cholesterol			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	1.48 (1.33-1.64)	1.38 (1.31-1.46)	1.38 (1.29-1.48)

ARR=other arrhythmias, HF=heart failure, HR=hazard ratio, CVD=cardiovascular disease, BMI=body mass index

Table D7 Comparison of multivariable hazard ratios for incident AF-STR to the averaged effects calculated from HRs for component individual CVDs

Risk factors	HR for AF-STR	Simple average	Inverse variance weighted average
BMI (kg/m^2)			
<20	0.94 (0.83-1.07)	1.10 (1.04-1.16)	1.09 (0.92-1.29)
20-24	1.00 (ref)	1.00 (ref)	1.00 (ref)
25-29	1.20 (1.14-1.27)	1.15 (1.13-1.18)	1.15 (0.98-1.34)
30-34	1.66 (1.55-1.78)	1.52 (1.47-1.57)	1.49 (1.01-2.19)
35+	2.57 (2.35-2.81)	2.30 (2.21-2.39)	2.15 (1.06-4.37)
Height (cm)			
<155	1.00 (ref)	1.00 (ref)	1.00 (ref)
155-159	1.02 (0.94-1.11)	1.04 (1.01-1.08)	1.04 (0.93-1.16)
160-164	1.11 (1.04-1.19)	1.12 (1.09-1.16)	1.11 (0.81-1.52)
165-169	1.31 (1.22-1.41)	1.29 (1.25-1.34)	1.25 (0.76-2.05)
170+	1.59 (1.47-1.72)	1.59 (1.54-1.64)	1.48 (0.70-3.11)
Smoking			
Never	1.00 (ref)	1.00 (ref)	1.00 (ref)
Past	1.11 (1.05-1.17)	1.12 (1.10-1.15)	1.12 (1.11-1.14)
Current <15 per day	1.60 (1.48-1.72)	1.63 (1.58-1.69)	1.60 (1.13-2.28)
Current 15+ per day	1.94 (1.79-2.10)	2.02 (1.96-2.09)	1.97 (1.28-3.04)
Alcohol (units/week)			
1 to 2	1.00 (ref)	1.00 (ref)	1.00 (ref)
3 to 6	0.98 (0.91-1.05)	0.98 (0.95-1.02)	0.98 (0.96-1.00)
7 to 14	1.04 (0.97-1.12)	1.02 (0.99-1.05)	1.02 (0.99-1.05)
15+	1.30 (1.16-1.44)	1.15 (1.10-1.21)	1.15 (1.12-1.19)
Strenuous physical activity			
None	1.00 (ref)	1.00 (ref)	1.00 (ref)
Up to once per week	0.93 (0.88-0.99)	0.90 (0.89-0.93)	0.90 (0.88-0.92)
2-3 times per week	0.92 (0.86-0.99)	0.90 (0.87-0.92)	0.90 (0.85-0.95)
4+ times per week	0.98 (0.89-1.08)	0.95 (0.90-0.99)	0.94 (0.92-0.97)
Being treated for high blood pressure			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	2.52 (2.40-2.64)	1.82 (1.79-1.87)	1.82 (1.63-2.03)
Being treated for diabetes			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	2.46 (2.23-2.72)	2.09 (1.98-2.19)	2.07 (1.74-2.47)
Being treated for high blood cholesterol			
No	1.00 (ref)	1.00 (ref)	1.00 (ref)
Yes	1.19 (1.06-1.33)	1.24 (1.18-1.31)	1.24 (1.20-1.29)

AF=atrial fibrillation, STR=stroke, HR=hazard ratio, CVD=cardiovascular disease, BMI=body mass index

Appendix E: Socioeconomic status and cardiovascular multimorbidity

Table E1 IHD-AF incidence by education after adjustment for different covariates

			Educational attainment				
	LR[1]	**attenuation in LR[1] (%)**	**Tertiary**	**Secondary**	**Technical**	**Compulsory schooling**	**< Compulsory schooling**
Number of incident IHD-AF events			2,208	5,203	3,902	12,971	1,021
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	1061	-	1.00 (-)	1.18 (1.12 - 1.24)	1.29 (1.23 - 1.36)	1.70 (1.63 - 1.78)	2.12 (1.96 - 2.28)
Age, region, deprivation	661	38	1.00 (-)	1.16 (1.11 - 1.22)	1.26 (1.20 - 1.33)	1.56 (1.49 - 1.63)	1.88 (1.74 - 2.02)
Age, region, smoking	668	37	1.00 (-)	1.15 (1.09 - 1.21)	1.23 (1.17 - 1.30)	1.55 (1.48 - 1.63)	1.90 (1.75 - 2.05)
Age, region, alcohol	835	21	1.00 (-)	1.17 (1.11 - 1.23)	1.26 (1.20 - 1.33)	1.62 (1.55 - 1.70)	1.99 (1.85 - 2.15)
Age, region, physical activity	692	35	1.00 (-)	1.15 (1.09 - 1.21)	1.25 (1.18 - 1.32)	1.57 (1.49 - 1.64)	1.93 (1.78 - 2.08)
Age, region, body mass index	657	38	1.00 (-)	1.15 (1.09 - 1.21)	1.24 (1.17 - 1.31)	1.54 (1.47 - 1.62)	1.87 (1.73 - 2.03)
Age, region, height	1151	-8	1.00 (-)	1.19 (1.13 - 1.25)	1.32 (1.26 - 1.39)	1.76 (1.68 - 1.84)	2.21 (2.05 - 2.39)
Age, region, parity	976	8	1.00 (-)	1.18 (1.12 - 1.24)	1.30 (1.23 - 1.37)	1.68 (1.61 - 1.76)	2.05 (1.90 - 2.21)
Age, region, menopausal hormone therapy	1054	1	1.00 (-)	1.17 (1.12 - 1.23)	1.29 (1.22 - 1.36)	1.70 (1.63 - 1.78)	2.11 (1.96 - 2.28)
Age, region, health behaviours[2]	207	80	1.00 (-)	1.10 (1.04 - 1.16)	1.14 (1.08 - 1.21)	1.30 (1.24 - 1.37)	1.55 (1.42 - 1.68)
Age, region, health behaviours[2], plus deprivation	142	87	1.00 (-)	1.09 (1.03 - 1.15)	1.13 (1.07 - 1.20)	1.25 (1.19 - 1.32)	1.47 (1.35 - 1.60)
Fully adjusted[3]	186	83	1.00 (-)	1.10 (1.04 - 1.16)	1.16 (1.10 - 1.23)	1.30 (1.24 - 1.37)	1.54 (1.41 - 1.67)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, deprivation, height, parity, and ever use of menopausal hormone therapy;
IHD=ischaemic heart disease; AF=atrial fibrillation

Table E2 IHD-AF incidence by area deprivation after adjustment for different covariates

			Townsend index of area deprivation (fifths)				
	LR[1]	**attenuation in LR[1] (%)**	**1 (Least deprived)**	**2**	**3**	**4**	**5 (Most deprived)**
Number of incident IHD-AF events			4,198	4,594	5,033	5,528	6,719
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	1102	-	1.00 (-)	1.07 (1.03 - 1.12)	1.20 (1.15 - 1.25)	1.35 (1.30 - 1.41)	1.78 (1.71 - 1.85)
Age, region, education	653	41	1.00 (-)	1.05 (1.00 - 1.09)	1.16 (1.11 - 1.21)	1.26 (1.21 - 1.31)	1.59 (1.52 - 1.65)
Age, region, smoking	638	42	1.00 (-)	1.04 (1.00 - 1.09)	1.16 (1.11 - 1.21)	1.28 (1.22 - 1.33)	1.57 (1.51 - 1.64)
Age, region, alcohol	932	15	1.00 (-)	1.06 (1.02 - 1.11)	1.18 (1.13 - 1.23)	1.31 (1.26 - 1.37)	1.71 (1.64 - 1.77)
Age, region, physical activity	825	25	1.00 (-)	1.07 (1.02 - 1.11)	1.18 (1.13 - 1.23)	1.31 (1.26 - 1.37)	1.68 (1.61 - 1.75)
Age, region, body mass index	687	38	1.00 (-)	1.05 (1.00 - 1.09)	1.15 (1.10 - 1.20)	1.26 (1.21 - 1.32)	1.60 (1.53 - 1.66)
Age, region, height	1125	-2	1.00 (-)	1.07 (1.02 - 1.12)	1.20 (1.15 - 1.25)	1.36 (1.30 - 1.42)	1.80 (1.73 - 1.87)
Age, region, parity	965	12	1.00 (-)	1.07 (1.02 - 1.11)	1.19 (1.14 - 1.24)	1.33 (1.28 - 1.39)	1.72 (1.65 - 1.79)
Age, region, menopausal hormone therapy	1094	1	1.00 (-)	1.07 (1.03 - 1.12)	1.20 (1.15 - 1.25)	1.35 (1.30 - 1.41)	1.78 (1.72 - 1.85)
Age, region, health behaviours[2]	249	77	1.00 (-)	1.01 (0.97 - 1.06)	1.09 (1.04 - 1.14)	1.16 (1.11 - 1.21)	1.35 (1.29 - 1.41)
Age, region, health behaviours[2], plus education	172	84	1.00 (-)	1.01 (0.96 - 1.05)	1.07 (1.03 - 1.12)	1.12 (1.07 - 1.17)	1.29 (1.23 - 1.35)
Fully adjusted[3]	170	85	1.00 (-)	1.01 (0.96 - 1.05)	1.07 (1.03 - 1.12)	1.13 (1.08 - 1.18)	1.29 (1.23 - 1.35)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, education, height, parity, and ever use of menopausal hormone therapy;
IHD=ischaemic heart disease; AF=atrial fibrillation

Table E3 IHD-HF incidence by education after adjustment for different covariates

			Educational attainment				
	LR[1]	**attenuation in LR[1] (%)**	**Tertiary**	**Secondary**	**Technical**	**Compulsory schooling**	**< Compulsory schooling**
Number of incident IHD-HF events			1,406	3,651	2,946	10,896	908
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	1819	-	1.00 (-)	1.31 (1.23 - 1.39)	1.54 (1.45 - 1.64)	2.27 (2.15 - 2.40)	3.06 (2.81 - 3.33)
Age, region, deprivation	1095	40	1.00 (-)	1.28 (1.20 - 1.36)	1.48 (1.39 - 1.58)	1.98 (1.87 - 2.09)	2.53 (2.32 - 2.75)
Age, region, smoking	1044	43	1.00 (-)	1.25 (1.17 - 1.33)	1.44 (1.35 - 1.53)	1.92 (1.82 - 2.04)	2.51 (2.30 - 2.74)
Age, region, alcohol	1287	29	1.00 (-)	1.28 (1.20 - 1.36)	1.46 (1.37 - 1.56)	2.05 (1.93 - 2.16)	2.66 (2.44 - 2.89)
Age, region, physical activity	1160	36	1.00 (-)	1.26 (1.18 - 1.34)	1.44 (1.35 - 1.54)	1.99 (1.88 - 2.11)	2.62 (2.40 - 2.85)
Age, region, body mass index	1146	37	1.00 (-)	1.27 (1.20 - 1.36)	1.47 (1.38 - 1.57)	2.01 (1.90 - 2.13)	2.56 (2.34 - 2.80)
Age, region, height	1650	9	1.00 (-)	1.31 (1.23 - 1.40)	1.54 (1.44 - 1.64)	2.24 (2.12 - 2.37)	2.96 (2.72 - 3.23)
Age, region, parity	1685	7	1.00 (-)	1.31 (1.23 - 1.39)	1.55 (1.45 - 1.65)	2.24 (2.12 - 2.37)	2.93 (2.70 - 3.19)
Age, region, menopausal hormone therapy	1748	4	1.00 (-)	1.30 (1.23 - 1.39)	1.53 (1.44 - 1.63)	2.26 (2.13 - 2.39)	3.03 (2.78 - 3.30)
Age, region, health behaviours[2]	265	85	1.00 (-)	1.16 (1.08 - 1.24)	1.26 (1.18 - 1.35)	1.45 (1.37 - 1.55)	1.74 (1.58 - 1.91)
Age, region, health behaviours[2], plus deprivation	174	90	1.00 (-)	1.15 (1.07 - 1.23)	1.24 (1.16 - 1.33)	1.38 (1.29 - 1.46)	1.61 (1.46 - 1.77)
Fully adjusted[3]	160	91	1.00 (-)	1.15 (1.08 - 1.23)	1.24 (1.16 - 1.33)	1.37 (1.29 - 1.46)	1.58 (1.44 - 1.75)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, deprivation, height, parity, and ever use of menopausal hormone therapy;
IHD=ischaemic heart disease; HF=heart failure

Table E4 IHD-HF incidence by area deprivation after adjustment for different covariates

			Townsend index of area deprivation (fifths)				
	LR[1]	**attenuation in LR[1] (%)**	**1 (Least deprived)**	**2**	**3**	**4**	**5 (Most deprived)**
Number of incident IHD-HF events			2,845	3,237	3,726	4,467	6,199
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	2091	-	1.00 (-)	1.12 (1.07 - 1.18)	1.33 (1.26 - 1.39)	1.63 (1.55 - 1.71)	2.42 (2.31 - 2.53)
Age, region, education	1259	40	1.00 (-)	1.09 (1.03 - 1.15)	1.25 (1.19 - 1.32)	1.46 (1.39 - 1.53)	2.04 (1.95 - 2.14)
Age, region, smoking	1124	46	1.00 (-)	1.09 (1.04 - 1.15)	1.25 (1.18 - 1.31)	1.46 (1.39 - 1.53)	1.97 (1.88 - 2.06)
Age, region, alcohol	1678	20	1.00 (-)	1.11 (1.05 - 1.16)	1.29 (1.22 - 1.35)	1.54 (1.47 - 1.62)	2.22 (2.12 - 2.33)
Age, region, physical activity	1547	26	1.00 (-)	1.11 (1.05 - 1.17)	1.28 (1.22 - 1.35)	1.54 (1.47 - 1.62)	2.20 (2.10 - 2.30)
Age, region, body mass index	1289	38	1.00 (-)	1.10 (1.04 - 1.16)	1.25 (1.19 - 1.32)	1.49 (1.42 - 1.56)	2.06 (1.97 - 2.16)
Age, region, height	1899	9	1.00 (-)	1.12 (1.07 - 1.18)	1.31 (1.25 - 1.37)	1.61 (1.53 - 1.69)	2.35 (2.25 - 2.46)
Age, region, parity	1874	10	1.00 (-)	1.12 (1.07 - 1.18)	1.31 (1.25 - 1.38)	1.60 (1.53 - 1.68)	2.33 (2.22 - 2.43)
Age, region, menopausal hormone therapy	2032	3	1.00 (-)	1.12 (1.07 - 1.18)	1.32 (1.25 - 1.38)	1.62 (1.54 - 1.69)	2.41 (2.30 - 2.52)
Age, region, health behaviours[2]	381	82	1.00 (-)	1.05 (0.99 - 1.10)	1.13 (1.07 - 1.19)	1.26 (1.19 - 1.32)	1.53 (1.45 - 1.60)
Age, region, health behaviours[2], plus education	273	87	1.00 (-)	1.04 (0.98 - 1.09)	1.12 (1.06 - 1.18)	1.21 (1.15 - 1.27)	1.45 (1.37 - 1.52)
Fully adjusted[3]	243	88	1.00 (-)	1.03 (0.98 - 1.09)	1.11 (1.05 - 1.17)	1.19 (1.13 - 1.26)	1.42 (1.35 - 1.50)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, education, height, parity, and ever use of menopausal hormone therapy;
IHD=ischaemic heart disease; HF=heart failure

Table E5 IHD-ARR incidence by education after adjustment for different covariates

	LR[1]	attenuation in LR[1] (%)	Educational attainment				
			Tertiary	**Secondary**	**Technical**	**Compulsory schooling**	**< Compulsory schooling**
Number of incident IHD-ARR events			1,589	3,568	2,817	8,827	703
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	716	-	1.00 (-)	1.12 (1.05 - 1.18)	1.32 (1.24 - 1.40)	1.64 (1.56 - 1.73)	2.13 (1.95 - 2.33)
Age, region, deprivation	445	38	1.00 (-)	1.10 (1.04 - 1.17)	1.28 (1.21 - 1.37)	1.50 (1.42 - 1.59)	1.89 (1.73 - 2.07)
Age, region, smoking	443	38	1.00 (-)	1.10 (1.03 - 1.16)	1.27 (1.19 - 1.35)	1.50 (1.42 - 1.58)	1.91 (1.74 - 2.10)
Age, region, alcohol	522	27	1.00 (-)	1.10 (1.03 - 1.16)	1.27 (1.20 - 1.35)	1.54 (1.46 - 1.62)	1.95 (1.78 - 2.14)
Age, region, physical activity	483	33	1.00 (-)	1.10 (1.03 - 1.16)	1.28 (1.20 - 1.36)	1.53 (1.45 - 1.62)	1.93 (1.76 - 2.12)
Age, region, body mass index	483	33	1.00 (-)	1.09 (1.02 - 1.16)	1.27 (1.20 - 1.36)	1.52 (1.44 - 1.60)	1.91 (1.74 - 2.09)
Age, region, height	692	3	1.00 (-)	1.12 (1.05 - 1.18)	1.32 (1.24 - 1.40)	1.64 (1.56 - 1.74)	2.12 (1.94 - 2.32)
Age, region, parity	636	11	1.00 (-)	1.11 (1.04 - 1.18)	1.31 (1.23 - 1.40)	1.60 (1.52 - 1.69)	2.04 (1.86 - 2.23)
Age, region, menopausal hormone therapy	729	-2	1.00 (-)	1.12 (1.05 - 1.19)	1.32 (1.24 - 1.40)	1.65 (1.57 - 1.74)	2.16 (1.98 - 2.37)
Age, region, health behaviours[2]	142	80	1.00 (-)	1.04 (0.98 - 1.11)	1.17 (1.09 - 1.25)	1.26 (1.19 - 1.34)	1.53 (1.38 - 1.69)
Age, region, health behaviours[2], plus deprivation	98	86	1.00 (-)	1.03 (0.97 - 1.10)	1.15 (1.08 - 1.23)	1.21 (1.14 - 1.29)	1.45 (1.31 - 1.60)
Fully adjusted[3]	100	86	1.00 (-)	1.03 (0.97 - 1.10)	1.15 (1.08 - 1.23)	1.22 (1.15 - 1.29)	1.45 (1.30 - 1.60)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, deprivation, height, parity, and ever use of menopausal hormone therapy;
IHD=ischaemic heart disease; ARR=other arrhythmias

Table E6 IHD-ARR incidence by area deprivation after adjustment for different covariates

			Townsend index of area deprivation (fifths)				
	LR[1]	**attenuation in LR[1] (%)**	**1 (Least deprived)**	**2**	**3**	**4**	**5 (Most deprived)**
Number of incident IHD-ARR events			2,847	3,209	3,517	3,803	4,637
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	757	-	1.00 (-)	1.10 (1.05 - 1.16)	1.24 (1.18 - 1.30)	1.38 (1.31 - 1.44)	1.80 (1.72 - 1.89)
Age, region, education	459	39	1.00 (-)	1.08 (1.02 - 1.13)	1.20 (1.14 - 1.26)	1.28 (1.22 - 1.34)	1.61 (1.54 - 1.70)
Age, region, smoking	424	44	1.00 (-)	1.07 (1.02 - 1.13)	1.19 (1.13 - 1.26)	1.28 (1.21 - 1.34)	1.58 (1.51 - 1.66)
Age, region, alcohol	613	19	1.00 (-)	1.09 (1.03 - 1.14)	1.21 (1.15 - 1.27)	1.33 (1.27 - 1.40)	1.71 (1.63 - 1.79)
Age, region, physical activity	578	24	1.00 (-)	1.09 (1.03 - 1.15)	1.22 (1.16 - 1.28)	1.33 (1.27 - 1.40)	1.70 (1.62 - 1.79)
Age, region, body mass index	507	33	1.00 (-)	1.09 (1.03 - 1.15)	1.19 (1.13 - 1.25)	1.30 (1.23 - 1.37)	1.65 (1.57 - 1.73)
Age, region, height	718	5	1.00 (-)	1.10 (1.05 - 1.16)	1.23 (1.17 - 1.30)	1.37 (1.30 - 1.44)	1.79 (1.71 - 1.88)
Age, region, parity	652	14	1.00 (-)	1.10 (1.04 - 1.15)	1.22 (1.16 - 1.29)	1.35 (1.29 - 1.42)	1.73 (1.65 - 1.82)
Age, region, menopausal hormone therapy	770	-2	1.00 (-)	1.10 (1.05 - 1.16)	1.24 (1.18 - 1.30)	1.37 (1.31 - 1.44)	1.82 (1.73 - 1.91)
Age, region, health behaviours[2]	176	77	1.00 (-)	1.04 (0.98 - 1.10)	1.13 (1.07 - 1.19)	1.16 (1.10 - 1.22)	1.37 (1.30 - 1.45)
Age, region, health behaviours[2], plus education	128	83	1.00 (-)	1.03 (0.98 - 1.09)	1.12 (1.06 - 1.18)	1.13 (1.07 - 1.19)	1.32 (1.25 - 1.39)
Fully adjusted[3]	121	84	1.00 (-)	1.03 (0.97 - 1.09)	1.11 (1.05 - 1.17)	1.12 (1.06 - 1.18)	1.31 (1.24 - 1.39)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, education, height, parity, and ever use of menopausal hormone therapy;
IHD=ischaemic heart disease; ARR=other arrhythmias

Table E7 AF-HF incidence by education after adjustment for different covariates

			Educational attainment				
	LR[1]	**attenuation in LR[1] (%)**	**Tertiary**	**Secondary**	**Technical**	**Compulsory schooling**	**< Compulsory schooling**
Number of incident AF-HF events			1,421	3,369	2,522	9,317	732
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	1024	-	1.00 (-)	1.19 (1.12 - 1.26)	1.29 (1.21 - 1.37)	1.86 (1.76 - 1.97)	2.28 (2.08 - 2.49)
Age, region, deprivation	622	39	1.00 (-)	1.17 (1.10 - 1.25)	1.25 (1.17 - 1.33)	1.67 (1.58 - 1.77)	1.97 (1.80 - 2.16)
Age, region, smoking	646	37	1.00 (-)	1.14 (1.07 - 1.22)	1.22 (1.14 - 1.30)	1.66 (1.56 - 1.75)	2.01 (1.83 - 2.20)
Age, region, alcohol	799	22	1.00 (-)	1.17 (1.10 - 1.25)	1.26 (1.18 - 1.34)	1.76 (1.66 - 1.86)	2.11 (1.92 - 2.31)
Age, region, physical activity	625	39	1.00 (-)	1.15 (1.08 - 1.23)	1.23 (1.15 - 1.32)	1.66 (1.56 - 1.75)	2.01 (1.84 - 2.21)
Age, region, body mass index	560	45	1.00 (-)	1.15 (1.08 - 1.23)	1.22 (1.14 - 1.30)	1.62 (1.53 - 1.72)	1.91 (1.74 - 2.09)
Age, region, height	1124	-10	1.00 (-)	1.21 (1.13 - 1.28)	1.32 (1.24 - 1.41)	1.94 (1.84 - 2.06)	2.42 (2.21 - 2.65)
Age, region, parity	980	4	1.00 (-)	1.20 (1.13 - 1.28)	1.30 (1.22 - 1.39)	1.86 (1.76 - 1.97)	2.24 (2.05 - 2.46)
Age, region, menopausal hormone therapy	958	7	1.00 (-)	1.19 (1.11 - 1.26)	1.28 (1.20 - 1.37)	1.84 (1.74 - 1.94)	2.23 (2.04 - 2.44)
Age, region, health behaviours[2]	163	84	1.00 (-)	1.08 (1.01 - 1.16)	1.12 (1.04 - 1.20)	1.32 (1.24 - 1.40)	1.51 (1.37 - 1.67)
Age, region, health behaviours[2], plus deprivation	108	89	1.00 (-)	1.08 (1.01 - 1.15)	1.10 (1.03 - 1.18)	1.27 (1.19 - 1.35)	1.43 (1.29 - 1.58)
Fully adjusted[3]	147	86	1.00 (-)	1.11 (1.04 - 1.19)	1.16 (1.08 - 1.24)	1.34 (1.26 - 1.43)	1.53 (1.38 - 1.69)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, deprivation, height, parity, and ever use of menopausal hormone therapy;
AF=atrial fibrillation; HF=heart failure

Table E8 AF-HF incidence by area deprivation after adjustment for different covariates

			Townsend index of area deprivation (fifths)				
	LR[1]	**attenuation in LR[1] (%)**	**1 (Least deprived)**	**2**	**3**	**4**	**5 (Most deprived)**
Number of incident AF-HF events			4,198	4,594	5,033	5,528	6,719
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	1115	-	1.00 (-)	1.09 (1.03 - 1.14)	1.24 (1.18 - 1.31)	1.49 (1.41 - 1.56)	1.99 (1.90 - 2.09)
Age, region, education	670	40	1.00 (-)	1.05 (1.00 - 1.11)	1.18 (1.12 - 1.24)	1.36 (1.29 - 1.43)	1.74 (1.65 - 1.83)
Age, region, smoking	641	43	1.00 (-)	1.05 (1.00 - 1.11)	1.18 (1.12 - 1.25)	1.37 (1.30 - 1.44)	1.72 (1.64 - 1.81)
Age, region, alcohol	950	15	1.00 (-)	1.08 (1.02 - 1.13)	1.22 (1.16 - 1.28)	1.44 (1.37 - 1.51)	1.90 (1.81 - 2.00)
Age, region, physical activity	818	27	1.00 (-)	1.08 (1.02 - 1.14)	1.22 (1.15 - 1.28)	1.42 (1.35 - 1.50)	1.85 (1.76 - 1.94)
Age, region, body mass index	601	46	1.00 (-)	1.05 (1.00 - 1.11)	1.16 (1.10 - 1.22)	1.35 (1.28 - 1.42)	1.68 (1.60 - 1.77)
Age, region, height	1130	-1	1.00 (-)	1.09 (1.03 - 1.14)	1.24 (1.18 - 1.31)	1.50 (1.42 - 1.57)	2.02 (1.93 - 2.12)
Age, region, parity	1006	10	1.00 (-)	1.08 (1.03 - 1.14)	1.23 (1.17 - 1.30)	1.47 (1.40 - 1.54)	1.93 (1.84 - 2.03)
Age, region, menopausal hormone therapy	1074	4	1.00 (-)	1.08 (1.03 - 1.14)	1.23 (1.17 - 1.30)	1.47 (1.40 - 1.55)	1.98 (1.89 - 2.08)
Age, region, health behaviours[2]	204	82	1.00 (-)	1.01 (0.96 - 1.07)	1.08 (1.02 - 1.14)	1.20 (1.14 - 1.26)	1.38 (1.31 - 1.45)
Age, region, health behaviours[2], plus education	141	87	1.00 (-)	1.00 (0.95 - 1.06)	1.06 (1.01 - 1.12)	1.16 (1.10 - 1.23)	1.31 (1.24 - 1.38)
Fully adjusted[3]	140	87	1.00 (-)	1.00 (0.95 - 1.06)	1.06 (1.00 - 1.12)	1.16 (1.10 - 1.23)	1.31 (1.24 - 1.39)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, education, height, parity, and ever use of menopausal hormone therapy;
AF=atrial fibrillation; HF=heart failure

Table E9 HF-ARR incidence by education after adjustment for different covariates

			Educational attainment				
	LR[1]	**attenuation in LR[1] (%)**	**Tertiary**	**Secondary**	**Technical**	**Compulsory schooling**	**< Compulsory schooling**
Number of incident HF-ARR events			786	1,806	1,398	5,004	407
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	577	-	1.00 (-)	1.14 (1.05 - 1.24)	1.30 (1.20 - 1.42)	1.83 (1.70 - 1.98)	2.38 (2.11 - 2.68)
Age, region, deprivation	366	37	1.00 (-)	1.13 (1.04 - 1.23)	1.27 (1.16 - 1.39)	1.66 (1.54 - 1.80)	2.07 (1.83 - 2.34)
Age, region, smoking	350	39	1.00 (-)	1.12 (1.03 - 1.22)	1.25 (1.14 - 1.37)	1.64 (1.52 - 1.78)	2.06 (1.81 - 2.33)
Age, region, alcohol	425	26	1.00 (-)	1.12 (1.03 - 1.22)	1.25 (1.15 - 1.37)	1.70 (1.57 - 1.84)	2.14 (1.90 - 2.42)
Age, region, physical activity	367	36	1.00 (-)	1.11 (1.02 - 1.21)	1.24 (1.13 - 1.35)	1.65 (1.53 - 1.78)	2.09 (1.84 - 2.37)
Age, region, body mass index	336	42	1.00 (-)	1.11 (1.02 - 1.21)	1.23 (1.13 - 1.35)	1.62 (1.50 - 1.75)	1.99 (1.75 - 2.26)
Age, region, height	570	1	1.00 (-)	1.15 (1.05 - 1.25)	1.31 (1.20 - 1.44)	1.85 (1.71 - 2.00)	2.42 (2.14 - 2.73)
Age, region, parity	533	8	1.00 (-)	1.14 (1.05 - 1.24)	1.30 (1.19 - 1.42)	1.81 (1.68 - 1.95)	2.31 (2.04 - 2.60)
Age, region, menopausal hormone therapy	550	5	1.00 (-)	1.14 (1.05 - 1.24)	1.31 (1.20 - 1.43)	1.83 (1.69 - 1.97)	2.35 (2.08 - 2.66)
Age, region, health behaviours[2]	88	85	1.00 (-)	1.05 (0.96 - 1.15)	1.11 (1.01 - 1.22)	1.30 (1.19 - 1.41)	1.49 (1.30 - 1.71)
Age, region, health behaviours[2], plus deprivation	63	89	1.00 (-)	1.04 (0.96 - 1.14)	1.10 (1.00 - 1.21)	1.25 (1.15 - 1.36)	1.42 (1.23 - 1.63)
Fully adjusted[3]	66	89	1.00 (-)	1.05 (0.96 - 1.15)	1.12 (1.02 - 1.23)	1.28 (1.17 - 1.39)	1.43 (1.24 - 1.65)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, deprivation, height, parity, and ever use of menopausal hormone therapy;
HF=heart failure; ARR=other arrhythmias

Table E10 HF-ARR incidence by area deprivation after adjustment for different covariates

			Townsend index of area deprivation (fifths)				
	LR[1]	**attenuation in LR[1] (%)**	**1 (Least deprived)**	**2**	**3**	**4**	**5 (Most deprived)**
Number of incident HF-ARR events			4,198	4,594	5,033	5,528	6,719
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	561	-	1.00 (-)	1.11 (1.03 - 1.19)	1.31 (1.23 - 1.41)	1.46 (1.37 - 1.57)	1.99 (1.86 - 2.12)
Age, region, education	322	43	1.00 (-)	1.07 (1.00 - 1.15)	1.25 (1.16 - 1.34)	1.33 (1.24 - 1.43)	1.72 (1.61 - 1.84)
Age, region, smoking	303	46	1.00 (-)	1.07 (0.99 - 1.15)	1.24 (1.16 - 1.33)	1.33 (1.24 - 1.42)	1.69 (1.58 - 1.81)
Age, region, alcohol	451	20	1.00 (-)	1.09 (1.01 - 1.17)	1.28 (1.19 - 1.37)	1.40 (1.31 - 1.50)	1.86 (1.74 - 1.99)
Age, region, physical activity	401	29	1.00 (-)	1.08 (1.01 - 1.16)	1.27 (1.19 - 1.37)	1.39 (1.30 - 1.49)	1.82 (1.70 - 1.95)
Age, region, body mass index	313	44	1.00 (-)	1.08 (1.00 - 1.16)	1.25 (1.16 - 1.34)	1.33 (1.24 - 1.43)	1.71 (1.60 - 1.83)
Age, region, height	533	5	1.00 (-)	1.10 (1.03 - 1.19)	1.31 (1.22 - 1.40)	1.46 (1.36 - 1.56)	1.97 (1.85 - 2.11)
Age, region, parity	487	13	1.00 (-)	1.10 (1.03 - 1.18)	1.30 (1.21 - 1.39)	1.44 (1.35 - 1.54)	1.91 (1.79 - 2.04)
Age, region, menopausal hormone therapy	550	2	1.00 (-)	1.10 (1.02 - 1.18)	1.30 (1.21 - 1.39)	1.43 (1.34 - 1.53)	1.98 (1.86 - 2.12)
Age, region, health behaviours[2]	92	84	1.00 (-)	1.01 (0.94 - 1.10)	1.15 (1.07 - 1.24)	1.15 (1.07 - 1.24)	1.35 (1.26 - 1.46)
Age, region, health behaviours[2], plus education	63	89	1.00 (-)	1.01 (0.93 - 1.09)	1.13 (1.05 - 1.22)	1.11 (1.03 - 1.20)	1.29 (1.20 - 1.39)
Fully adjusted[3]	58	90	1.00 (-)	1.00 (0.93 - 1.09)	1.12 (1.04 - 1.20)	1.09 (1.01 - 1.18)	1.28 (1.19 - 1.38)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, education, height, parity, and ever use of menopausal hormone therapy;
HF=heart failure; ARR=other arrhythmias

Table E11 AF-STR incidence by education after adjustment for different covariates

	LR[1]	attenuation in LR[1] (%)	Educational attainment				
			Tertiary	**Secondary**	**Technical**	**Compulsory schooling**	**< Compulsory schooling**
Number of incident AF-STR events			903	1,923	1,494	4,560	338
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	205	-	1.00 (-)	1.07 (0.99 - 1.16)	1.19 (1.10 - 1.30)	1.44 (1.34 - 1.54)	1.65 (1.46 - 1.88)
Age, region, deprivation	123	40	1.00 (-)	1.07 (0.98 - 1.15)	1.18 (1.08 - 1.28)	1.35 (1.25 - 1.45)	1.51 (1.33 - 1.71)
Age, region, smoking	129	37	1.00 (-)	1.07 (0.99 - 1.16)	1.18 (1.09 - 1.29)	1.35 (1.26 - 1.46)	1.58 (1.39 - 1.80)
Age, region, alcohol	187	9	1.00 (-)	1.07 (0.99 - 1.16)	1.19 (1.10 - 1.30)	1.43 (1.33 - 1.54)	1.63 (1.44 - 1.85)
Age, region, physical activity	146	29	1.00 (-)	1.05 (0.96 - 1.13)	1.16 (1.06 - 1.26)	1.36 (1.26 - 1.47)	1.56 (1.37 - 1.78)
Age, region, body mass index	127	38	1.00 (-)	1.06 (0.98 - 1.15)	1.16 (1.06 - 1.26)	1.35 (1.25 - 1.45)	1.51 (1.32 - 1.72)
Age, region, height	245	-19	1.00 (-)	1.10 (1.01 - 1.19)	1.23 (1.14 - 1.34)	1.51 (1.40 - 1.62)	1.76 (1.55 - 2.00)
Age, region, parity	200	2	1.00 (-)	1.08 (1.00 - 1.17)	1.21 (1.11 - 1.31)	1.45 (1.35 - 1.55)	1.65 (1.45 - 1.87)
Age, region, menopausal hormone therapy	192	6	1.00 (-)	1.07 (0.99 - 1.16)	1.18 (1.08 - 1.28)	1.42 (1.32 - 1.53)	1.64 (1.44 - 1.86)
Age, region, health behaviours[2]	59	71	1.00 (-)	1.05 (0.97 - 1.15)	1.14 (1.04 - 1.25)	1.25 (1.16 - 1.35)	1.42 (1.24 - 1.64)
Age, region, health behaviours[2], plus deprivation	40	80	1.00 (-)	1.05 (0.97 - 1.14)	1.14 (1.04 - 1.24)	1.21 (1.12 - 1.31)	1.35 (1.18 - 1.56)
Fully adjusted[3]	55	73	1.00 (-)	1.07 (0.98 - 1.17)	1.17 (1.07 - 1.28)	1.26 (1.16 - 1.37)	1.44 (1.24 - 1.66)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, deprivation, height, parity, and ever use of menopausal hormone therapy;
AF=atrial fibrillation; STR=stroke

Table E12 AF-STR incidence by area deprivation after adjustment for different covariates

			Townsend index of area deprivation (fifths)				
	LR[1]	**attenuation in LR[1] (%)**	**1 (Least deprived)**	**2**	**3**	**4**	**5 (Most deprived)**
Number of incident AF-STR events			1,663	1,729	1,798	2,032	2,292
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	257	-	1.00 (-)	1.02 (0.95 - 1.09)	1.08 (1.01 - 1.15)	1.25 (1.17 - 1.33)	1.55 (1.45 - 1.65)
Age, region, education	156	39	1.00 (-)	1.00 (0.93 - 1.07)	1.05 (0.98 - 1.12)	1.19 (1.11 - 1.27)	1.41 (1.32 - 1.51)
Age, region, smoking	152	41	1.00 (-)	0.98 (0.91 - 1.05)	1.03 (0.97 - 1.11)	1.17 (1.10 - 1.25)	1.39 (1.30 - 1.49)
Age, region, alcohol	235	8	1.00 (-)	1.01 (0.95 - 1.08)	1.08 (1.01 - 1.15)	1.23 (1.16 - 1.32)	1.52 (1.43 - 1.63)
Age, region, physical activity	201	22	1.00 (-)	1.01 (0.95 - 1.08)	1.07 (1.00 - 1.15)	1.22 (1.14 - 1.30)	1.49 (1.40 - 1.59)
Age, region, body mass index	176	32	1.00 (-)	1.00 (0.93 - 1.07)	1.03 (0.96 - 1.10)	1.18 (1.10 - 1.26)	1.44 (1.34 - 1.53)
Age, region, height	271	-5	1.00 (-)	1.02 (0.95 - 1.09)	1.08 (1.01 - 1.15)	1.26 (1.18 - 1.34)	1.57 (1.47 - 1.68)
Age, region, parity	237	8	1.00 (-)	1.01 (0.95 - 1.08)	1.08 (1.01 - 1.15)	1.24 (1.16 - 1.32)	1.52 (1.43 - 1.62)
Age, region, menopausal hormone therapy	249	3	1.00 (-)	1.02 (0.95 - 1.09)	1.07 (1.00 - 1.15)	1.25 (1.17 - 1.33)	1.54 (1.45 - 1.64)
Age, region, health behaviours[2]	81	69	1.00 (-)	0.97 (0.90 - 1.04)	1.00 (0.93 - 1.07)	1.10 (1.02 - 1.18)	1.28 (1.19 - 1.37)
Age, region, health behaviours[2], plus education	60	77	1.00 (-)	0.96 (0.89 - 1.03)	0.98 (0.91 - 1.06)	1.08 (1.00 - 1.16)	1.23 (1.15 - 1.33)
Fully adjusted[3]	60	76	1.00 (-)	0.96 (0.89 - 1.03)	0.97 (0.90 - 1.05)	1.08 (1.00 - 1.16)	1.23 (1.15 - 1.33)

[1]Likelihood-ratio test statistic;
[2]smoking, alcohol consumption, physical activity, body mass index;
[3]adjusted for all health behaviours, education, height, parity, and ever use of menopausal hormone therapy;
AF=atrial fibrillation; STR=stroke

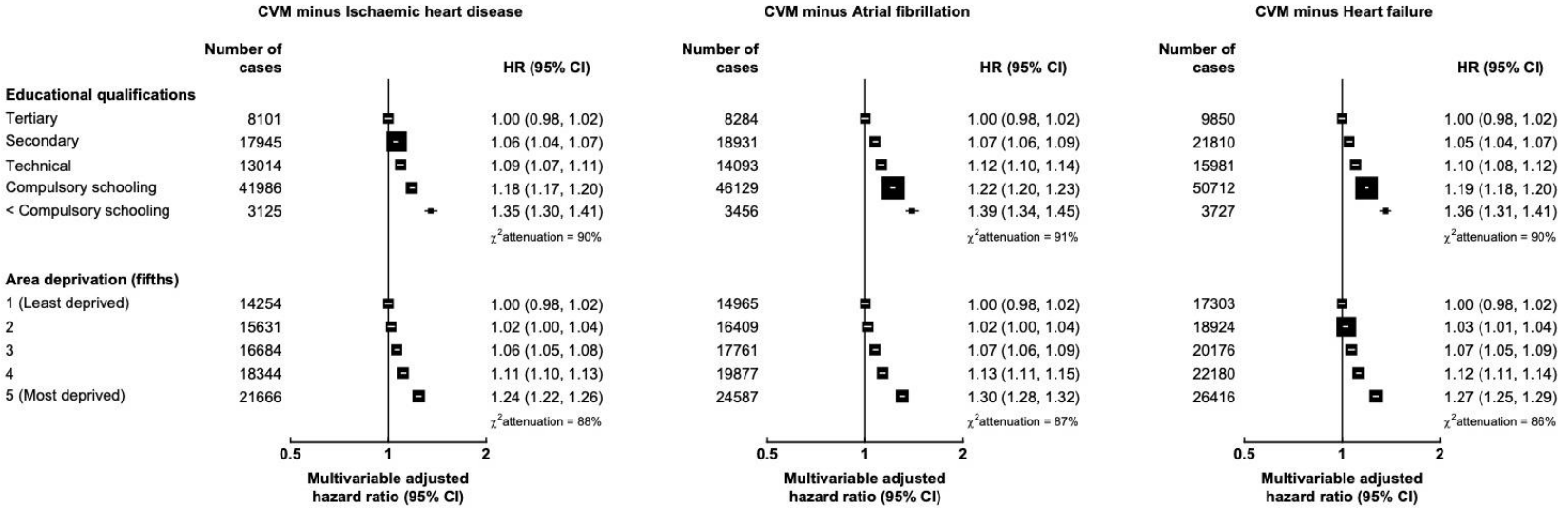


Figure E1 Relative risks of CVM minus ischaemic heart disease, atrial fibrillation, or heart failure in relation to socioeconomic factors

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation or educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed.

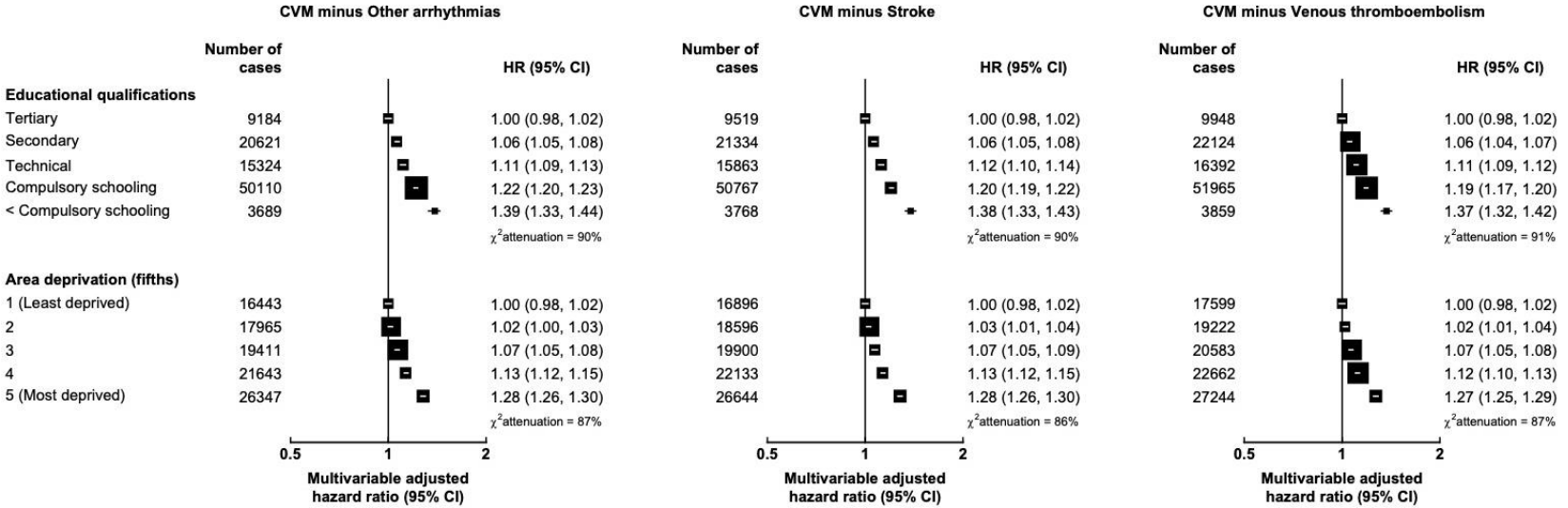


Figure E2 Relative risks of CVM minus other arrhythmias, stroke, or venous thromboembolism in relation to socioeconomic factors

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation or educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed.

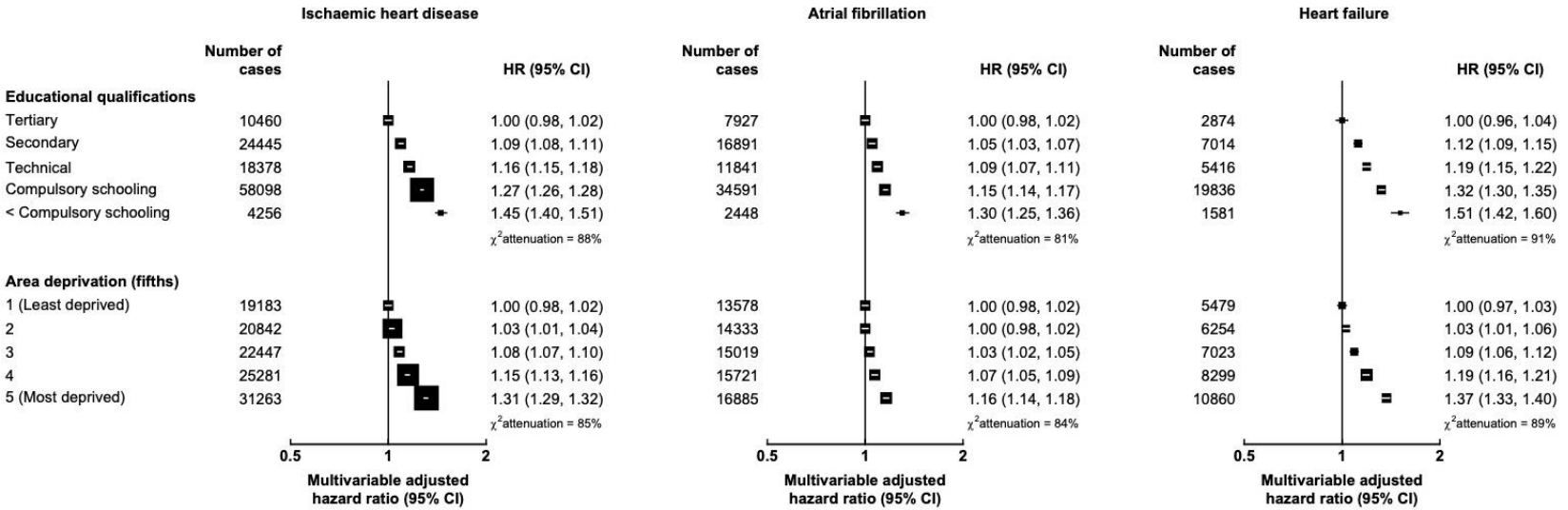


Figure E3 Relative risks of ischaemic heart disease, atrial fibrillation, and heart failure in relation to socioeconomic factors

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation or educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed.

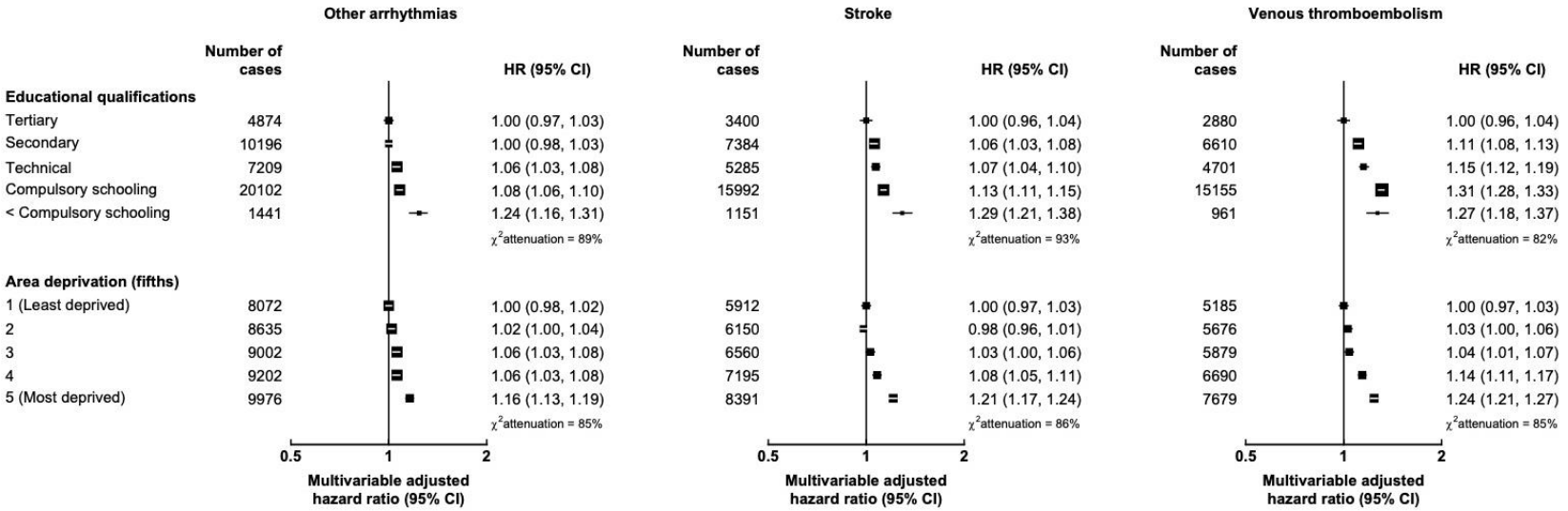


Figure E4 Relative risks of other arrhythmias, stroke, and venous thromboembolism in relation to socioeconomic factors

Minimally adjusted models were stratified by age at recruitment and area of residence only. Multivariable adjusted models were additionally adjusted for area deprivation or educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy. The size of the square boxes was inversely proportional to the standard error of the effect estimate. $\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models. Floated confidence intervals were displayed.

Table E13 CVM and complex CVM incidence by education, excluding women who were being treated for hypertension or diabetes at baseline

Education	LR[1]	attenuation in LR[1] (%)	Tertiary	Secondary	Technical	Compulsory schooling	< Compulsory schooling
INCIDENT CVM (n)			7,880	17,135	12,256	38,480	2,734
HR (95% CI), adjusted for,							
Age, region (minimally adjusted)	2712	-	1.00 (-)	1.12 (1.09 - 1.15)	1.22 (1.19 - 1.26)	1.58 (1.55 - 1.62)	1.89 (1.81 - 1.98)
Age, region, deprivation	1640	40	1.00 (-)	1.11 (1.08 - 1.14)	1.20 (1.17 - 1.24)	1.46 (1.42 - 1.50)	1.70 (1.62 - 1.77)
Age, region, smoking	1302	52	1.00 (-)	1.08 (1.05 - 1.11)	1.16 (1.12 - 1.19)	1.39 (1.35 - 1.42)	1.63 (1.56 - 1.71)
Age, region, alcohol	2136	21	1.00 (-)	1.11 (1.08 - 1.14)	1.20 (1.17 - 1.24)	1.52 (1.48 - 1.56)	1.79 (1.71 - 1.87)
Age, region, physical activity	1775	35	1.00 (-)	1.09 (1.07 - 1.12)	1.18 (1.15 - 1.22)	1.47 (1.43 - 1.51)	1.74 (1.66 - 1.82)
Age, region, body mass index	1934	29	1.00 (-)	1.10 (1.07 - 1.13)	1.20 (1.17 - 1.23)	1.49 (1.46 - 1.53)	1.77 (1.70 - 1.86)
Age, region, height	2748	-1	1.00 (-)	1.12 (1.09 - 1.16)	1.24 (1.21 - 1.28)	1.61 (1.57 - 1.65)	1.93 (1.85 - 2.02)
Age, region, parity	2507	8	1.00 (-)	1.12 (1.09 - 1.15)	1.23 (1.19 - 1.26)	1.57 (1.53 - 1.61)	1.85 (1.77 - 1.93)
Age, region, menopausal hormone therapy	2689	1	1.00 (-)	1.12 (1.09 - 1.15)	1.22 (1.18 - 1.25)	1.58 (1.54 - 1.62)	1.90 (1.82 - 1.99)
Age, region, health behaviours[2]	412	85	1.00 (-)	1.05 (1.02 - 1.08)	1.09 (1.06 - 1.13)	1.21 (1.18 - 1.25)	1.40 (1.33 - 1.47)
Age, region, health behaviours[2], plus deprivation	268	90	1.00 (-)	1.05 (1.02 - 1.08)	1.09 (1.06 - 1.12)	1.18 (1.14 - 1.21)	1.34 (1.27 - 1.41)
Fully adjusted[3]	317	88	1.00 (-)	1.06 (1.03 - 1.09)	1.11 (1.08 - 1.14)	1.20 (1.17 - 1.24)	1.37 (1.30 - 1.44)
INCIDENT COMPLEX CVM (n)			1,517	3,352	2,473	8,626	673
HR (95% CI), adjusted for,							
Age, region (minimally adjusted)	898	-	1.00 (-)	1.13 (1.06 - 1.20)	1.24 (1.17 - 1.33)	1.75 (1.66 - 1.85)	2.20 (2.01 - 2.41)
Age, region, deprivation	557	38	1.00 (-)	1.11 (1.05 - 1.18)	1.22 (1.14 - 1.30)	1.59 (1.50 - 1.68)	1.93 (1.76 - 2.11)
Age, region, smoking	431	52	1.00 (-)	1.08 (1.02 - 1.15)	1.16 (1.09 - 1.24)	1.48 (1.40 - 1.57)	1.82 (1.66 - 2.01)
Age, region, alcohol	707	21	1.00 (-)	1.12 (1.05 - 1.19)	1.22 (1.14 - 1.30)	1.67 (1.58 - 1.76)	2.05 (1.87 - 2.25)
Age, region, physical activity	610	32	1.00 (-)	1.10 (1.04 - 1.17)	1.19 (1.12 - 1.27)	1.61 (1.52 - 1.70)	1.99 (1.81 - 2.19)
Age, region, body mass index	653	27	1.00 (-)	1.11 (1.04 - 1.18)	1.21 (1.13 - 1.29)	1.63 (1.54 - 1.73)	2.04 (1.85 - 2.24)
Age, region, height	935	-4	1.00 (-)	1.14 (1.07 - 1.21)	1.26 (1.19 - 1.35)	1.79 (1.70 - 1.89)	2.29 (2.09 - 2.52)
Age, region, parity	845	6	1.00 (-)	1.13 (1.07 - 1.20)	1.25 (1.18 - 1.34)	1.74 (1.65 - 1.84)	2.15 (1.96 - 2.36)
Age, region, menopausal hormone therapy	872	3	1.00 (-)	1.13 (1.06 - 1.20)	1.24 (1.16 - 1.32)	1.74 (1.65 - 1.84)	2.19 (2.00 - 2.40)
Age, region, health behaviours[2]	155	83	1.00 (-)	1.04 (0.98 - 1.11)	1.09 (1.01 - 1.16)	1.27 (1.20 - 1.35)	1.53 (1.38 - 1.69)
Age, region, health behaviours[2], plus deprivation	108	88	1.00 (-)	1.03 (0.97 - 1.10)	1.08 (1.01 - 1.15)	1.22 (1.15 - 1.30)	1.46 (1.31 - 1.61)
Fully adjusted[3]	127	86	1.00 (-)	1.05 (0.98 - 1.12)	1.10 (1.03 - 1.18)	1.27 (1.19 - 1.35)	1.50 (1.35 - 1.67)

[1]Likelihood-ratio test statistic; [2]smoking, alcohol consumption, physical activity, body mass index; [3]adjusted for all health behaviours, deprivation, height, parity, and ever use of menopausal hormone therapy; CVM=cardiovascular multimorbidity

Table E14 CVM and complex CVM incidence by area deprivation, excluding women who were being treated for hypertension or diabetes at baseline

Area deprivation (fifths of Townsend index)	LR[1]	attenuation in LR[1] (%)	1 (Least deprived)	2	3	4	5 (Most deprived)
INCIDENT CVM (n)			13,667	14,666	15,592	16,975	19,773
HR (95% CI), adjusted for,							
Age, region (minimally adjusted)	3103	-	1.00 (-)	1.07 (1.05 - 1.10)	1.18 (1.16 - 1.21)	1.34 (1.31 - 1.37)	1.74 (1.70 - 1.78)
Age, region, education	1887	39	1.00 (-)	1.05 (1.03 - 1.08)	1.14 (1.12 - 1.17)	1.26 (1.23 - 1.29)	1.57 (1.53 - 1.60)
Age, region, smoking	1398	55	1.00 (-)	1.04 (1.02 - 1.07)	1.13 (1.10 - 1.16)	1.23 (1.21 - 1.26)	1.47 (1.44 - 1.50)
Age, region, alcohol	2624	15	1.00 (-)	1.07 (1.04 - 1.09)	1.17 (1.14 - 1.20)	1.31 (1.28 - 1.34)	1.67 (1.64 - 1.71)
Age, region, physical activity	2339	25	1.00 (-)	1.07 (1.04 - 1.09)	1.17 (1.14 - 1.19)	1.30 (1.27 - 1.33)	1.64 (1.61 - 1.68)
Age, region, body mass index	2253	27	1.00 (-)	1.07 (1.04 - 1.09)	1.16 (1.13 - 1.19)	1.29 (1.26 - 1.33)	1.63 (1.60 - 1.67)
Age, region, height	3026	2	1.00 (-)	1.08 (1.05 - 1.10)	1.19 (1.16 - 1.21)	1.35 (1.32 - 1.38)	1.74 (1.70 - 1.78)
Age, region, parity	2773	11	1.00 (-)	1.07 (1.05 - 1.10)	1.18 (1.15 - 1.20)	1.33 (1.30 - 1.36)	1.69 (1.66 - 1.73)
Age, region, menopausal hormone therapy	3067	1	1.00 (-)	1.07 (1.05 - 1.10)	1.19 (1.16 - 1.21)	1.35 (1.32 - 1.38)	1.74 (1.70 - 1.78)
Age, region, health behaviours[2]	623	80	1.00 (-)	1.03 (1.00 - 1.06)	1.09 (1.06 - 1.11)	1.16 (1.13 - 1.18)	1.32 (1.29 - 1.35)
Age, region, health behaviours[2], plus education	453	85	1.00 (-)	1.02 (1.00 - 1.05)	1.07 (1.05 - 1.10)	1.13 (1.10 - 1.16)	1.27 (1.24 - 1.30)
Fully adjusted[3]	420	86	1.00 (-)	1.02 (0.99 - 1.05)	1.07 (1.05 - 1.10)	1.13 (1.10 - 1.16)	1.26 (1.23 - 1.30)
INCIDENT COMPLEX CVM (n)			2,774	2,979	3,336	3,605	4,498
HR (95% CI), adjusted for,							
Age, region (minimally adjusted)	964	-	1.00 (-)	1.06 (1.00 - 1.11)	1.23 (1.17 - 1.30)	1.39 (1.32 - 1.46)	1.92 (1.83 - 2.02)
Age, region, education	563	42	1.00 (-)	1.03 (0.97 - 1.08)	1.17 (1.12 - 1.24)	1.27 (1.21 - 1.34)	1.68 (1.60 - 1.76)
Age, region, smoking	412	57	1.00 (-)	1.02 (0.97 - 1.08)	1.16 (1.10 - 1.22)	1.25 (1.18 - 1.31)	1.55 (1.48 - 1.63)
Age, region, alcohol	818	15	1.00 (-)	1.05 (1.00 - 1.11)	1.21 (1.15 - 1.27)	1.35 (1.28 - 1.42)	1.84 (1.75 - 1.93)
Age, region, physical activity	729	24	1.00 (-)	1.05 (1.00 - 1.11)	1.21 (1.15 - 1.28)	1.35 (1.28 - 1.42)	1.80 (1.72 - 1.90)
Age, region, body mass index	682	29	1.00 (-)	1.05 (0.99 - 1.10)	1.19 (1.13 - 1.25)	1.33 (1.26 - 1.40)	1.77 (1.69 - 1.86)
Age, region, height	934	3	1.00 (-)	1.06 (1.01 - 1.12)	1.24 (1.17 - 1.30)	1.40 (1.33 - 1.47)	1.93 (1.84 - 2.02)
Age, region, parity	852	12	1.00 (-)	1.05 (1.00 - 1.11)	1.22 (1.16 - 1.29)	1.37 (1.30 - 1.44)	1.86 (1.77 - 1.95)
Age, region, menopausal hormone therapy	940	2	1.00 (-)	1.06 (1.00 - 1.11)	1.23 (1.17 - 1.29)	1.38 (1.31 - 1.45)	1.92 (1.83 - 2.02)
Age, region, health behaviours[2]	173	82	1.00 (-)	1.00 (0.95 - 1.06)	1.10 (1.04 - 1.16)	1.16 (1.10 - 1.22)	1.35 (1.28 - 1.43)
Age, region, health behaviours[2], plus education	118	88	1.00 (-)	0.99 (0.94 - 1.05)	1.08 (1.02 - 1.14)	1.12 (1.06 - 1.18)	1.29 (1.22 - 1.36)
Fully adjusted[3]	110	89	1.00 (-)	0.99 (0.94 - 1.05)	1.08 (1.02 - 1.14)	1.11 (1.05 - 1.18)	1.28 (1.21 - 1.35)

[1]Likelihood-ratio test statistic; [2]smoking, alcohol consumption, physical activity, body mass index; [3]adjusted for all health behaviours, education, height, parity, and ever use of menopausal hormone therapy; CVM=cardiovascular multimorbidity

Table E15 CVM and complex CVM incidence by education, restricted to never smokers

Education	LR[1]	attenuation in LR[1] (%)	Tertiary	Secondary	Technical	Compulsory schooling	< Compulsory schooling
INCIDENT CVM (n)			5,632	11,413	7,719	19,775	1,508
HR (95% CI), adjusted for,							
Age, region (minimally adjusted)	784	-	1.00 (-)	1.07 (1.04 - 1.11)	1.13 (1.10 - 1.17)	1.35 (1.31 - 1.39)	1.68 (1.58 - 1.78)
Age, region, deprivation	485	38	1.00 (-)	1.07 (1.03 - 1.10)	1.12 (1.08 - 1.16)	1.28 (1.24 - 1.32)	1.54 (1.45 - 1.63)
Age, region, smoking	520	34	1.00 (-)	1.06 (1.02 - 1.09)	1.10 (1.06 - 1.14)	1.28 (1.24 - 1.32)	1.55 (1.46 - 1.64)
Age, region, alcohol	471	40	1.00 (-)	1.05 (1.02 - 1.09)	1.10 (1.06 - 1.14)	1.27 (1.23 - 1.31)	1.54 (1.45 - 1.63)
Age, region, physical activity	338	57	1.00 (-)	1.05 (1.01 - 1.08)	1.09 (1.06 - 1.13)	1.22 (1.19 - 1.26)	1.45 (1.37 - 1.54)
Age, region, body mass index	823	-5	1.00 (-)	1.08 (1.05 - 1.12)	1.16 (1.12 - 1.20)	1.38 (1.34 - 1.42)	1.72 (1.62 - 1.82)
Age, region, height	736	6	1.00 (-)	1.08 (1.04 - 1.11)	1.14 (1.10 - 1.18)	1.35 (1.31 - 1.39)	1.64 (1.55 - 1.74)
Age, region, parity	784	0	1.00 (-)	1.07 (1.04 - 1.11)	1.13 (1.09 - 1.17)	1.36 (1.32 - 1.40)	1.69 (1.60 - 1.79)
Age, region, menopausal hormone therapy	146	81	1.00 (-)	1.02 (0.99 - 1.06)	1.05 (1.02 - 1.09)	1.14 (1.10 - 1.17)	1.32 (1.24 - 1.40)
Age, region, health behaviours[2]	86	89	1.00 (-)	1.02 (0.99 - 1.06)	1.05 (1.01 - 1.08)	1.10 (1.07 - 1.14)	1.25 (1.18 - 1.34)
Age, region, health behaviours[2], plus deprivation	126	84	1.00 (-)	1.03 (1.00 - 1.07)	1.07 (1.03 - 1.11)	1.14 (1.10 - 1.18)	1.31 (1.23 - 1.40)
Fully adjusted[3]	784	-	1.00 (-)	1.07 (1.04 - 1.11)	1.13 (1.10 - 1.17)	1.35 (1.31 - 1.39)	1.68 (1.58 - 1.78)
INCIDENT COMPLEX CVM (n)			1,122	2,359	1,632	4,652	412
HR (95% CI), adjusted for,							
Age, region (minimally adjusted)	321	-	1.00 (-)	1.10 (1.03 - 1.18)	1.16 (1.07 - 1.25)	1.49 (1.39 - 1.59)	2.06 (1.84 - 2.31)
Age, region, deprivation	206	36	1.00 (-)	1.09 (1.02 - 1.17)	1.14 (1.05 - 1.23)	1.39 (1.30 - 1.48)	1.84 (1.64 - 2.07)
Age, region, smoking	223	30	1.00 (-)	1.08 (1.01 - 1.16)	1.12 (1.03 - 1.20)	1.39 (1.30 - 1.48)	1.87 (1.67 - 2.10)
Age, region, alcohol	216	33	1.00 (-)	1.07 (1.00 - 1.15)	1.12 (1.03 - 1.21)	1.38 (1.29 - 1.48)	1.88 (1.67 - 2.11)
Age, region, physical activity	155	52	1.00 (-)	1.07 (1.00 - 1.15)	1.10 (1.02 - 1.19)	1.32 (1.23 - 1.41)	1.74 (1.54 - 1.96)
Age, region, body mass index	336	-5	1.00 (-)	1.11 (1.04 - 1.20)	1.18 (1.09 - 1.28)	1.52 (1.43 - 1.63)	2.13 (1.90 - 2.39)
Age, region, height	305	5	1.00 (-)	1.11 (1.03 - 1.19)	1.17 (1.08 - 1.26)	1.49 (1.40 - 1.59)	2.02 (1.80 - 2.26)
Age, region, parity	301	6	1.00 (-)	1.10 (1.02 - 1.18)	1.15 (1.07 - 1.24)	1.48 (1.38 - 1.58)	2.02 (1.80 - 2.27)
Age, region, menopausal hormone therapy	87	73	1.00 (-)	1.04 (0.97 - 1.12)	1.06 (0.98 - 1.15)	1.22 (1.13 - 1.30)	1.58 (1.40 - 1.79)
Age, region, health behaviours[2]	59	82	1.00 (-)	1.03 (0.96 - 1.11)	1.05 (0.97 - 1.14)	1.17 (1.09 - 1.26)	1.48 (1.31 - 1.68)
Age, region, health behaviours[2], plus deprivation	68	79	1.00 (-)	1.05 (0.98 - 1.14)	1.08 (1.00 - 1.17)	1.22 (1.13 - 1.31)	1.53 (1.34 - 1.73)
Fully adjusted[3]	321	-	1.00 (-)	1.10 (1.03 - 1.18)	1.16 (1.07 - 1.25)	1.49 (1.39 - 1.59)	2.06 (1.84 - 2.31)

[1]Likelihood-ratio test statistic; [2]smoking, alcohol consumption, physical activity, body mass index; [3]adjusted for all health behaviours, deprivation, height, parity, and ever use of menopausal hormone therapy; CVM=cardiovascular multimorbidity

Table E16 CVM and complex CVM incidence by area deprivation, restricted to never smokers

Area deprivation (fifths of Townsend index)	LR[1]	attenuation in LR[1] (%)	1 (Least deprived)	2	3	4	5 (Most deprived)
INCIDENT CVM (n)			9,256	9,579	9,571	9,587	8,853
HR (95% CI), adjusted for,							
Age, region (minimally adjusted)	997	-	1.00 (-)	1.06 (1.03 - 1.09)	1.14 (1.11 - 1.17)	1.27 (1.23 - 1.30)	1.54 (1.49 - 1.58)
Age, region, education	664	33	1.00 (-)	1.05 (1.02 - 1.08)	1.11 (1.08 - 1.15)	1.22 (1.19 - 1.26)	1.44 (1.39 - 1.48)
Age, region, smoking	781	22	1.00 (-)	1.05 (1.02 - 1.08)	1.12 (1.09 - 1.15)	1.23 (1.20 - 1.27)	1.47 (1.42 - 1.51)
Age, region, alcohol	747	25	1.00 (-)	1.05 (1.02 - 1.09)	1.12 (1.09 - 1.16)	1.24 (1.20 - 1.28)	1.47 (1.42 - 1.51)
Age, region, physical activity	487	51	1.00 (-)	1.04 (1.01 - 1.07)	1.09 (1.06 - 1.13)	1.19 (1.15 - 1.22)	1.37 (1.32 - 1.41)
Age, region, body mass index	984	1	1.00 (-)	1.06 (1.03 - 1.09)	1.14 (1.11 - 1.17)	1.27 (1.23 - 1.31)	1.54 (1.50 - 1.59)
Age, region, height	901	10	1.00 (-)	1.06 (1.03 - 1.09)	1.13 (1.10 - 1.17)	1.26 (1.22 - 1.29)	1.51 (1.46 - 1.55)
Age, region, parity	989	1	1.00 (-)	1.06 (1.03 - 1.09)	1.14 (1.10 - 1.17)	1.27 (1.23 - 1.31)	1.54 (1.49 - 1.58)
Age, region, menopausal hormone therapy	321	68	1.00 (-)	1.03 (1.00 - 1.06)	1.08 (1.05 - 1.11)	1.15 (1.12 - 1.19)	1.30 (1.26 - 1.34)
Age, region, health behaviours[2]	249	75	1.00 (-)	1.03 (1.00 - 1.06)	1.07 (1.04 - 1.10)	1.14 (1.11 - 1.18)	1.26 (1.22 - 1.30)
Age, region, health behaviours[2], plus education	245	75	1.00 (-)	1.03 (1.00 - 1.06)	1.07 (1.04 - 1.10)	1.14 (1.11 - 1.18)	1.26 (1.22 - 1.30)
Fully adjusted[3]	997	-	1.00 (-)	1.06 (1.03 - 1.09)	1.14 (1.11 - 1.17)	1.27 (1.23 - 1.30)	1.54 (1.49 - 1.58)
INCIDENT COMPLEX CVM (n)			1,932	2,080	2,122	2,137	2,148
HR (95% CI), adjusted for,							
Age, region (minimally adjusted)	364	-	1.00 (-)	1.08 (1.01 - 1.15)	1.18 (1.11 - 1.26)	1.32 (1.24 - 1.40)	1.74 (1.63 - 1.85)
Age, region, education	232	36	1.00 (-)	1.06 (1.00 - 1.13)	1.15 (1.08 - 1.22)	1.25 (1.17 - 1.33)	1.58 (1.48 - 1.69)
Age, region, smoking	286	21	1.00 (-)	1.07 (1.01 - 1.14)	1.16 (1.09 - 1.23)	1.28 (1.20 - 1.36)	1.64 (1.54 - 1.74)
Age, region, alcohol	280	23	1.00 (-)	1.07 (1.01 - 1.14)	1.17 (1.09 - 1.24)	1.29 (1.21 - 1.37)	1.65 (1.55 - 1.76)
Age, region, physical activity	169	54	1.00 (-)	1.06 (1.00 - 1.13)	1.11 (1.04 - 1.18)	1.22 (1.14 - 1.30)	1.48 (1.39 - 1.58)
Age, region, body mass index	352	3	1.00 (-)	1.09 (1.02 - 1.16)	1.18 (1.11 - 1.26)	1.33 (1.25 - 1.42)	1.74 (1.63 - 1.85)
Age, region, height	326	10	1.00 (-)	1.08 (1.01 - 1.15)	1.18 (1.10 - 1.25)	1.31 (1.23 - 1.39)	1.69 (1.59 - 1.80)
Age, region, parity	350	4	1.00 (-)	1.08 (1.01 - 1.15)	1.18 (1.10 - 1.25)	1.31 (1.23 - 1.40)	1.73 (1.62 - 1.84)
Age, region, menopausal hormone therapy	112	69	1.00 (-)	1.05 (0.99 - 1.12)	1.09 (1.03 - 1.17)	1.18 (1.11 - 1.26)	1.39 (1.30 - 1.49)
Age, region, health behaviours[2]	82	77	1.00 (-)	1.04 (0.98 - 1.11)	1.08 (1.01 - 1.15)	1.15 (1.08 - 1.23)	1.34 (1.25 - 1.43)
Age, region, health behaviours[2], plus education	80	78	1.00 (-)	1.04 (0.98 - 1.12)	1.07 (1.00 - 1.15)	1.15 (1.08 - 1.23)	1.34 (1.25 - 1.43)
Fully adjusted[3]	364	-	1.00 (-)	1.08 (1.01 - 1.15)	1.18 (1.11 - 1.26)	1.32 (1.24 - 1.40)	1.74 (1.63 - 1.85)

[1]Likelihood-ratio test statistic; [2]smoking, alcohol consumption, physical activity, body mass index; [3]adjusted for all health behaviours, education, height, parity, and ever use of menopausal hormone therapy; CVM=cardiovascular multimorbidity

Appendix F: Reproductive factors and cardiovascular multimorbidity

Appendix

Table F1 Association of age at menarche with complex cardiovascular multimorbidity risk after adjustment for various covariates

	LR[1]	attenuation in LR[1] (%)	Age at menarche (years)				
			<12	12	13	14	15+
Number of incident CVM events			5,780	4,342	5,851	5,300	4,933
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	286	-	1.33 (1.29 - 1.38)	1.06 (1.01 - 1.10)	1.00 (-)	1.03 (0.99 - 1.07)	1.11 (1.07 - 1.15)
Age, region, deprivation	268	6	1.31 (1.27 - 1.36)	1.06 (1.01 - 1.10)	1.00 (-)	1.01 (0.97 - 1.05)	1.07 (1.03 - 1.11)
Age, region, education	275	4	1.31 (1.26 - 1.36)	1.06 (1.02 - 1.11)	1.00 (-)	0.99 (0.95 - 1.03)	1.04 (1.00 - 1.09)
Age, region, smoking	239	16	1.30 (1.25 - 1.35)	1.05 (1.00 - 1.09)	1.00 (-)	1.00 (0.97 - 1.04)	1.05 (1.01 - 1.10)
Age, region, alcohol	269	6	1.32 (1.27 - 1.37)	1.05 (1.01 - 1.09)	1.00 (-)	1.02 (0.98 - 1.06)	1.09 (1.05 - 1.13)
Age, region, physical activity	263	8	1.32 (1.27 - 1.37)	1.05 (1.01 - 1.10)	1.00 (-)	1.02 (0.98 - 1.05)	1.09 (1.05 - 1.13)
Age, region, body mass index	136	52	1.21 (1.17 - 1.26)	1.02 (0.98 - 1.07)	1.00 (-)	1.04 (1.00 - 1.08)	1.14 (1.10 - 1.19)
Age, region, height	300	-5	1.35 (1.30 - 1.40)	1.06 (1.02 - 1.10)	1.00 (-)	1.03 (0.99 - 1.07)	1.10 (1.06 - 1.15)
Age, region, parity	284	1	1.33 (1.28 - 1.38)	1.06 (1.02 - 1.10)	1.00 (-)	1.03 (0.99 - 1.07)	1.10 (1.06 - 1.14)
Age, region, menopausal hormone therapy	282	1	1.33 (1.29 - 1.38)	1.06 (1.02 - 1.10)	1.00 (-)	1.03 (0.99 - 1.07)	1.10 (1.06 - 1.15)
Age, region, health behaviours[2]	82	71	1.17 (1.12 - 1.22)	1.01 (0.97 - 1.06)	1.00 (-)	1.00 (0.96 - 1.04)	1.07 (1.03 - 1.12)
Age, region, health behaviours[2], plus socioeconomic factors	84	71	1.17 (1.12 - 1.21)	1.01 (0.97 - 1.06)	1.00 (-)	0.98 (0.94 - 1.02)	1.04 (0.99 - 1.08)
Fully adjusted[3]	99	63	1.18 (1.13 - 1.23)	1.02 (0.98 - 1.07)	1.00 (-)	0.97 (0.93 - 1.01)	1.02 (0.98 - 1.07)

[1]Likelihood-ratio test statistic;

[2]smoking, alcohol consumption, physical activity, body mass index;

[3]adjusted for all health behaviours, education, deprivation, height, parity, and ever use of menopausal hormone therapy;

CVM = cardiovascular multimorbidity

Appendix

Table F2 Association of parity with complex cardiovascular multimorbidity risk after adjustment for various covariates

	LR[1]	attenuation in LR[1] (%)	Parity (number of full-term pregnancies)				
			Nulliparous	**1**	**2**	**3**	**4+**
Number of incident CVM events			2,827	3,216	9,593	6,296	4,797
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region (minimally adjusted)	651	-	1.00 (-)	0.98 (0.93 - 1.03)	0.90 (0.86 - 0.93)	1.06 (1.02 - 1.11)	1.42 (1.35 - 1.49)
Age, region, deprivation	379	42	1.00 (-)	0.98 (0.93 - 1.03)	0.91 (0.87 - 0.95)	1.05 (1.01 - 1.10)	1.30 (1.24 - 1.36)
Age, region, education	447	31	1.00 (-)	0.92 (0.87 - 0.97)	0.86 (0.83 - 0.90)	1.01 (0.96 - 1.05)	1.27 (1.21 - 1.34)
Age, region, smoking	418	36	1.00 (-)	0.94 (0.89 - 0.99)	0.89 (0.85 - 0.93)	1.04 (0.99 - 1.08)	1.29 (1.23 - 1.36)
Age, region, alcohol	541	17	1.00 (-)	0.97 (0.92 - 1.02)	0.90 (0.86 - 0.94)	1.06 (1.02 - 1.11)	1.37 (1.31 - 1.44)
Age, region, physical activity	524	19	1.00 (-)	0.96 (0.92 - 1.02)	0.90 (0.86 - 0.94)	1.06 (1.01 - 1.11)	1.37 (1.31 - 1.44)
Age, region, body mass index	351	46	1.00 (-)	0.97 (0.92 - 1.03)	0.90 (0.87 - 0.94)	1.04 (1.00 - 1.09)	1.28 (1.22 - 1.35)
Age, region, height	620	5	1.00 (-)	0.98 (0.94 - 1.04)	0.90 (0.87 - 0.94)	1.07 (1.02 - 1.12)	1.42 (1.36 - 1.49)
Age, region, menopausal hormone therapy	613	6	1.00 (-)	0.98 (0.93 - 1.03)	0.90 (0.86 - 0.94)	1.06 (1.02 - 1.11)	1.41 (1.35 - 1.48)
Age, region, health behaviours[2]	143	78	1.00 (-)	0.92 (0.87 - 0.97)	0.90 (0.86 - 0.94)	1.01 (0.96 - 1.06)	1.13 (1.07 - 1.19)
Age, region, health behaviours[2], plus socioeconomic factors	82	87	1.00 (-)	0.90 (0.85 - 0.95)	0.89 (0.85 - 0.93)	0.97 (0.93 - 1.02)	1.04 (0.99 - 1.10)
Fully adjusted[3]	78	88	1.00 (-)	0.91 (0.86 - 0.96)	0.89 (0.85 - 0.94)	0.98 (0.93 - 1.03)	1.05 (1.00 - 1.11)

[1]Likelihood-ratio test statistic;

[2]smoking, alcohol consumption, physical activity, body mass index;

[3]adjusted for all health behaviours, education, deprivation, height, and ever use of menopausal hormone therapy;

CVM = cardiovascular multimorbidity

Appendix

Table F3 Association of total breastfeeding duration with complex cardiovascular multimorbidity risk in parous women after adjustment for various covariates

In parous women	LR[1]	attenuation in LR[1] (%)	Total breastfeeding duration (months)				
			None	1	2 to 5	6 to 11	12+
Number of incident CVM events			5,627	2,264	4,212	2,887	2,915
Cox regression models adjusted for,			Hazard Ratio (95% Confidence Interval)				
Age, region, parity (minimally adjusted)	318	-	1.00 (-)	0.81 (0.77 - 0.85)	0.81 (0.78 - 0.85)	0.74 (0.71 - 0.78)	0.68 (0.65 - 0.71)
Age, region, parity, deprivation	245	23	1.00 (-)	0.82 (0.78 - 0.86)	0.83 (0.80 - 0.86)	0.77 (0.74 - 0.81)	0.71 (0.68 - 0.75)
Age, region, parity, education	180	44	1.00 (-)	0.83 (0.79 - 0.87)	0.84 (0.81 - 0.88)	0.79 (0.76 - 0.83)	0.74 (0.71 - 0.78)
Age, region, parity, smoking	183	42	1.00 (-)	0.83 (0.79 - 0.87)	0.83 (0.80 - 0.87)	0.78 (0.75 - 0.82)	0.75 (0.71 - 0.78)
Age, region, parity, alcohol	269	15	1.00 (-)	0.81 (0.78 - 0.86)	0.83 (0.80 - 0.86)	0.76 (0.73 - 0.80)	0.70 (0.67 - 0.73)
Age, region, parity, physical activity	211	34	1.00 (-)	0.82 (0.78 - 0.86)	0.84 (0.80 - 0.87)	0.78 (0.75 - 0.82)	0.72 (0.69 - 0.76)
Age, region, parity, body mass index	218	32	1.00 (-)	0.83 (0.78 - 0.87)	0.84 (0.80 - 0.87)	0.79 (0.75 - 0.82)	0.71 (0.68 - 0.75)
Age, region, parity, height	323	-1	1.00 (-)	0.81 (0.77 - 0.85)	0.81 (0.78 - 0.84)	0.74 (0.71 - 0.77)	0.67 (0.64 - 0.71)
Age, region, parity, menopausal hormone therapy	309	3	1.00 (-)	0.81 (0.77 - 0.85)	0.81 (0.78 - 0.85)	0.75 (0.71 - 0.78)	0.68 (0.65 - 0.71)
Age, region, parity, health behaviours	75	76	1.00 (-)	0.86 (0.82 - 0.91)	0.88 (0.84 - 0.92)	0.86 (0.82 - 0.91)	0.82 (0.78 - 0.86)
Age, region, parity, health behaviours, socioeconomic factors	55	83	1.00 (-)	0.87 (0.82 - 0.92)	0.89 (0.85 - 0.93)	0.88 (0.84 - 0.93)	0.84 (0.80 - 0.89)
Fully adjusted[3]	59	81	1.00 (-)	0.87 (0.82 - 0.92)	0.89 (0.85 - 0.93)	0.88 (0.83 - 0.92)	0.84 (0.79 - 0.88)

[1]Likelihood-ratio test statistic;

[2]smoking, alcohol consumption, physical activity, body mass index;

[3]adjusted for all health behaviours, education, deprivation, height, parity, and ever use of menopausal hormone therapy;

CVM = cardiovascular multimorbidity

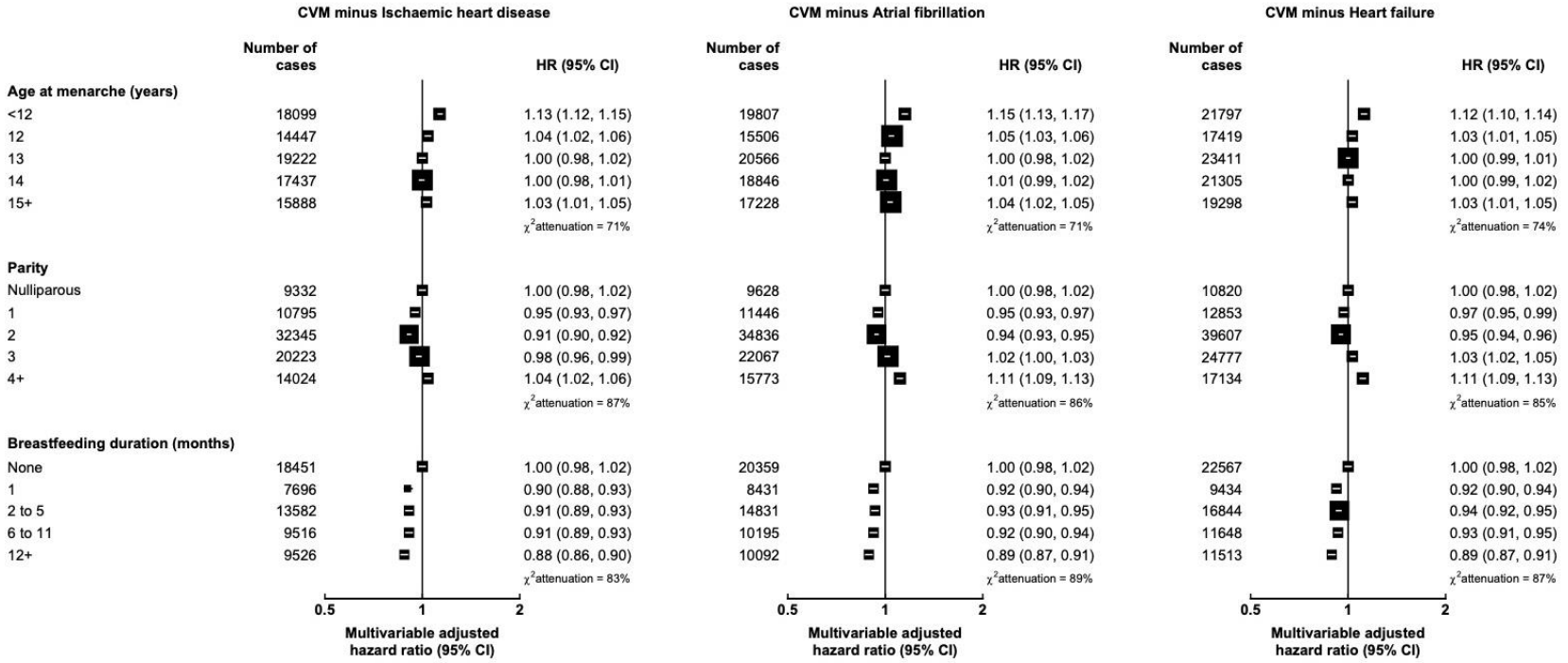


Figure F1 Relative risks of CVM minus ischaemic heart disease, atrial fibrillation, or heart failure in relation to reproductive factors

Minimally adjusted models for age at menarche and parity were stratified by age at recruitment and area of residence only. Minimally adjusted models for breastfeeding duration were also adjusted for parity.
Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy.
$\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models.
The size of the square boxes was inversely proportional to the standard error of the effect estimate. Floated confidence intervals were displayed.

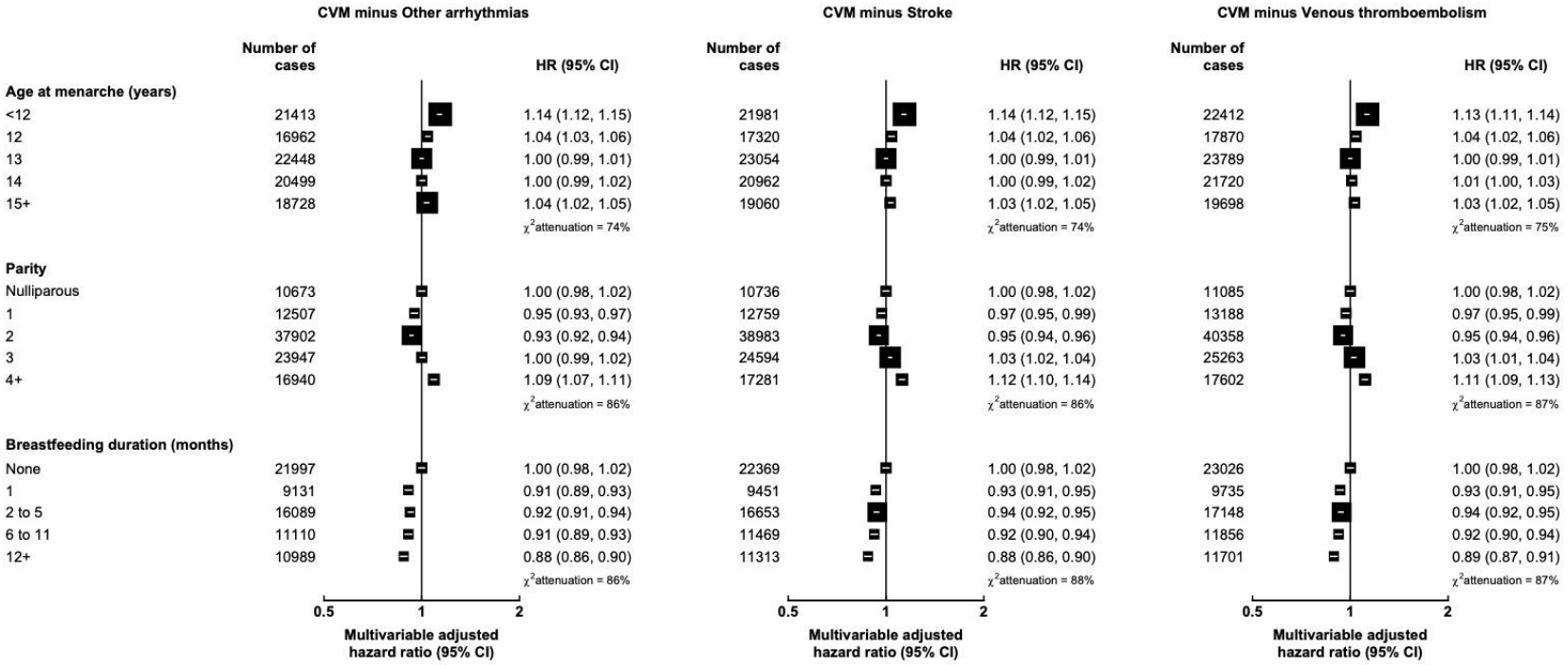


Figure F2 Relative risks of CVM minus other arrhythmias, stroke, or venous thromboembolism in relation to reproductive factors

Minimally adjusted models for age at menarche and parity were stratified by age at recruitment and area of residence only. Minimally adjusted models for breastfeeding duration were also adjusted for parity.
Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy.
$\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models.
The size of the square boxes was inversely proportional to the standard error of the effect estimate. Floated confidence intervals were displayed.

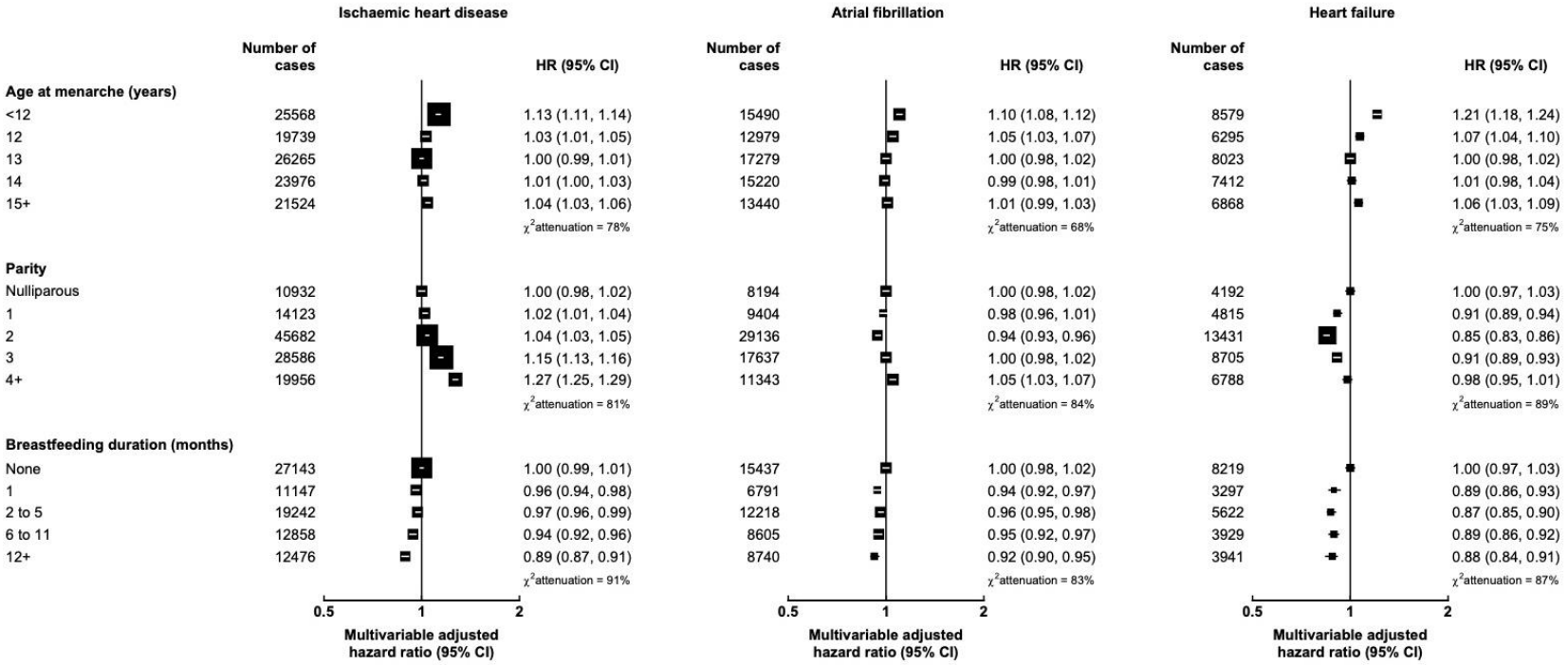


Figure F3 Relative risks of ischaemic heart disease, atrial fibrillation, and heart failure in relation to reproductive factors

Minimally adjusted models for age at menarche and parity were stratified by age at recruitment and area of residence only. Minimally adjusted models for breastfeeding duration were also adjusted for parity.
Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy.
$\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models.
The size of the square boxes was inversely proportional to the standard error of the effect estimate. Floated confidence intervals were displayed.

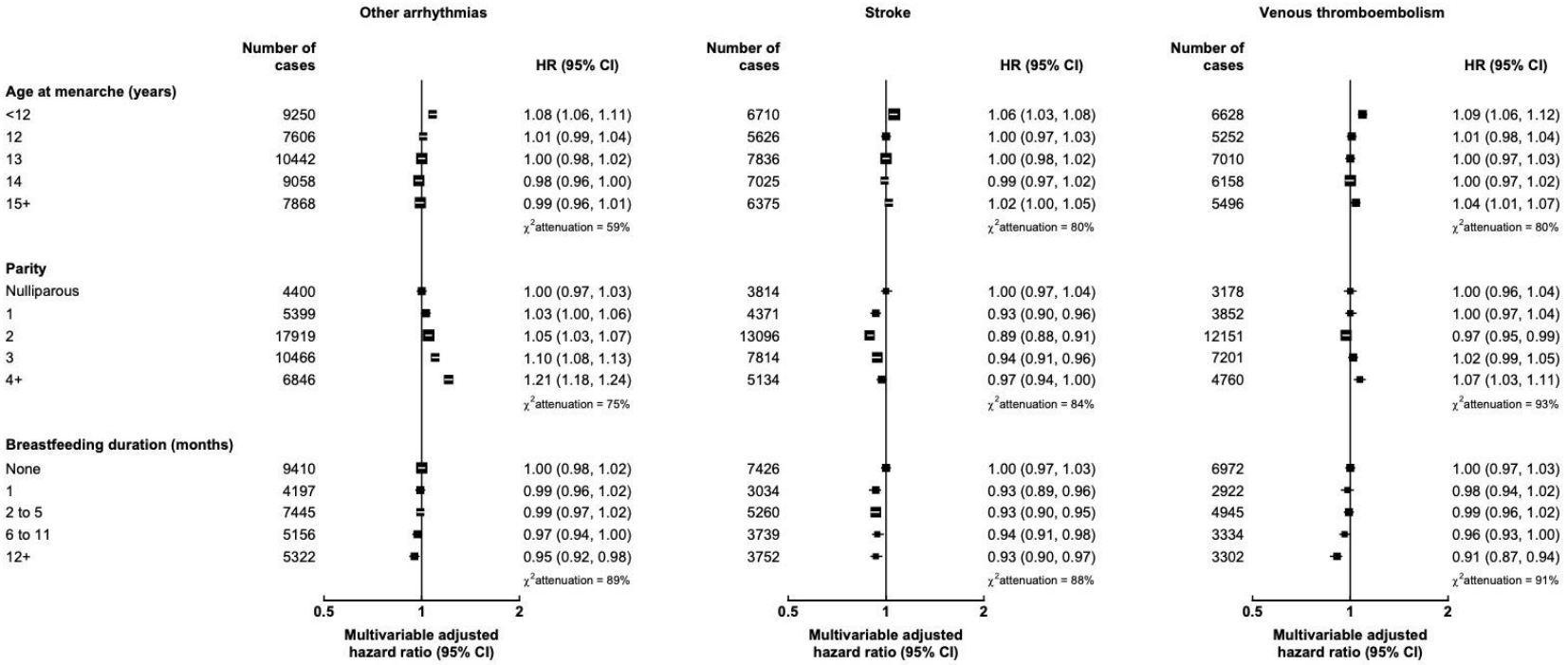


Figure F4 Relative risks of other arrhythmias, stroke, and venous thromboembolism in relation to reproductive factors

Minimally adjusted models for age at menarche and parity were stratified by age at recruitment and area of residence only. Minimally adjusted models for breastfeeding duration were also adjusted for parity.
Multivariable adjusted models were additionally adjusted for area deprivation, educational attainment, physical activity, smoking, alcohol consumption, height, body mass index, parity, and ever use of menopausal hormone therapy.
$\chi^2_{attenuation}$ was calculated as the relative change in the likelihood-ratio χ^2 statistic between the minimally and multivariable adjusted models.
The size of the square boxes was inversely proportional to the standard error of the effect estimate. Floated confidence intervals were displayed.