

**Blood glucose and diabetes mellitus as risk
factors for cardiovascular disease in the
China Kadoorie Biobank of 0.5 million people**

Fiona J Bragg

BSc(Hons) MBChB MRCP MSc MFPH

St. Anne's College, University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Trinity Term 2015

Abstract

Blood glucose and diabetes mellitus as risk factors for cardiovascular disease in the China Kadoorie Biobank of 0.5 million people

Fiona J Bragg, St. Anne's College, University of Oxford

A thesis submitted for the degree of Doctor of Philosophy, Trinity Term 2015

Background: In China, the incidence of diabetes is rising, but little is known about its impact on cardiovascular diseases (CVD). Moreover, substantial uncertainty remains, both in China and elsewhere, about the relevance to CVD of blood glucose levels below the glycaemic threshold for diabetes. This thesis assessed these associations in Chinese adults.

Methods: In 2004-08, the China Kadoorie Biobank recruited 0.5 million adults aged 30-79 years from 10 diverse regions, recording 25 488 deaths by 1st January 2014. Logistic and Cox regression yielded adjusted odds (ORs) and hazard ratios (HRs), respectively, relating self-reported diabetes and, among people without self-reported diabetes, random plasma glucose (RPG) levels to prevalent and incident CVD.

Results: At baseline, 5.9% of participants had self-reported (3.2%) or screen-detected (2.8%) diabetes. Self-reported diabetes was associated with approximately 2-fold increased odds of prevalent ischaemic heart disease (IHD) (n=14 771, OR 2.16, 95% CI 2.05-2.28) and stroke/transient ischaemic attack (n=8430, 1.88, 1.75-2.02). There were also significant excess risks of CVD mortality (n=6916, HR 2.07, 95% CI 1.90-2.26) and incident IHD (n=3465, 2.44, 2.18-2.73), ischaemic stroke (n=19 778, 1.68, 1.60-1.77) and intra-cerebral haemorrhage (ICH) (n=4423, 1.24, 1.07-1.44). These risks increased with longer diabetes duration and number of other presenting risk factors. Screen-detected diabetes was also associated with significant, though more modest, excess incident CVD risks. Among participants without self-reported diabetes, there was a positive log-linear association of RPG with prevalent and incident CVDs. In prospective analyses the association continued down to baseline RPG of ~4.0 mmol/L. After adjusting for regression dilution bias, each 1 mmol/L higher usual RPG was associated with ~10% increased risks of incident ischaemic CVDs and CVD mortality, and ~5% increased risks of incident ICH. In ~89 000 genotyped individuals, a RPG-related genetic risk score was associated significantly with RPG. It showed no significant association with ischaemic CVD risks, but power to detect such an association was low.

Conclusions: Among adult Chinese, diabetes is associated with significantly increased risks of major CVDs. In individuals without diabetes, there is a positive log-linear association between RPG and CVDs, with no evidence of a threshold.

Acknowledgements

I am indebted to the China Kadoorie Biobank participants for their vital contributions, to the members of the survey teams in the 10 study areas and to the project development and management teams at these centres and in Beijing and Oxford.

I would like to thank my primary supervisor, Professor Zhengming Chen, my second supervisor, Dr Derrick Bennett, and my previous second supervisor, Dr Margaret Smith, for their support and guidance during my DPhil.

Thanks are due, also, to members of the China Kadoorie Biobank team and to other staff in the Clinical Trial Service Unit and Epidemiological Studies Unit, including Professor Robert Clarke, Dr Christina Davies, Dr Huaidong Du, Mr. Alex Hacker, Dr Sarah Lewington, Dr Iona Millwood, Mr. Sam Sansome, Mr. Dan Schmidt, Mr. Paul Sherliker, Dr Robin Walters and Dr Ling Yang.

Finally I would like to thank the British Heart Foundation Centre of Research Excellence in Oxford and the Clinical Trial Service Unit and Epidemiological Studies Unit, University of Oxford for providing funding to enable me to undertake this DPhil.

Statement of contribution to work

I was responsible for conceiving and developing the research questions addressed in this thesis. I performed all aspects of the literature reviews and planned and designed the study. I conducted all statistical analyses (some analyses used statistical programmes developed within the Clinical Trial Service Unit and Epidemiological Studies Unit, as reported at the relevant points in the thesis), and produced all the tables and figures presented. I interpreted the findings and wrote-up all aspects of the study, including the thesis and related publications.

Contents

Abstract	ii
Acknowledgements.....	iii
Statement of contribution to work.....	iv
List of tables	x
List of figures	xiv
List of abbreviations	xx
Chapter 1. Introduction	1
1.1 Cardiovascular disease and its subtypes.....	1
1.1.1 Ischaemic heart disease	2
1.1.2 Stroke	3
1.2 Epidemiology of cardiovascular diseases	4
1.2.1 International perspective	4
1.2.2 Epidemiology of cardiovascular diseases in China.....	6
1.3 Diabetes, blood glucose and cardiovascular disease	7
1.3.1 Glucose metabolism and homeostasis.....	7
1.3.2 Diabetes and cardiovascular disease.....	9
1.3.3 Blood glucose and cardiovascular disease.....	10
Chapter 2. Literature review.....	12
2.1 Objectives	12
2.2 Methodology.....	12
2.3 Diabetes and cardiovascular disease	12
2.3.1 Cross-sectional and case-control evidence	12
2.3.2 Prospective evidence	19
2.4 Blood glucose levels and cardiovascular disease risk.....	27
2.4.1 Cross-sectional and case-control evidence	27
2.4.2 Prospective evidence	29
2.4.3 Mendelian randomisation studies assessing the causal relevance of blood glucose to cardiovascular disease	37
2.5 Summary of findings and main research questions	41
2.6 Aims and objectives of thesis	42
Chapter 3. China Kadoorie Biobank methods and study population.....	67
3.1 Background	67
3.2 Study design	68
3.3 Study population.....	68
3.4 Baseline data collection.....	69

3.5 DNA extraction and genotyping	71
3.6 Resurvey	71
3.7 Follow-up and outcome adjudication	73
3.8 Baseline characteristics of China Kadoorie Biobank participants	74
3.9 Baseline characteristics of the China Kadoorie Biobank resurvey populations	75
3.10 Baseline characteristics of the China Kadoorie Biobank genotyped population ...	75
3.11 Discussion	76
Chapter 4. Prevalence, variability and correlates of diabetes and plasma glucose	83
4.1 Background	83
4.2 Methods	84
4.2.1 Assessment of diabetes status	84
4.2.2 Assessment of plasma glucose	84
4.2.3 Participant exclusions	85
4.2.4 Statistical analyses	85
4.3 Results	89
4.3.1 Prevalence of diabetes	89
4.3.2 Cross-sectional associations of diabetes	91
4.3.3 Plasma glucose distribution	93
4.3.4 Cross-sectional associations of plasma glucose	94
4.3.5 Within-person variation in plasma glucose	95
4.4 Discussion	96
4.4.1 Prevalence and correlates of diabetes	96
4.4.2 Validity of diabetes diagnoses	99
4.4.3 Distribution and correlates of random plasma glucose	100
4.4.4 Intra-individual variation in random plasma glucose levels	101
4.4.5 Random plasma glucose as a glycaemic measure	102
4.4.6 Conclusion	104
Chapter 5. Diabetes, plasma glucose and prevalent cardiovascular disease	133
5.1 Background	133
5.2 Methods	134
5.2.1 Exposure definitions	134
5.2.2 Outcome definitions	134
5.2.3 Selection of confounders	134
5.2.4 Participant exclusions	135
5.2.5 Statistical analyses	135
5.3 Results	137

5.3.1 Prevalence and correlates of cardiovascular disease.....	137
5.3.2 Diabetes and prevalent cardiovascular disease	139
5.3.3 Blood glucose and prevalent cardiovascular disease	140
5.4 Discussion.....	141
5.4.1 Diabetes and prevalent cardiovascular disease	141
5.4.2 Blood glucose and prevalent cardiovascular disease	144
5.4.3 Limitations	145
5.4.4 Conclusion	147
Chapter 6. Diabetes and the risk of incident cardiovascular disease.....	168
6.1 Background	168
6.2 Methods	169
6.2.1 Exposure definitions.....	169
6.2.2 Outcome definitions	169
6.2.3 Selection of confounders.....	170
6.2.4 Participant exclusions	170
6.2.5 Statistical analyses	170
6.3 Results	173
6.3.1 Incidence and correlates of cardiovascular diseases	174
6.3.2 Diabetes and risk of cardiovascular disease	175
6.4 Discussion.....	178
6.4.1 Incidence of cardiovascular diseases.....	179
6.4.2 Diabetes and risk of cardiovascular disease	180
6.4.3 Accuracy of ischaemic heart disease and stroke diagnoses.....	185
6.4.4 Limitations	187
6.4.5 Conclusion.....	188
Chapter 7. Plasma glucose and the risk of incident cardiovascular disease.....	216
7.1 Background	216
7.2 Methods	217
7.2.1 Exposure definitions.....	217
7.2.2 Outcome definitions	217
7.2.3 Selection of confounders.....	218
7.2.4 Participant exclusions	218
7.2.5 Statistical analyses	218
7.3 Results	222
7.3.1 Incidence and correlates of cardiovascular diseases	222
7.3.2 Plasma glucose and risk of cardiovascular disease	222

7.4 Discussion.....	226
7.4.1 Plasma glucose and risk of cardiovascular disease	226
7.4.2 Limitations	232
7.4.3 Conclusion	234
Chapter 8. Assessing the causal association between random plasma glucose and cardiovascular disease: a Mendelian randomisation study.....	261
8.1 Background	261
8.2 Methods	263
8.2.1 Population.....	263
8.2.2 Genotyping	263
8.2.3 Exposure definitions.....	264
8.2.4 Outcome definitions	264
8.2.5 Participant exclusions	264
8.2.6 Statistical analyses	265
8.3 Results	267
8.3.1 Random plasma glucose and risk of cardiovascular diseases: observational analyses	267
8.3.2 Genotype and random plasma glucose.....	267
8.3.3 Genotype and risk of cardiovascular diseases	268
8.3.4 Random plasma glucose and cardiovascular diseases: genetic analyses	268
8.4 Discussion.....	269
8.4.1. Strengths and limitations.....	272
8.4.2. Conclusion	274
Chapter 9. Discussion.....	285
9.1 Summary of main findings	285
9.1.1 Distribution and correlates of diabetes and plasma glucose.....	285
9.1.2 Diabetes, plasma glucose and prevalent cardiovascular disease	286
9.1.3 Diabetes, plasma glucose and incident cardiovascular diseases	286
9.1.4 The causal relevance of plasma glucose in cardiovascular disease.....	287
9.2 Consistency with published literature.....	287
9.2.1 Distribution and correlates of diabetes and plasma glucose.....	288
9.2.2 Diabetes and cardiovascular disease.....	288
9.2.3 Plasma glucose and cardiovascular disease.....	290
9.2.4 The causal relevance of plasma glucose in cardiovascular disease.....	291
9.3 Causality	291
9.4 Strengths and limitations of the study	295

9.5 Clinical and public health implications	299
9.6 Future research	300
9.7 Conclusion	302
References	303
Appendix A. List of papers authored during DPhil	338
Appendix B. Additional China Kadoorie Biobank methods	340
Appendix C. Additional methods and results: Intra-individual variation in plasma glucose levels	343
Appendix D. Additional results: Diabetes, plasma glucose and prevalent cardiovascular disease	349
Appendix E. Additional methods and results: Diabetes and the risk of incident cardiovascular disease	360
Appendix F. Additional methods and results: Plasma glucose and the risk of incident cardiovascular disease	385
Appendix G. Additional methods and results: Assessing the causal association between random plasma glucose and cardiovascular disease: a Mendelian randomisation study	404

List of tables

Table 1.1. Diagnostic criteria for diabetes and pre-diabetes	11
Table 2.1. Literature review search terms	44
Table 2.2. Selection of articles for inclusion in the literature review	45
Table 2.3. Characteristics of studies examining the relationship between diabetes and prevalent ischaemic heart disease	46
Table 2.4. Key findings of studies examining the relationship between diabetes and prevalent ischaemic heart disease	47
Table 2.5. Characteristics of studies examining the relationship between diabetes and prevalent stroke	48
Table 2.6. Key findings of studies examining the relationship between diabetes and prevalent stroke	49
Table 2.7. Characteristics of studies examining the relationship between diabetes and incident ischaemic heart disease.....	50
Table 2.8. Key findings of studies examining the relationship between diabetes and incident ischaemic heart disease.....	51
Table 2.9. Characteristics of studies examining the relationship between diabetes and incident stroke.....	52
Table 2.10. Key findings of studies examining the relationship between diabetes and incident stroke.....	54
Table 2.11. Characteristics of studies examining the relationship between blood glucose levels and prevalent ischaemic heart disease	56
Table 2.12. Key findings of studies examining the relationship between blood glucose levels and prevalent ischaemic heart disease	57
Table 2.13. Characteristics of studies examining the relationship between blood glucose levels and incident ischaemic heart disease.....	58
Table 2.14. Key findings of studies examining the relationship between blood glucose levels and incident ischaemic heart disease.....	59
Table 2.15. Characteristics of studies examining the relationship between blood glucose levels and incident stroke	61
Table 2.16. Key findings of studies examining the relationship between blood glucose levels and incident stroke	62
Table 2.17. Characteristics of Mendelian randomisation studies examining the relationship between blood glucose and cardiovascular disease.....	65
Table 2.18. Key findings of Mendelian randomisation studies examining the relationship between blood glucose and cardiovascular disease.....	66
Table 3.1. Proportion of participants recruited in study areas, by sex.....	78
Table 3.2. Baseline demographic and lifestyle characteristics of participants by sex....	79
Table 3.3. Mean values and distributions of physical measurements and plasma glucose by sex	80
Table 3.4. Baseline characteristics of the entire study population and of resurvey participants	81
Table 3.5. Baseline characteristics of the entire study population and genotyped population	82

Table 4.1. Frequency of missing baseline random plasma glucose levels and random plasma glucose levels >27.8 mmol/L by study area in individuals without self-reported diabetes at baseline	105
Table 4.2. Baseline characteristics of participants by availability of baseline random plasma glucose data in individuals without self-reported diabetes at baseline.....	106
Table 4.3. Diabetes status at baseline and resurvey among participants included in first and second resurveys	110
Table 4.4. Self-reported diabetes medication use in men and women with self-reported diabetes at baseline	111
Table 4.5. Baseline diabetes prevalence by lifestyle characteristics	112
Table 4.6. Baseline diabetes prevalence by dietary consumption.....	113
Table 4.7. Mean adjusted difference in baseline plasma glucose level between individuals with fasting period ≥8 hours and <8 hours among individuals without self-reported diabetes at baseline	120
Table 4.8. Mean baseline fasting plasma glucose and glycaemic categories among individuals without self-reported diabetes at baseline.....	121
Table 4.9. Baseline random plasma glucose by study area in participants without self-reported diabetes at baseline	123
Table 4.10. Baseline random plasma glucose by lifestyle characteristics in participants without self-reported diabetes at baseline	124
Table 4.11. Baseline random plasma glucose by dietary consumption in participants without self-reported diabetes at baseline	125
Table 4.12. Age and sex-specific regression dilution ratios based on baseline and first resurvey in participants without self-reported diabetes at baseline, estimated by different methods	130
Table 5.1. Self-reported ischaemic heart disease at baseline and resurvey among participants included in both first and second resurveys.....	151
Table 5.2. Self-reported stroke/transient ischaemic attack at baseline and resurvey among participants included in both first and second resurveys.....	152
Table 5.3. Adjusted odds ratios for prevalent cardiovascular diseases by self-reported diabetes status in men and women	153
Table 5.4. Adjusted odds ratios for prevalent cardiovascular diseases by screen-detected diabetes status in men and women.....	160
Table 5.5. Adjusted odds ratios for prevalent cardiovascular diseases per 1 mmol/L higher baseline random plasma glucose at random plasma glucose levels <11.1 mmol/L in individuals without self-reported diabetes.....	162
Table 6.1. Adjusted hazard ratios for incident cardiovascular diseases by self-reported diabetes status.....	195
Table 6.2. Adjusted hazard ratios for incident cardiovascular diseases by screen-detected diabetes status	214
Table 7.1. Adjusted hazard ratios for incident cardiovascular diseases by baseline random plasma glucose	235
Table 7.2. Adjusted hazard ratios for incident cardiovascular diseases per 1 mmol/L higher usual random plasma glucose.....	241

Table 7.3. Adjusted hazard ratios for incident cardiovascular diseases by baseline random plasma glucose excluding individuals recruited from Zhejiang.....	258
Table 7.4. Adjusted hazard ratios for incident cardiovascular diseases by baseline random plasma glucose excluding individuals with new-onset of diabetes during follow-up.....	259
Table 7.5. Adjusted hazard ratios for incident cardiovascular diseases by baseline random plasma glucose excluding individuals with screen-detected diabetes at baseline	260
Table 8.1. Adjusted hazard ratios for incident cardiovascular diseases per 1 mmol/L higher baseline and usual random plasma glucose	276
Table 8.2. Association of previously published plasma glucose-related SNPs with baseline random plasma glucose	277
Table 8.3. Adjusted hazard ratios for incident cardiovascular disease per glucose-increasing allele	281
Table 8.4. Observed and Mendelian randomisation expected hazard ratios for incident cardiovascular diseases per glucose-increasing allele in the random plasma glucose genetic risk score	284
Table B.1. Type 2 diabetes-related SNPs genotyped in the CKB	341
Table C1. Age and sex-specific regression dilution ratios based on baseline and first resurvey in participants without self-reported diabetes at baseline or first resurvey, estimated by different methods	344
Table C2. Mean baseline and resurvey plasma glucose levels and regression dilution ratios by baseline and resurvey fasting period in participants without self-reported diabetes at baseline	348
Table D.1. Association of baseline ischaemic heart disease prevalence with lifestyle characteristics	351
Table D.2. Association of baseline ischaemic heart disease prevalence with dietary consumption.....	352
Table D.3. Association of baseline stroke/transient ischaemic attack prevalence with lifestyle characteristics	355
Table D.4. Association of baseline stroke/transient ischaemic attack prevalence with dietary consumption	356
Table E.1. Cardiovascular disease outcome definitions	361
Table E.2. Self-reported cardiovascular disease medication use in men and women with self-reported diabetes at baseline	366
Table E.3. Association of major occlusive vascular disease with lifestyle characteristics	367
Table E.4. Association of major occlusive vascular disease with dietary consumption.....	368
Table E.5. Association of intra-cerebral haemorrhage with lifestyle characteristics	371
Table E.6. Association of intra-cerebral haemorrhage with dietary consumption	372
Table E.7. Adjusted hazard ratios for incident cardiovascular diseases by self-reported diabetes status, excluding individuals with diabetes newly diagnosed during follow-up from the comparator group.....	375

Table E.8. Cardiovascular disease risk factors in men and women with and without self-reported diabetes at baseline	384
Table F.1. Adjusted hazard ratios for incident cardiovascular diseases by baseline fasting time-adjusted baseline plasma glucose	403
Table G.1. Adjusted hazard ratios for incident cardiovascular diseases per 1 mmol/L higher random plasma glucose in genotyped population	406
Table G.2. Association of random plasma glucose-related SNPs with baseline characteristics	407
Table G.3. Power estimates for detection of an association between genetically determined random plasma glucose and risk of incident cardiovascular diseases	409
Table G.4. Effect allele frequencies by study area	410

List of figures

Figure 3.1. Locations of the China Kadoorie Biobank recruitment centres.....	77
Figure 4.1. Participant exclusions for presented analyses	107
Figure 4.2. Baseline prevalence of self-reported and screen-detected diabetes by age.....	108
Figure 4.3. Baseline prevalence of self-reported and screen-detected diabetes by study area	109
Figure 4.4. Baseline self-reported diabetes prevalence by adiposity	115
Figure 4.5. Baseline screen-detected diabetes prevalence by adiposity	116
Figure 4.6. Baseline diabetes prevalence by blood pressure.....	117
Figure 4.7. Frequency distribution of baseline random plasma glucose in individuals without self-reported diabetes at baseline	118
Figure 4.8. Mean baseline plasma glucose level by self-reported fasting time in individuals without self-reported diabetes at baseline	119
Figure 4.9. Baseline random plasma glucose by age in participants without self-reported diabetes at baseline	122
Figure 4.10. Associations of baseline random plasma glucose with adiposity measures in participants without self-reported diabetes at baseline	127
Figure 4.11. Associations of baseline random plasma glucose with blood pressure in participants without self-reported diabetes at baseline	128
Figure 4.12. Changes in mean random plasma glucose between baseline and first resurvey in five groups defined by baseline random plasma glucose in participants without self-reported diabetes at baseline	129
Figure 4.13. Changes in mean random plasma glucose between baseline and first and second resurveys in five groups defined by baseline random plasma glucose in participants without self-reported diabetes at baseline	131
Figure 4.14. Relationship between baseline and first resurvey random plasma glucose levels in participants without self-reported diabetes at baseline	132
Figure 5.1. Baseline prevalence of cardiovascular diseases by age and sex.....	148
Figure 5.2. Baseline prevalence of ischaemic heart disease by study area	149
Figure 5.3. Baseline prevalence of stroke/transient ischaemic attack by study area...	150
Figure 5.4. Odds ratios for prevalent cardiovascular diseases by self-reported diabetes status, progressively adjusted for potential confounding factors.....	154
Figure 5.5. Adjusted odds ratios for prevalent cardiovascular diseases by self-reported diabetes status, by study area.....	155
Figure 5.6. Adjusted odds ratios for prevalent ischaemic heart disease by self-reported diabetes status in population subgroups	156
Figure 5.7. Adjusted odds ratios for prevalent stroke/transient ischaemic attack by self-reported diabetes status in population subgroups	157
Figure 5.8. Adjusted odds ratios for prevalent cardiovascular diseases by duration of diabetes diagnosis	158
Figure 5.9. Adjusted odds ratios for prevalent cardiovascular diseases by use of hypoglycaemic medication in self-reported diabetes	159

Figure 5.10. Adjusted odds ratios for prevalent cardiovascular diseases by baseline random plasma glucose levels in participants without self-reported diabetes	161
Figure 5.11. Adjusted odds ratios for prevalent cardiovascular diseases per 1 mmol/L higher baseline random plasma glucose at random plasma glucose levels <11.1 mmol/L in individuals without self-reported diabetes, progressively adjusted for potential confounding factors.....	163
Figure 5.12. Adjusted odds ratios for prevalent cardiovascular diseases per 1 mmol/L higher baseline random plasma glucose at random plasma glucose levels <11.1 mmol/L in individuals without self-reported diabetes, by study area	164
Figure 5.13. Adjusted odds ratios for prevalent ischaemic heart disease per 1mmol/L higher baseline random plasma glucose at random plasma glucose levels <11.1 mmol/L in population subgroups among participants without self-reported diabetes...	165
Figure 5.14. Adjusted odds ratios for prevalent stroke/ transient ischaemic attack per 1mmol/L higher baseline random plasma glucose at random plasma glucose levels <11.1 mmol/L in population subgroups among participants without self-reported diabetes	166
Figure 5.15. Adjusted odds ratios for prevalent cardiovascular diseases by baseline random plasma glucose levels stratified by fasting time in participants without self-reported diabetes	167
Figure 6.1. Incidence of cardiovascular diseases by age and sex	189
Figure 6.2. Incidence of cardiovascular disease mortality by study area	190
Figure 6.3. Incidence of major occlusive vascular disease by study area	191
Figure 6.4. Incidence of major coronary event by study area	192
Figure 6.5. Incidence of ischaemic stroke by study area	193
Figure 6.6. Incidence of intracerebral haemorrhage by study area	194
Figure 6.7. Age- and sex-specific adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease by self-reported diabetes status.....	196
Figure 6.8. Age- and sex-specific adjusted hazard ratios for major coronary event and ischaemic stroke by self-reported diabetes status	197
Figure 6.9. Sex-specific adjusted hazard ratios for incident cardiovascular diseases by self-reported diabetes status	198
Figure 6.10. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease by self-reported diabetes status progressively adjusted for potential confounding factors	199
Figure 6.11. Adjusted hazard ratios for major coronary event and ischaemic stroke by self-reported diabetes status progressively adjusted for potential confounding factors	200
Figure 6.12. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease by self-reported diabetes status, by study area	201
Figure 6.13. Adjusted hazard ratios for major coronary event and ischaemic stroke by self-reported diabetes status, by study area	202
Figure 6.14. Adjusted hazard ratios for cardiovascular disease mortality by self-reported diabetes status in population subgroups	203

Figure 6.15. Adjusted hazard ratios for major occlusive vascular disease by self-reported diabetes status in population subgroups	204
Figure 6.16. Adjusted hazard ratios for ischaemic stroke by self-reported diabetes status in population subgroups.....	205
Figure 6.17. Adjusted hazard ratios for major coronary event by self-reported diabetes status in population subgroups.....	206
Figure 6.18. Age- and sex-specific adjusted hazard ratios for intracerebral haemorrhage by self-reported diabetes status	207
Figure 6.19. Adjusted hazard ratios for intracerebral haemorrhage by self-reported diabetes status, by study area.....	208
Figure 6.20. Adjusted hazard ratios for intracerebral haemorrhage by self-reported diabetes status progressively adjusted for potential confounding factors.....	209
Figure 6.21. Adjusted hazard ratios for intracerebral haemorrhage by self-reported diabetes status in population subgroups	210
Figure 6.22. Adjusted hazard ratios for cardiovascular diseases by duration of diabetes diagnosis.....	211
Figure 6.23. Adjusted hazard ratios for cardiovascular diseases by hypoglycaemic medication treatment status in self-reported diabetes	212
Figure 6.24. Adjusted hazard ratios for cardiovascular diseases by baseline plasma glucose level in self-reported diabetes	213
Figure 6.25. Adjusted hazard ratios for major occlusive vascular disease by number of cardiovascular risk factors at baseline among individuals with diabetes	215
Figure 7.1. Adjusted hazard ratios for cardiovascular diseases by baseline random plasma glucose	236
Figure 7.2. Adjusted hazard ratios for intracerebral haemorrhage by baseline random plasma glucose	237
Figure 7.3. Adjusted hazard ratios for cardiovascular diseases by usual random plasma glucose	238
Figure 7.4. Age- and sex-specific adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease per 1 mmol/L higher usual random plasma glucose	239
Figure 7.5. Adjusted hazard ratios for cardiovascular disease mortality per 1 mmol/L higher usual random plasma glucose in population subgroups	240
Figure 7.6. Age- and sex-specific adjusted hazard ratios for major coronary event and ischaemic stroke per 1 mmol/L higher usual random plasma glucose	242
Figure 7.7. Adjusted hazard ratios for ischaemic stroke per 1 mmol/L higher usual random plasma glucose in population subgroups.....	243
Figure 7.8. Adjusted hazard ratios for major occlusive vascular disease per 1 mmol/L higher usual random plasma glucose in population subgroups	244
Figure 7.9. Adjusted hazard ratios for major coronary event per 1 mmol/L higher usual random plasma glucose in population subgroups.....	245
Figure 7.10. Adjusted hazard ratios for intracerebral haemorrhage by usual random plasma glucose	246

Figure 7.11. Age- and sex-specific adjusted hazard ratios for intracerebral haemorrhage per 1 mmol/L higher usual random plasma glucose	247
Figure 7.12. Adjusted hazard ratios for intracerebral haemorrhage per 1 mmol/L higher usual random plasma glucose in population subgroups	248
Figure 7.13. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease by baseline random plasma glucose, by fasting time	249
Figure 7.14. Adjusted hazard ratios for major coronary event and ischaemic stroke by baseline random plasma glucose, by fasting time	250
Figure 7.15. Adjusted hazard ratios for intracerebral haemorrhage by baseline random plasma glucose, by fasting time	251
Figure 7.16. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease per 1 mmol/L higher baseline random plasma glucose progressively adjusted for potential confounding factors	252
Figure 7.17. Adjusted hazard ratios for major coronary event and ischaemic stroke per 1 mmol/L higher baseline random plasma glucose progressively adjusted for potential confounding factors.....	253
Figure 7.18. Adjusted hazard ratios for intracerebral haemorrhage per 1 mmol/L higher baseline random plasma glucose progressively adjusted for potential confounding factors	254
Figure 7.19. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease per 1 mmol/L higher baseline random plasma glucose, by study area	255
Figure 7.20. Adjusted hazard ratios for major coronary event and ischaemic stroke per 1 mmol/L higher baseline random plasma glucose, by study area.....	256
Figure 7.21. Adjusted hazard ratios for intracerebral haemorrhage per 1 mmol/L higher baseline random plasma glucose, by study area.....	257
Figure 8.1. “Triangulation” approach for calculation of predicted association of the random plasma glucose genetic risk score with cardiovascular disease risk	275
Figure 8.2. Distribution of random plasma glucose genetic risk score	278
Figure 8.3. Mean baseline random plasma glucose levels by random plasma glucose genetic risk score	279
Figure 8.4. Mean baseline random plasma glucose by categories of random plasma glucose genetic risk score	280
Figure 8.5. Adjusted hazard ratios for major occlusive vascular disease and major coronary event by categories of random plasma glucose genetic risk score.....	282
Figure 8.6. Adjusted hazard ratios for ischaemic stroke by categories of random plasma glucose genetic risk score	283
Figure D.1. Causal diagram demonstrating the relationships of diabetes and plasma glucose with cardiovascular disease	350
Figure D.2. Association of baseline ischaemic heart disease prevalence with adiposity.....	353
Figure D.3. Association of baseline ischaemic heart disease prevalence with blood pressure.....	354

Figure D.4. Association of baseline stroke/transient ischaemic attack prevalence with adiposity.....	357
Figure D.5. Association of baseline stroke/transient ischaemic attack prevalence with blood pressure	358
Figure D.6. Adjusted odds ratios for prevalent cardiovascular disease by baseline random plasma glucose in participants without self-reported diabetes	359
Figure E.1. Association of major occlusive vascular disease with adiposity.....	369
Figure E.2. Association of major occlusive vascular disease with blood pressure	370
Figure E.3. Association of intra-cerebral haemorrhage with adiposity	373
Figure E.4. Association of intra-cerebral haemorrhage with blood pressure	374
Figure E.5. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease by screen-detected diabetes status, by study area	376
Figure E.6. Adjusted hazard ratios for major coronary event and ischaemic stroke by screen-detected diabetes status, by study area	377
Figure E.7. Adjusted hazard ratios for intracerebral haemorrhage by screen-detected diabetes status, by study area.....	378
Figure E.8. Adjusted hazard ratios for major occlusive vascular disease by screen-detected diabetes status in population subgroups.....	379
Figure E.9. Adjusted hazard ratios for ischaemic stroke by screen-detected diabetes status in population subgroups.....	380
Figure E.10. Adjusted hazard ratios for cardiovascular disease mortality by screen-detected diabetes status in population subgroups.....	381
Figure E.11. Adjusted hazard ratios for major coronary event by screen-detected diabetes status in population subgroups	382
Figure E.12. Adjusted hazard ratios for intracerebral haemorrhage by screen-detected diabetes status in population subgroups	383
Figure F.1. Incidence of cardiovascular diseases by age and sex in prospective plasma glucose analyses population	386
Figure F.2. Incidence of cardiovascular disease mortality by study area in prospective plasma glucose analyses population.....	387
Figure F.3. Incidence of major occlusive vascular disease by study area in prospective plasma glucose analyses population.....	388
Figure F.4. Incidence of major coronary event by study area in prospective plasma glucose analyses population	389
Figure F.5. Incidence of ischaemic stroke by study area in prospective plasma glucose analyses population	390
Figure F.6. Incidence of intracerebral haemorrhage by study area in prospective plasma glucose analyses population	391
Figure F.7. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease by baseline random plasma glucose in men and women.....	392
Figure F.8. Adjusted hazard ratios for major coronary event and ischaemic stroke by baseline random plasma glucose in men and women	393
Figure F.9. Adjusted hazard ratios for intracerebral haemorrhage by baseline random plasma glucose in men and women	394

Figure F.10. Adjusted hazard ratios for cardiovascular diseases by baseline random plasma glucose 395

Figure F.11. Adjusted hazard ratios for intracerebral haemorrhage by baseline random plasma glucose 396

Figure F.12. Mean baseline plasma glucose levels by self-reported fasting time.....400

Figure F.13. Adjusted hazard ratios for cardiovascular diseases by fasting time-adjusted baseline random plasma glucose 401

Figure F.14. Adjusted hazard ratios for intracerebral haemorrhage by fasting time-adjusted baseline random plasma glucose 402

Figure G.1. Participant exclusions for genetic analyses 405

List of abbreviations

ADA	American Diabetes Association
AMI	Acute myocardial infarction
APCSC	Asia Pacific Cohort Studies Collaboration
ARIC	Atherosclerosis Risk in Communities
BMI	Body mass index
CCHS	Copenhagen City Heart Study
CDC	Centre for Disease Control and Prevention
CI	Confidence interval
CIMT	Carotid intima media thickness
CKB	China Kadoorie Biobank
CMCS	Chinese Multi-provincial Cohort Study
CNDMDS	China National Diabetes and Metabolic Disorders Study
CPS-II	Cancer Prevention Study-II
CT	Computed tomography
CTSU	Clinical Trial Service Unit and Epidemiological Studies Unit
CVD	Cardiovascular disease
DALY	Disability-adjusted life year
DBP	Diastolic blood pressure
DECODE	Diabetes Epidemiology: Collaborative analysis of Diagnostic criteria in Europe
DNA	Deoxyribonucleic acid
ECG	Electrocardiogram
EDTA	Ethylenediaminetetraacetic acid
ERFC	Emerging Risk Factors Collaboration
FBG	Fasting blood glucose
FPG	Fasting plasma glucose
GRS	Genetic risk score
GWAS	Genome-wide association study
HDL	High-density lipoprotein
HI	Health insurance
HIC	High-income countries
HR	Hazard ratio
HS	Haemorrhagic stroke
HSE	Health Survey for England
HWE	Hardy Weinberg equilibrium
ICD	International Classification of Diseases

ICH	Intracerebral haemorrhage
IFG	Impaired fasting glucose
IGT	Impaired glucose tolerance
IHD	Ischaemic heart disease
IS	Ischaemic stroke
KCPS	Korean Cancer Prevention Study
LD	Linkage disequilibrium
LDL	Low-density lipoprotein
LMIC	Low- and middle-income countries
MAGIC	Meta-Analyses of Glucose and Insulin-Related Traits Consortium
MCE	Major coronary event
MET	Metabolic equivalent of task
MI	Myocardial infarction
MOVD	Major occlusive vascular disease
MR	Magnetic resonance
MWS	Million Women Study
OGTT	Oral glucose tolerance test
OR	Odds ratio
PPV	Positive predictive value
RBG	Random blood glucose
RCT	Randomised controlled trial
RPG	Random plasma glucose
SAH	Subarachnoid haemorrhage
SBP	Systolic blood pressure
SD	Standard deviation
SE	Standard error
SNP	Single nucleotide polymorphisms
TIA	Transient ischaemic attack
TS	Total stroke
USA-PRC	USA-People's Republic of China Collaborative Study of Cardiovascular and Cardiopulmonary Epidemiology
WHO	World Health Organization

Chapter 1. Introduction

Cardiovascular diseases (CVD) are the leading cause of premature death, and in 2013 were estimated to account for about 17 million deaths worldwide,¹ which is expected to increase to over 23 million by 2030.² They are also the leading cause of morbidity worldwide.³ Although rates of CVD have been decreasing in many developed countries, they are increasing in many low- and middle-income countries (LMIC),⁴ including in China, where, in 2010, they accounted for over 3 million deaths.⁵

Diabetes has long been recognised as an independent risk factor for CVD in Western populations,⁶ but there is limited evidence about its relevance for CVD in China. Moreover, substantial uncertainty remains—in both Western and Chinese populations—about the association of blood glucose levels below the diabetes threshold with CVDs.⁶ This thesis aims to investigate the association of diabetes, and of blood glucose levels in individuals without diabetes, with the risk of major CVDs in Chinese adults, using cross-sectional and prospective data from the China Kadoorie Biobank (CKB), a nationwide prospective cohort study of over 0.5 million Chinese adults.⁷

This chapter provides a brief overview of CVDs and their epidemiology. It also provides an introduction to diabetes and its pathophysiology, and to glycaemia more generally, and their relationships with CVD. A detailed systematic review of the published evidence on these relationships is presented in Chapter 2.

1.1 Cardiovascular disease and its subtypes

CVD refers to diseases of the heart and blood vessels. It comprises a diverse group of conditions, including ischaemic heart disease (IHD), cerebrovascular disease (including stroke and transient ischaemic attack [TIA]), heart failure, cardiomyopathies, cardiac arrhythmias and sudden cardiac death, peripheral vascular disease, abdominal aortic aneurysm and congenital heart disease.⁸ This thesis focusses mainly on IHD and stroke, as these have the greatest impact on disease burden.^{1, 3}

1.1.1 Ischaemic heart disease

IHD includes several forms, most notably angina pectoris, acute myocardial infarction (AMI) and sudden cardiac death, which differ largely by severity and chronicity. Each results from reduced blood supply to cardiac muscle, due to coronary artery atherosclerosis.⁹

Pathophysiology

Atherosclerosis is a complex, multifactorial process. A detailed description of this is beyond the scope of this thesis, but a simplified description is presented here. Atherosclerosis critically involves damage to the arterial endothelium, with multiple consequences, including reduced release of nitric oxide, leading to impaired vascular relaxation,¹⁰ prothrombotic changes,¹⁰ adherence of inflammatory cells to the endothelium and increased endothelial permeability, facilitating migration of inflammatory cells, including monocytes, and lipids from the blood into the intima, or inner layer, of the arterial wall.^{10, 11} Once in the intima, monocytes evolve into macrophages which take up oxidised low-density lipoprotein (LDL) particles to form foam cells, ultimately resulting in formation of a fibro-fatty plaque in the vessel intima.^{10, 12} Smooth muscle cell migration from other vessel wall layers results in plaque expansion and production of extra-cellular matrix,¹³ increasing plaque size further, with protrusion into the vessel lumen and restriction of blood flow, leading to myocardial ischaemia.¹² The atheromatous plaque may become unstable and rupture, or become eroded,¹² leading to thrombus formation¹³ and more severe arterial occlusion, resulting in unstable angina or myocardial infarction (MI).¹²

Risk factors

The aetiology of IHD involves a complex interplay of genetic, lifestyle and environmental factors.¹⁴ Several risk factors have been identified, many of which cause endothelial damage, including advancing age,¹⁵ hypertension,¹⁶ dyslipidaemia,¹⁷⁻¹⁹ tobacco,²⁰ heavy alcohol use,²¹ physical inactivity,^{22, 23} elevated blood glucose levels,⁶ and adiposity.^{24, 25} Many of these risk factors are modifiable, providing scope for primary and secondary

prevention. In addition, multiple genetic variants associated with the risk of IHD have been identified,²⁶⁻²⁸ although each has only a small effect on IHD risk.²⁹

Clinical diagnosis

Diagnosis and subtyping of IHD in the clinical setting is typically through electrocardiograms (ECG) and measurement of cardiac enzymes, supplemented in some instances by more complex procedures including coronary angiography and stress testing.^{30, 31}

1.1.2 Stroke

Stroke is defined by the World Health Organization (WHO) as “rapidly developing clinical signs of focal (at times global) disturbance of cerebral function, lasting more than 24 hours or leading to death with no apparent cause other than that of vascular origin”,³² or, more simply, as “a clinical syndrome defined by acute neurological deficit (usually focal) with a vascular basis”.³³ Stroke is a complex syndrome comprising a heterogeneous group of conditions.³⁴ Broadly, stroke can be divided into ischaemic stroke (IS), resulting from occlusion or stenosis of arteries supplying the brain, or haemorrhagic stroke (HS), characterised by rupture of arteries supplying the brain resulting in bleeding into the cranial cavity.³⁵ There are various IS subcategories, defined by anatomical location³⁶ or underlying aetiology.³⁷ HS can be subdivided into intracerebral haemorrhage (ICH), characterised by bleeding into the parenchyma of the brain, and subarachnoid haemorrhage (SAH), in which bleeding occurs into the subarachnoid space surrounding the brain.^{34, 38}

Pathophysiology

IS results most commonly from thrombosis, due to atherothrombosis in major arteries of the brain or neck, or from embolism, where thrombus (typically) detaches from a lesion in another part of the cardiovascular system and causes arterial occlusion in the cerebrovascular system.³⁸ The underlying pathophysiology of IS is thus frequently similar to that of IHD. In contrast, ICH results from rupture of weakened cerebral blood vessels,³⁹ due mainly to hypertension or cerebral amyloid angiopathy, typically associated with

ageing.³⁵ SAH usually follows rupture of aneurysms (generally of unknown aetiology but also resulting from hypertension) or of vascular malformations.^{35, 38}

Risk factors

Given their established pathophysiologies, risk factor profiles of stroke subtypes differ somewhat. Risk factors for IS are comparable with those for IHD and include age, hypertension,¹⁶ smoking,²⁰ adiposity,²⁴ physical inactivity⁴⁰ and possibly dyslipidaemias,^{18, 19} as well as risk factors for embolism, such as atrial fibrillation and carotid artery atherosclerosis.³⁸ The main identified risk factors for ICH are hypertension,¹⁶ ageing, smoking,⁴¹ heavy alcohol consumption⁴² and coagulopathies.⁴³ SAH shares many of these risk factors, but additionally vascular malformations.³⁸ As for IHD, many risk factors for stroke are modifiable, providing opportunities for prevention. Genetic factors also play a role; genome-wide association studies (GWAS) have identified several genetic variants for stroke subtypes, but each with weak effect on risk.⁴⁴⁻⁴⁶

Clinical diagnosis

Stroke is usually diagnosed by computed tomography (CT) or magnetic resonance (MR) scanning or, in some fatal cases, by autopsy, which additionally enable subtyping. Accurate subtyping is essential for appropriate clinical management,³⁵ but also for accurate ascertainment of the distribution and determinants of stroke and stroke subtypes in epidemiological studies.⁴⁷

1.2 Epidemiology of cardiovascular diseases

1.2.1 International perspective

Worldwide, CVD is the major cause of death, disability and overall disease burden.¹ Estimates from the Global Burden of Disease Study placed IHD as the leading cause of death worldwide, accounting for over 8 million deaths, and stroke as the third leading cause, responsible for over 6 million deaths, in 2013.¹ IHD and stroke were similarly ranked for their impact on disability.³ While IS was estimated to account for the greatest proportion of

stroke deaths,¹ HS and other non-IS combined were estimated to account for more disability-adjusted life years (DALYs, a measure combining premature mortality, in the form of years of life lost, and years lived with disability) than IS, as a consequence, not of the relative frequency of stroke subtypes, but of the younger average age of onset of HS and thus more years of life lost per death.³

Globally, the burden of IHD and stroke, both IS and HS, has increased since 1990.^{1, 3, 48, 49} Population growth and ageing account for some of this increase, however, and the same period has seen reductions in age-standardised incidence, mortality and disability due to both IS and HS.⁴⁸ The global patterns and trends in CVDs mask interesting differences between populations. Since the 1980s, IHD mortality rates in high-income countries (HIC) have steadily declined.⁵⁰ Similar trends have been observed for stroke, and between 1990 and 2010 there were estimated 13%, 37% and 35% reductions in IS incidence, mortality and DALYs, and 19%, 38% and 39% reductions in the same measures for HS in these countries.⁴⁸ In contrast, many LMIC, including those of East Asia, experienced significant increases in age-standardised IHD mortality over this period.⁵⁰ For stroke, although mortality rates and disability have decreased in this time,⁴⁸ the incidence of IS, and possibly of HS, has increased modestly.⁴⁸ Consequently, LMIC were estimated to account for approximately 70%,⁵¹ 64% and 86% of the worldwide burden of disease attributable to IHD, IS and HS, respectively, in 2010.⁴⁸

Some of the reasons for these trends, and the contrasting patterns in HIC and LMIC, can be conceptualised in the epidemiological transition.⁵² Industrialisation and economic development are accompanied by shifts in patterns of disease.⁵³ In the context of CVD, this is been characterised by a transition from diseases resulting from infection, eg, rheumatic heart disease, and nutritional deficiency-related abnormalities of the cardiac muscle, to hypertension-related disorders, including HS and hypertensive heart disease, to atherosclerotic CVDs, including IHD and IS, frequently associated with ageing and with the typical lifestyle traits of more developed, industrialised countries.⁵³

1.2.2 Epidemiology of cardiovascular diseases in China

Consistent with the global picture, CVDs are the leading cause of death in China.⁵ Over recent decades age-standardised mortality from IHD has increased significantly.⁵⁴ Partly because of this, and population growth, the number of IHD deaths has more than doubled, from ~450 000 in 1990 to ~950 000 in 2010, as has disability due to IHD.⁵ Despite these trends, IHD is not the most common manifestation of CVD in China. Rather, that is stroke,⁵⁴ which accounted for an estimated ~1.7 million deaths and over 30 million DALYs in China in 2010, increasing from ~1.3 million deaths and ~25 million DALYs in 1990.⁵ Furthermore, ICH accounts for ~30% of all stroke events in China,⁵⁵ compared with ~10-15% in Western populations.⁵⁵ For IS, small vessel disease is thought to be more prominent in China than in many other populations.⁵⁶ These differences in patterns of CVDs between China and the West cannot be explained fully by known CVD risk factors.^{55, 57} However, China is at an earlier stage in the epidemiological transition; higher rates of stroke, particularly HS, reflect the influence of hypertension, and increasing rates of IHD reflect the adverse consequences of increasing development and urbanisation, and an ageing population.⁵⁴

Substantial variation in rates of CVDs exists across China.^{54, 55} Northern areas have been shown to experience higher rates of both IHD and stroke than southern areas,^{58, 59} with higher CVD rates also seen in urban, than rural, areas.^{54, 55} Higher levels of CVD risk factors, including dyslipidaemia, hypertension and obesity, in northern and urban regions account for some variation,⁵⁹ but some remains unexplained.^{55, 57} As the typical Western lifestyle becomes more widespread, disparities between areas will likely lessen,⁵⁹ with increasing levels of CVDs in rural and southern areas. With an expanding and aging population, marked increases in CVD burden in China are predicted over coming decades.⁶⁰ However, effective preventive measures may succeed in limiting this.⁵⁴

1.3 Diabetes, blood glucose and cardiovascular disease

1.3.1 Glucose metabolism and homeostasis

Under normal physiological conditions, glucose concentrations in the blood are tightly regulated, with fasting levels typically maintained between 3.5 mmol/L and 5.5 mmol/L and post-prandial levels below 8 mmol/L.⁶¹ This is achieved through regulation of glucose absorption from the intestine, storage and production in the liver and uptake in other organs and tissues.^{62, 63} These processes are, in turn, controlled by multiple hormones including insulin, glucagon, growth hormone, cortisol, thyroid hormones and catecholamines.^{62, 63} In diabetes, the most important hormone, and the only one with hypoglycaemic actions, is insulin.⁶³ Insulin is produced by the beta-cells of the pancreatic islets of Langerhans,⁶³ and when released stimulates uptake, storage and utilisation of glucose in most body tissues, but particularly in the liver, adipose tissue and skeletal muscle.⁶²

Diabetes

In diabetes mellitus (from hereon in referred to as diabetes), the tight control of blood glucose concentrations is lost, leading to chronic hyperglycaemia.⁶¹ Type 1 diabetes is characterised by auto-immune destruction of the pancreatic beta-cells, resulting in an absolute deficiency of insulin, whilst type 2 diabetes is characterised by relative insulin deficiency resulting from insulin resistance and inadequate compensatory insulin secretion.⁶⁴ Other forms of diabetes exist, but types 1 and 2 are most prevalent, accounting for 5-10%⁶⁴ and 85-95%, respectively.⁶⁵

A current global epidemic of diabetes is widely acknowledged.⁶⁶⁻⁶⁹ It was estimated that there were over 380 million people worldwide with diabetes in 2013, although in an estimated 175 million the condition remained undiagnosed.⁶⁵ The incidence and prevalence of both type 1 and type 2 diabetes have increased over recent decades,^{65, 70} and the number of people living with diabetes is predicted to increase further to over 590 million by 2035, with the greatest increase in LMIC.⁶⁵ Genetic factors play an important role

in determining diabetes risk, with over 90 genetic variants associated with type 1 or type 2 diabetes identified.^{71, 72} Even collectively, however, these variants account for only a small proportion of diabetes risk,⁶⁸ and cannot account for the escalation in diabetes incidence and prevalence over recent decades. Relatively little is known of the environmental risk factors for type 1 diabetes, but the non-genetic determinants of type 2 diabetes are better understood and include age,⁶⁵ adiposity,^{73, 74} physical inactivity⁷⁵ and smoking.⁷⁶ In addition to population growth and ageing, increased prevalence of these risk factors provides some explanation for the escalating burden of diabetes over recent decades.^{67, 68}

Diagnosis of diabetes requires demonstration of persistently elevated blood glucose levels beyond certain thresholds (Table 1.1). However, this represents one extreme of a continuous glycaemic range. Glycaemic thresholds for diagnosing diabetes were defined based upon an elevated risk of microvascular diabetes complications, particularly retinopathy,⁷⁷ but blood glucose levels below the diabetic threshold may have important implications for health. Pre-diabetes refers collectively to the states of impaired glucose tolerance (IGT) and impaired fasting glucose (IFG), in which blood glucose levels are above the strictly normal range, but below the diabetic threshold (Table 1.1).^{64, 78} Individuals with blood glucose levels in this range have been shown to be at increased risk of future development of diabetes.⁶⁴

In addition to escalating diabetes prevalence worldwide, mean levels of fasting plasma glucose (FPG) have increased steadily, rising by 0.07 mmol/L and 0.09 mmol/L per decade in men and women, respectively, since 1980.⁷⁹ Despite regional differences, the increase in age-standardised mean FPG levels was evident in most regions.⁷⁹ Both genetic and environmental factors influence blood glucose levels in individuals without diabetes, with some overlap with those factors influencing diabetes.⁸⁰⁻⁸⁴

Diabetes in China

As recently as the 1980s, diabetes was said to be rare in China.⁸⁵ However, due, at least in part, to population size and ageing, and an increasingly obesogenic environment,⁸⁶ China now has the world's largest diabetes population.⁶⁵ A nationally representative survey in 2010 found that over 10% of Chinese adults had diabetes.⁸⁷ Diabetes in China has typically affected leaner individuals, with more marked pancreatic beta-cell dysfunction^{88, 89} and greater susceptibility to microvascular diseases and cancer, than to the characteristic cardiovascular complications seen in the West.^{86, 88} Although considerable literature exists on the prevalence of diabetes in China, there is none describing the distribution of blood glucose levels in the Chinese population.

1.3.2 Diabetes and cardiovascular disease

CVD is the most common cause of death among people with diabetes.⁶⁵ Studies in predominantly Western populations have shown independent two- to three-fold increases in the risk of ischaemic CVDs in individuals with diabetes,^{6, 90, 91} but less clear associations with HS.^{6, 90, 91}

Various biological mechanisms underlie the association of diabetes with increased risk of ischaemic CVDs. Higher concentrations of blood glucose have been shown to induce endothelial dysfunction,^{92, 93} impair nitric oxide synthesis^{92, 94, 95} and promote thrombosis⁹² and inflammation.⁹²⁻⁹⁵ The processes involved are complex and incompletely understood, but are thought to include production of advanced glycation end-products and increased activity in the polyol and hexosamine pathways.^{92, 94, 96, 97} Damage to the vasa vasorum (small blood vessels supplying walls of larger vessels) of large arteries may also contribute to atherosclerosis.⁹⁸ The association of diabetes, especially type 2 diabetes, with other known CVD risk factors, including hypertension and dyslipidaemia,^{95, 97} also contributes to the elevated CVD risk in diabetes. These associations, particularly that with hypertension,

may provide a mechanism through which diabetes influences the risk of non-atherosclerotic CVDs.⁹⁹

1.3.3 Blood glucose and cardiovascular disease

For the purposes of clinical management, CVD risk factors, including blood pressure and cholesterol, are frequently dichotomised. However, this may overlook important cardiovascular risk-related information below the frequently arbitrary clinical thresholds.⁵³ This may also be the case for blood glucose.¹⁰⁰ The same pathophysiological mechanisms underlying the associations of diabetes with CVDs are potentially relevant to higher blood glucose levels below the diabetic threshold,¹⁰¹ although epidemiological evidence on their associations with the risk of CVDs has been inconsistent.^{6, 102-106}

Diabetes is an important risk factor for CVDs, particularly ischaemic CVDs, in Western populations.⁶ Despite the increasing prevalence of diabetes⁸⁵⁻⁸⁷ relatively little is known of its role in CVD in China, where CVD patterns typically differ importantly from those in the West.^{54, 55} Moreover, in China and elsewhere, the nature of the relationship between blood glucose levels and CVDs remains controversial.^{6, 103} Large-scale, population-based studies are needed to enable robust characterisation of these relationships in China. In turn, this will enable effective CVD prevention, not only in China, but more widely.

Table 1.1. Diagnostic criteria for diabetes and pre-diabetes^{64, 78, 107}

Glycaemic category	Diagnostic criteria
Diabetes	• Fasting plasma glucose: ≥ 7.0 mmol/L • 2-hour post-load plasma glucose: ≥ 11.1 mmol/L^a • HbA1c: $\geq 6.5\%$ • Random plasma glucose: ≥ 11.1 mmol/L in a patient with classic symptoms of hyperglycaemia or hyperglycaemic crisis
Impaired glucose Tolerance	• 2-hour post-load plasma glucose: ≥ 7.8 and < 11.1 mmol/L^a
Impaired fasting glucose	• Fasting plasma glucose: - WHO guidance 6.1-6.9 mmol/L - ADA guidance 5.6-6.9 mmol/L

^aBased on 75g oral glucose tolerance test.

ADA, American Diabetes Association; WHO, World Health Organization.

Chapter 2. Literature review

2.1 Objectives

The objectives of this literature review were to identify, appraise and assimilate published evidence on the relationship of diabetes, and, among individuals without known diabetes, of blood glucose levels, with both prevalent and incident IHD and stroke. This was with the aim of identifying current gaps in knowledge to inform new research based on data from the CKB. An additional review was undertaken of Mendelian randomisation studies of the association between blood glucose and CVD.

2.2 Methodology

Five individual searches were undertaken using MEDLINE and Embase databases. The search terms used are detailed in Table 2.1, and the numbers of articles identified are presented in Table 2.2. Two additional studies with papers published in Chinese were identified, examining the association of diabetes with prevalent CVD. For practical reasons, and due to small sample sizes and study design limitations (a cross-sectional study including 7577 participants¹⁰⁸ and a case-control study including 315 stroke cases and 117 controls¹⁰⁹), these were not included in the review.

2.3 Diabetes and cardiovascular disease

2.3.1 Cross-sectional and case-control evidence

Ischaemic heart disease

10 articles describing the association between diabetes and prevalent IHD were identified, including four cross-sectional studies,¹¹⁰⁻¹¹³ five case-control studies¹¹⁴⁻¹¹⁸ and one meta-analysis¹¹⁹ (Tables 2.3 and 2.4). Two articles reported on the INTERHEART case-control study, one focussing on the Chinese study population¹¹⁵ and one describing findings from the entire study;¹¹⁷ both were included due to the particular relevance of findings from the Chinese study population.

Overall, the included studies suggest approximately two- to four-fold increases in the odds of prevalent IHD associated with diabetes, generally higher in women than men.^{111, 112, 114, 116, 119} The non-prospective design of the studies greatly limits the reliability of the evidence obtained, particularly due to lack of temporal clarity between exposure and outcome, and to reverse causality and selection biases,^{111, 112, 114} likely inflating risk estimates.

Three identified studies described the association of diabetes with prevalent IHD in Chinese populations.^{110, 113, 115} In a cross-sectional survey of 58 308 individuals (mean age 56 years) from rural Beijing, 6.1% of participants had a history of IHD (based on self-report and medical record review) and 8.6% had diabetes, defined as self-reported current treatment with insulin or oral hypoglycaemic medication.¹¹⁰ Diabetes was associated with an adjusted odds ratio (OR) of 2.51 (95% confidence interval [CI] 2.29-2.75) for IHD. This is consistent with findings from the China National Diabetes and Metabolic Disorders Study (CNDMDS), a nationally representative survey of over 45 000 men and women aged 20 years or older.¹¹³ This study found that diabetes—defined through self-report or oral glucose tolerance testing (OGTT)—was associated with an OR of 2.44 (1.31-4.53) for self-reported IHD. In contrast, in the large INTERHEART case-control study, among the Chinese population subgroup, self-reported diabetes was associated with an adjusted OR of 5.10 (4.03-6.33) for hospital-diagnosed AMI.¹¹⁵ INTERHEART-China included ~3000 first AMI cases from 26 centres and a similar number of age-, sex- and location-matched hospital and community controls, with no history of IHD or exertional chest pain. The OR observed in China was much greater than in the other 51 INTERHEART countries combined (2.84, 2.61-3.09, *p* for heterogeneity <0.0001), which included over 9000 AMI cases and 11 000 controls.¹¹⁷ The higher diabetes-associated odds in China were attributed to greater severity of diabetes, compared with other countries, resulting from later diagnosis and/or less adequate management.¹¹⁵ The IHD cases included in the cross-sectional studies in China, in which the ORs were more modest,^{110, 113} may be less severe or more susceptible to misclassification than the AMI cases in INTERHEART.^{110, 113} Similarly, inclusion of screen-detected cases in the diabetes

definition in one study,¹¹³ or differences in study populations may partly account for the different risk estimates.

A cross-sectional study of 141 400 Swedish men and women aged 45-74 years reported risk estimates which were reasonably consistent with those from the cross-sectional Chinese studies.¹¹¹ This study, based mainly on administrative healthcare data, included 10 364 individuals with diabetes, assessed through plasma glucose measurement, and over 11 000 IHD events. The age-adjusted prevalence rate ratios of IHD associated with diabetes were 2.65 (95% CI 2.52-2.79) in men, and 3.97 (3.69-4.27) in women, with attenuation of the sex difference at older ages. In another small cross-sectional study of 5000 Jordanian men and women (mean age 50.1 and 54.1 years, respectively) with 1534 IHD cases and 1410 individuals with diabetes (based on a self-reported diagnosis or current use of hypoglycaemic medication), men with diabetes had 2.9-fold (2.93, 2.50-3.44) greater odds of self-reported IHD, with 5.6-fold greater odds (5.59, 4.33-7.23) in women.¹¹² These ORs appear to have been adjusted only for age, smoking and hypertension.

The remaining four articles related to case-control studies. A meta-analysis of individual participant data from three Italian hospital-based studies included 1737 cases and 2317 controls, and found that self-reported diabetes was associated with a just greater than doubling of the odds of non-fatal MI (OR 2.3, 95% CI 1.8-2.9) after extensive adjustment.¹¹⁹ Additional analyses found non-significantly higher diabetes-associated odds in individuals with a longer duration of diabetes and with diabetes diagnosed at a younger age. The Japanese Acute Coronary Syndrome Study comprised 1925 hospital-diagnosed AMI cases, and 2279 age- and sex-matched controls, recruited from a single hospital.¹¹⁴ It reported an unadjusted OR of 3.44 (95% CI 2.50-4.75) for AMI associated with diabetes, assessed through fasting blood glucose (FBG) measurement. A smaller, hospital-based study in Qatar examined the association of self-reported diabetes with MI in 512 cases—ascertained through measurement of cardiac enzymes—and 382 controls, and reported a diabetes-associated OR of 4.01 (2.51-6.42) for MI in a predominantly (78%) male

population.¹¹⁸ Finally, the population-based Stockholm Heart Epidemiology Program Study examined risk factors for MI amongst 2246 cases and ~3000 age-, sex- and area-matched controls who had no record of MI.¹¹⁶ Diabetes was defined based on self-report, use of diabetes medications or FBG measurement, and was shown to be associated with an OR for MI of 2.45 (1.86-3.22) in men and 4.36 (2.72-7.01) in women, after adjustment for smoking, overweight and job strain.

Limitations of the evidence and implications for future study

The definitions of diabetes employed in these previous studies varied, and included self-reported diabetes,^{115, 117-119} self-reported diabetes on medication,¹¹⁰ measured blood glucose consistent with diabetes,^{111, 114} or various combinations of these.^{112, 113, 116} The definition used could impact on diabetes-associated risk estimates. A self-reported diagnosis would be expected to be less specific than self-reported diabetes on medication, but the latter approach would exclude individuals with diabetes managed through lifestyle interventions alone. Diabetes defined through a combination of self-report and blood glucose measurement would be more sensitive, but may include individuals with false positive diagnoses, the extent of which would depend upon the glycaemic measure used. Conversely, use of blood glucose measurement alone, such as in the Swedish study,¹¹¹ would fail to capture well-controlled diabetes cases. Whilst comparisons of included studies using different diabetes definitions do not reveal systematic differences in risk estimates, additional differences in study methods, populations and disease outcomes complicate these comparisons.

Included case-control studies investigated MI¹¹⁴⁻¹¹⁹ whereas the cross-sectional studies examined broader IHD outcomes.¹¹⁰⁻¹¹³ MI was mainly hospital-diagnosed through variable combinations of symptoms, ECG, cardiac enzyme measurement and autopsy,¹¹⁴⁻¹¹⁸ whilst IHD was mainly determined through self-report,^{110, 112, 113} supplemented in one study by review of medical records.¹¹⁰ As such, disease outcomes included in the case-control studies were arguably more specific and severe, which may explain the stronger

associations shown in some of these studies. Conversely, some case-control studies used hospital controls, and the absence of IHD in controls was usually determined through self-report,^{114, 115, 117} raising the possibility of misclassification.

Whilst these cross-sectional and case-control studies provide a valuable starting point for understanding the association of diabetes with IHD, there is clearly a need for more robust prospective evidence on the association.

Stroke

10 articles were identified which reported associations of diabetes with prevalent stroke. These included three cross-sectional studies^{110, 113, 120} and seven case-control studies^{118, 121-126} (Tables 2.5 and 2.6). Three studies were in Chinese populations.^{110, 113, 124} The studies were generally small, and were again limited by lack of temporal clarity between exposure and outcome, and susceptibility to reverse causality and selection biases. Studies mostly reported on total stroke (TS) outcomes,^{110, 113, 118, 120-122, 124} showing inconsistent, but generally positive, associations with diabetes.^{110, 118, 120-122, 124} The findings of the limited studies reporting on stroke subtypes were suggestive of a positive association of diabetes with prevalent IS,^{122, 126} but not HS.^{123, 125, 126}

The two previously described nationally representative, cross-sectional surveys in China additionally examined the association of diabetes with prevalent stroke, with quite different findings.^{110, 113} Whilst the rural Beijing study of 58 308 men and women, including an unspecified number of TS events, found an approximate doubling of the odds of self-reported TS associated with diabetes (OR 2.24, 95% CI 1.98-2.52),¹¹⁰ the CNDMDS, of ~46 000 participants and including 472 TS events found no significant association (0.97, 0.59-1.61),¹¹³ which contrasts with the greater than two-fold increased odds of IHD demonstrated in this study. More extensive adjustment for confounding and a less specific diabetes definition in the latter study may have contributed to the null findings, although the requirement for CT/MR scan confirmation of TS diagnoses might have been expected to

have resulted in a stronger association with diabetes.¹¹³ The third Chinese study comprised ~1000 TS cases and an equal number of age-, sex-, ethnicity- and area-matched controls from five communities in Chongqing, South West China.¹²⁴ Consistent with the rural Beijing study, there was an overall approximate doubling of the odds of TS associated with self-reported diabetes, with higher odds in women, than in men (3.65, 2.34-5.69 vs. 1.81, 1.14-2.88), after adjustment for most established potential confounding factors, but not for smoking.

A Cuban cross-sectional study—the 10/66 Population Based Study—examined the association of diabetes with TS in almost 3000 men and women aged 65 years or older, of whom 229 had a history of stroke, ascertained through self-report, physical examination or CT scan.¹²⁰ After extensive adjustment, the odds of TS were approximately 2.5-times greater (OR 2.6, 95% CI 2.0-3.5) in individuals with self-reported diabetes compared to those without diabetes. The remaining six studies used a case-control design.^{118, 121-123, 125,}

¹²⁶ These included a small UK-based study, which used routine primary care data to identify 244 women aged 45 to 69 years with a recorded stroke diagnosis, and 488 age- and general practice-matched controls.¹²¹ There was a striking 10-fold (10.5, 4.0-27.9) greater odds of TS associated with diabetes. However, the estimate was minimally adjusted and the CI was extremely wide. The previously described Qatari study also included 262 scan-confirmed TS cases. Comparing these with controls with no history of CVD, diabetes was associated with an OR of 5.12 (3.11-8.60) for TS, but again adjustment for confounding was minimal.¹¹⁸

Most of the studies described above did not report on the association of diabetes with stroke subtypes, despite frequent use of CT or MR scans for diagnosis.^{110, 113, 118, 120, 124} In the few studies which did report on associations with stroke subtypes, the results were inconsistent.^{122, 123, 125, 126} The INTERSTROKE case-control study included over 2000 IS cases and a similar number of age, sex and area-matched hospital and community controls, and was the largest of its kind reporting on the association with diabetes.¹²² It demonstrated 60% (99% CI 29-99%) higher odds of IS in individuals with self-reported diabetes compared

with those without, after adjustment for hypertension, smoking, alcohol intake, physical activity, diet, obesity and psychosocial factors. This was lower than the OR of 2.5 ($p < 0.0002$) observed in a Canadian case-control study of ~400 IS cases—ascertained from routine data and subsequent medical record review—and the same number of hospital-based age-, sex- and residence-matched controls.¹²⁶ Secular changes, for example in diabetes management, may provide some explanation for the lower odds in the INTERSTROKE study which was initiated in 2007,¹²² almost 30 years later than the Canadian study.¹²⁶

Three small studies, each with just over 100 cases, reported on the association of diabetes with prevalent HS.^{123, 125, 126} The same Canadian case-control study included 137 HS cases (ICH, SAH or haemorrhage of unknown origin), and found no significant association with diabetes (ascertained through medical record review) (OR 1.3, $p \geq 0.05$).¹²⁶ A similarly-sized OR was found in a French population-based case-control study of 130 individuals with cerebral haemorrhage and 130 age- and sex-matched healthy controls, although the extent of adjustment was unclear.¹²³ Finally, in an Indian hospital-based case-control study including 166 scan-confirmed ICH cases and the same number of age- and sex-matched controls, diabetes was associated with more than doubling of the odds of HS (2.39, 0.75-7.59), but this was not significant due to the small sample size.¹²⁵

Limitations of the evidence and implications for future study

Many of the described limitations of studies examining the association of diabetes with prevalent IHD apply to these studies of diabetes and prevalent stroke. Small sample sizes in many studies^{118, 120, 121, 123-126} may explain the relatively inconsistent findings, and there was again a lack of temporal clarity between exposure and outcome. Similarly, whilst the majority of studies ascertained diabetes through self-report alone,^{110, 118, 120, 122, 124} this was not universal.^{113, 121, 123, 125} Moreover, there was little assessment of stroke subtypes or of potential effect modification, apart from by sex.¹²⁴ This review of existing evidence thus

clearly highlights the importance of prospective analyses of the relationships of diabetes with incident stroke risk, including stroke subtypes.

2.3.2 Prospective evidence

Ischaemic heart disease

The search identified nine articles describing the association of diabetes with risk of incident IHD; seven were from cohort studies,^{90, 127-132} including three in China,¹³⁰⁻¹³² and two were meta-analyses^{6, 133} (Tables 2.7 and 2.8). Several large, well-designed studies were identified; despite differences in study protocols, including diabetes definitions, exclusion criteria and IHD definitions, they provide reasonably consistent evidence across populations—including more limited Chinese study populations—of an approximate doubling of the risk of IHD in diabetes.

In the Chinese Multi-provincial Cohort Study (CMCS) of over 30 000 men and women with no prior history of CVD from 11 Chinese provinces, diabetes, defined through self-report or blood glucose measurement, was associated with a hazard ratio (HR) of 1.81 (95% CI 1.20-2.71) for IHD, defined as AMI or coronary or sudden death.¹³¹ Although adjusted for multiple potential confounders, socioeconomic status, alcohol consumption and blood pressure were not considered, and residual confounding may have been substantial. In a prospective study of almost 67 000 Chinese men and women aged 65 years or older from Hong Kong, who were followed up for about 11 years, self-reported diabetes was associated with more than two-fold increased risk of IHD mortality in men (2.20, 1.89-2.57) and women (2.25, 1.98-2.55), after adjustment for several potential confounders, but not for blood pressure.¹³⁰ In the smaller USA-People's Republic of China Collaborative Study of Cardiovascular and Cardiopulmonary Epidemiology (USA-PRC), including 9903 men and women aged 35-59 years and without prior MI or stroke from Beijing and Guangzhou, there was a three-fold (2.99, $p=0.0145$) increased risk of IHD associated with diabetes (defined as a self-reported history of diabetes on medication or a FBG level ≥ 7.0 mmol/L) in women, but not in men

(1.51, $p=0.4318$).¹³² CIs and SEs were not provided for risk estimates in this study, but the non-significant findings in men may chiefly reflect the relatively small number of IHD events ($n=105$), as well as low prevalence of diabetes (2%) at the time of recruitment in the early 1980s.^{130, 131}

Western studies have found diabetes-associated IHD risks consistent with those of later Chinese studies, including the prospective US Cancer Prevention Study-II (CPS-II) which included over one million men and women aged 30 years and older. Over 50 000 participants reported a history of diabetes,¹²⁷ and more than 100 000 IHD deaths were recorded during an average 26-year follow-up. IHD mortality among individuals with diabetes was more than double that of individuals without diabetes, with slightly more extreme excess risk in women (adjusted HR 2.46, 95% CI 2.39-2.53) than in men (2.11, 2.05-2.16). Similarly, in the UK Million Women Study (MWS) of over 1.2 million women, with mean age 56 years at baseline, there was a 2.4-fold (2.44, 2.32-2.57) increased risk of IHD associated with diabetes.¹²⁹ The study, which recruited between 1996 and 2001, included almost 30 000 women with self-reported diabetes among those without prior CVD or pulmonary embolism, and, after six years of follow-up, 21 928 IHD events (fatal and non-fatal) were recorded. Another UK study, using routine general practice data from a representative sample (with regards to age and sex) of the UK population, demonstrated approximately two-fold (2.13, 2.01-2.26) and three-fold (2.95, 2.75-3.17) increased risks of MI associated with diabetes in men and women, respectively.¹²⁸ These estimates were based on 235 640 individuals aged 35-89 years and without a prior history of MI or coronary revascularisation (including over 40 000 with a recorded diagnosis of type 2 diabetes), of whom 9058 experienced a first MI during nine years of follow-up. A recent, much larger UK study found a somewhat lower diabetes-associated risk of MI (1.54, 1.42-1.67) after extensive adjustment for potential confounders.⁹⁰ This study was based on data from over 1.9 million participants without prior CVD in the CALIBER programme, linking four electronic health data sources, and including ~16 000 non-fatal MI events.⁹⁰ Similar risk estimates

were reported for other IHD outcomes in this study, including stable (1.62, 1.49-1.77) and unstable (1.53, 1.32-1.76) angina and unheralded coronary death (1.43, 1.23-1.65).⁹⁰ Censoring of individuals at the time of the first of any of the 12 investigated cardiovascular endpoints, including TIA, stable angina and peripheral vascular disease, may explain some of the lower risk estimated in this study.

Two meta-analyses examined the association of diabetes with incident IHD.^{6, 133} The Emerging Risk Factors Collaboration (ERFC) individual participant data meta-analysis included ~530 000 participants without prior IHD or stroke from 102 prospective studies in mainly Western, and no Chinese, populations.⁶ Approximately 39 000 participants had a known history of diabetes at baseline, and over 26 000 experienced an incident IHD event during an average 13.7-year follow-up. Diabetes was associated with an approximate doubling of the risk of incident fatal and non-fatal IHD (HR 2.00, 95% CI 1.83-2.19) and of non-fatal MI (1.82, 1.64-2.03) after adjustment for age, sex, smoking, blood pressure, body mass index (BMI) and lipids. The second meta-analysis, combining published data from 20 studies with individual participant data from four studies (one of which included data from 36 individual cohorts), included over 850 000 individuals and 28 000 IHD events.¹³³ Individuals with diabetes had more than two-fold increased risk of incident fatal and non-fatal IHD, with a more extreme adjusted HR in women (2.82, 2.35-3.38) than in men (2.16, 1.82-2.56). Comparable estimates were reported for fatal IHD (men: 2.04, 1.72-2.43; women: 2.83, 2.25-3.54). This study included several East Asian, including Chinese, studies, although the precise number is unclear.

Limitations of the evidence and implications for future study

Individual cohort studies most commonly defined diabetes through self-report^{127, 129, 130} or a combination of self-report and FBG measurement.^{131, 132} As described, diabetes definitions may impact on associated IHD risks. Although most studies did not address the validity of diabetes diagnoses,^{128, 130, 131} a small number described reasonable validity of self-reported diabetes.^{127, 129} Diabetes diagnoses derived from routine health services data^{90, 128} may be

less accurate, although diabetes-associated risks based on different definitions (recorded diagnoses or HbA1c measurement) in the CALIBER study were comparable.⁹⁰

Ascertainment of IHD events varied between studies, and included individually validated events,¹³¹ IHD mortality based on death registry data,^{127, 130} and non-fatal IHD events determined through linkage with routine hospital datasets without specific validation.^{128, 129} Despite differences in IHD definition, the strength of the associations appeared to be reasonably consistent between studies, showing at least two-fold increased risk of IHD associated with diabetes,^{6, 90, 127-133} although the ERFC meta-analysis showed a moderately greater excess risk for fatal (HR 2.31, 95% CI 2.05-2.60), than for non-fatal (1.82, 1.64-2.03), IHD.⁶

Despite inclusion of several large studies, investigation of effect modification was limited mainly to sex and age. The largest study, the CALIBER study based on routine healthcare data, showed no effect modification by sex.⁹⁰ Many other studies, however, demonstrated approximately 20-40% higher diabetes-associated risks of IHD in women than in men.^{6, 127, 128, 133} Several studies demonstrated a stronger association of diabetes with IHD at younger ages,^{6, 90, 128, 130} and one meta-analysis found higher diabetes-associated IHD HRs in non-smokers than in smokers, and at below average BMI and systolic blood pressure (SBP).⁶ The three Chinese studies¹³⁰⁻¹³² reported diabetes-associated IHD risks broadly similar to those reported in Western studies.^{6, 90, 127-129} However, in comparison with Western studies, the studies were relatively small¹³⁰⁻¹³² and were established in the 1980s¹³² and 1990s,^{130, 131} when rates of diabetes were much lower (eg, <1% prevalence in 1980) than currently.⁶⁸

Two earlier Chinese studies described relatively high loss to follow-up (14%¹³¹ and 5%¹³²). They did not address the likelihood of resulting bias, although the risk estimates in these studies did not clearly differ from those with more complete follow-up. Whilst acknowledging the potential limitations of the use of routine data sources for morbidity and mortality follow-

up, a potential advantage of such data is lower loss to follow-up, as demonstrated (0.2%),¹²⁷ or assumed,¹²⁸⁻¹³⁰ in several studies.

The described studies and their limitations demonstrate the need for large, contemporary, prospective studies in Chinese populations, particularly given the escalating prevalence of diabetes in China.⁸⁶ With the exception of the ERFC, investigation of effect modification has been limited, and warrants investigation specifically in Chinese populations in whom relevant factors, such as adiposity, blood pressure and physical activities, frequently differ in comparison with Western populations.

Stroke

12 studies were identified which described the association of diabetes with incident stroke risk, including 10 cohort studies^{90, 127, 129-132, 134-137}—two in mainland China,^{130-132, 134} two in Hong Kong^{130-132, 134} and one in Taiwan¹³⁶—and two meta-analyses^{6, 138} (Tables 2.9 and 2.10). Despite investigation of the relationship between diabetes and incident stroke in several large studies, understanding of the relationship is less well-advanced than of that between diabetes and IHD, and this is particularly true of stroke subtypes. The available evidence suggests the diabetes-associated risk of IS^{6, 90, 129, 131, 136} is broadly consistent with the approximate doubling of risk seen for IHD. Associations between diabetes and HS have been much less consistent, but generally weaker.^{6, 90, 129}

Six studies of diabetes and stroke reported on the relationship with TS only.^{127, 130, 134, 135, 137, 138} The previously described CPS-II included almost 35 000 TS deaths, and found a significantly increased risk in men (HR 1.87, 95% CI 1.77-1.97) and women (1.70, 1.61-1.80) with self-reported diabetes, after adjustment for potential confounding factors.¹²⁷ Another smaller North American study included ~74 000 men and women aged 25 years or older from a single province of Canada of whom almost 600 died or were hospitalised due to stroke.¹³⁷ After adjustment for area of residence, socioeconomic status and a general comorbidity score, and exclusion of individuals with a CVD diagnosis in the previous three

years, diabetes was associated with a 62% (1.62, 1.35-1.94) increased risk of TS. A study based on over 80 000 Health Survey for England (HSE) participants without prior CVD, ascertained TS mortality through linkage to national death registry data. Self-reported diabetes was associated with an 87% (1.87, 1.28-2.73) increased risk of TS mortality.¹³⁵ Although comparable with other studies, minimal adjustment for potential confounders (age and sex only) likely over-estimated diabetes-associated risks. The previously described study of 66 820 elderly Hong Kong residents recorded almost 2000 stroke deaths during approximately 11 years of follow-up.¹³⁰ After adjusting for potential confounders, there was a 58% increased risk of stroke mortality (1.58, 1.40-1.77) associated with self-reported diabetes. Risks were comparable in men and women (1.53, 1.26-1.85 vs, 1.60, 1.38-1.85, respectively) and were attenuated with increasing age. A further Hong Kong cohort study included just 304 participants and seven stroke events and thus provided little valuable information on the association.¹³⁴ Although many individual cohort studies showed comparable diabetes-associated risks in men and women, this contrasts with the findings of a meta-analysis of a combination of published and individual participant data undertaken with the specific aim of investigating sex-differences. Including over 775 000 participants from 64 cohorts, this study found a greater diabetes-associated TS (fatal or non-fatal) risk in women (2.28, 1.93-2.69), than in men (1.83, 1.60-2.08), based on approximately 12 500 stroke events.¹³⁸

Reporting only of TS risks in some studies may reflect the relatively small numbers of stroke events included,^{130, 134, 135, 137} or lack of adequate scanning facilities, particularly in older studies. Investigation of the association of diabetes with TS may mask associations with stroke subtypes, but only six studies reported separately on the association of diabetes with IS or HS.^{6, 90, 129, 131, 132, 136} Several Chinese studies described the association of diabetes with IS risk. The USA-PRC included 266 fatal or non-fatal IS events, 81% of which were confirmed by CT or MR scan,¹³² and found almost three-fold excess risk in women (n=108, HR 2.89; p=0.003) but no association in men (n=158, 1.08; p=0.8). In the later CMCS of

~30 000 participants (and an unspecified number of IS events) there was a two-fold increased risk of IS (2.06, 95% CI 1.54-2.76) associated with diabetes defined through self-report and blood glucose measurement.¹³¹ A much larger study, based on National Health Insurance (HI) claim data in Taiwan, included over 1 million participants and 100 000 fatal and non-fatal IS events over 6 years of follow-up.¹³⁶ Analyses showed a 2.3-fold increased risk (2.32, 2.27-2.36) of IS associated with diabetes in women and 1.9-fold increased risk (1.89, 1.85-1.92) in men, and risks were also higher at younger ages. These risk estimates were not adjusted for important confounding factors, reflecting the limitations of the routine data source. The study did not exclude individuals with a prior history of stroke, and reverse causality may have contributed to the association.

Several major studies in Western populations have also investigated diabetes-associated IS risk. The large CALIBER study recorded ~6000 fatal and non-fatal IS events, and found a 72% excess risk of IS (HR 1.72, 95% CI 1.52-1.95) associated with diabetes, with no sex-difference.⁹⁰ In the MWS, with 2572 IS cases after an average 6-year follow-up, diabetes was associated with a 2.5-fold (2.55, 2.19-2.96) increased risk of fatal and non-fatal IS after extensive adjustment for potential confounding factors.¹²⁹ In the ERFC meta-analysis, including ~3800 IS events, diabetes was associated with a 2.3-fold (2.27, 1.95-2.65) increased risk of fatal and non-fatal IS, with a greater excess risk in women (2.83, 2.35-3.40) than in men (2.16, 1.84-2.52) (p for interaction=0.0089).⁶

Five studies examining the association of diabetes with IS also reported on the association with HS,^{6, 90, 129, 131, 136} but most included only small numbers of HS events, producing less consistent results. However, three larger studies—the ERFC meta-analysis⁶ and studies based on the CALIBER programme⁹⁰ and on Taiwan National HI data¹³⁶—each reported significant, positive associations of diabetes with HS, more modest than with IS. HRs ranged from 1.16 (1.11-1.22) in men and 1.20 (1.15-1.26) in women in the Taiwanese study,¹³⁶ to 1.28 (1.02-1.62) in the CALIBER study⁹⁰ and 1.56 (95% CI 1.19-2.05)⁶ in the ERFC meta-analysis. When ICH was investigated separately, there was clearer evidence

of a significant positive association with diabetes^{6, 90} than when a composite HS outcome (ICH and SAH) was investigated.¹²⁹ Evidence of the association of diabetes with SAH is more limited, but the CALIBER study, with ~1200 SAH cases, found a reduced risk (0.48, 0.26-0.89) among people with diabetes.⁹⁰

Limitations of the evidence and implications for future study

The validity of diabetes diagnoses was reported on favourably in stroke-specific studies,^{127, 129, 136, 137} but comparisons between studies are again affected by the use of different diabetes definitions. Although most studies excluded individuals with prevalent stroke, or CVD more broadly, at baseline,^{6, 90, 129, 131, 132, 135, 137} this was not universal,^{130, 134, 136} with subsequent risk of reverse causality. Risk estimates may also have been distorted by bias resulting from substantial loss to follow-up in a small number of studies, including the CMCS (14%)¹³¹ and the HSE (10.4%) study.¹³⁵ The CMCS made no comparisons between those lost to follow-up and the remainder of the study population.¹³¹ In the HSE, study participants not followed-up were generally from lower socio-economic groups experiencing poorer health,¹³⁵ which could result in underestimation of effect sizes. As with studies examining IHD, there was minimal reporting of adjudication or validation of stroke outcomes.

Thus, although the association of diabetes with stroke risk has been investigated in several relatively large prospective studies, there are clear gaps in our understanding. Furthermore, there are no estimates from large, contemporary studies in Chinese populations. Although there are several older studies in China,^{131, 132} these are relatively small and may not accurately reflect the current situation. More recent studies have included restricted populations¹³⁰ or have been limited by use of routine data.¹³⁶ Given differences in the typical patterns of CVDs compared with Western populations, prospective studies with detailed phenotyping of CVD events in Chinese populations may provide unique insight, for example into the association of diabetes with ICH, which is seen more frequently in Chinese populations.⁵⁵ Again, and as previously noted, sufficiently large studies would enable

detailed investigation of effect modification in the association of diabetes with stroke subtypes, which has, to-date, been limited.

2.4 Blood glucose levels and cardiovascular disease risk

2.4.1 Cross-sectional and case-control evidence

Ischaemic heart disease

Three studies reported on the association of blood glucose levels with prevalent IHD in the general population,¹³⁹⁻¹⁴¹ but none in a Chinese population (Tables 2.11 and 2.12). Overall, the quality of evidence is poor; studies were small and in relatively young populations, and were susceptible to reverse causality and regression dilution bias. As such, few conclusions can be reached about the nature of any association.

The Cleveland Clinic Foundation study recruited ~2500 men and women (mean age 57 years) without diabetes—defined as a prior diagnosis or FBG >6.9 mmol/L—from a US preventive cardiology clinic.¹³⁹ There was a continuous positive association between FBG levels and prevalent IHD, with univariate ORs of 1.00 (reference), 1.17 (95% CI 0.92-1.50), 1.53 (1.19-1.96), 1.70 (1.32-2.19) and 2.50 (1.92-3.25) across quintiles of FBG. After adjustment for age, sex, smoking, SBP and diastolic blood pressure (DBP), LDL-cholesterol, high-density lipoprotein (HDL)-cholesterol and BMI, the odds of prevalent IHD were reported to be significantly elevated in the third, fourth and fifth quintiles only ($p<0.01$, $p<0.001$, $p<0.001$, respectively), although precise ORs were not presented. About half of the participants in this study had IHD, likely reflecting recruitment from a preventive cardiology clinic. This may limit the generalisability of the findings or may have biased risk estimates. The validity of the non-fatal outcomes investigated was not described, but a reasonably clear IHD definition was used—based on self-report or clinical investigation—which was likely reasonably robust against misclassification.

In contrast, there was no evidence of an association between OGTT post-load plasma glucose levels and prevalent IHD in the similarly small Finnish Social Insurance Institution Coronary Heart Disease Study.¹⁴⁰ IHD was defined based on symptoms and ECG findings, and three outcomes—definite MI and two different definitions of IHD—were investigated, affecting up to approximately 650 men. HRs for each outcome were 1.00 or 1.01 per 0.06 mmol/L higher blood glucose in multi-variable analyses.

Finally, a small, single-centre, hospital-based case-control study in India examined the association of plasma glucose in individuals without diabetes with first MI. The study included ~200 cases, aged 30-69 years, and a similar number of age- and sex-matched hospital controls with conditions unrelated to CVD and no prior history of heart disease or ischaemic ECG changes.¹⁴¹ Plasma glucose measurements were undertaken at nine or 10 days post-admission in cases. After adjusting for confounders, there was no association between FPG and the odds of MI. A positive association was seen, however, with post-load plasma glucose; each standard deviation (SD) (2.3 mmol/L) increase was associated with a two-fold increase in the odds of MI (OR 2.01, 95% CI 1.56-2.59). These analyses were limited by lack of adjustment for socioeconomic status and alcohol.

Limitations of the evidence and implications for future study

This review demonstrates the current lack of evidence on the association of blood glucose levels with prevalent IHD. Furthermore, the evidence is restricted by limitations inherent in the design of the studies, and large-scale prospective analyses are required, including in Chinese populations.

Stroke

No studies were identified describing the relationship of blood glucose levels with prevalent stroke in the general population.

2.4.2 Prospective evidence

Ischaemic heart disease

Six papers were identified on the relationship between blood glucose levels and incident IHD, including five meta-analyses^{6, 103-105, 142} and one large prospective cohort study¹⁰² (Tables 2.13 and 2.14). No Chinese studies were identified. Among those studies presenting IHD risk in individuals with blood glucose levels in the pre-diabetic and diabetic ranges, all found significantly elevated risks when compared with various reference groups within the “normoglycaemic” range.^{6, 102, 103} Below this level, the findings were less consistent; several larger studies found positive associations between blood glucose and IHD risk, although different shapes and strengths of associations were described.^{6, 102, 103} This could reflect investigation of different glycaemic measures or study populations, use of varying inclusion and exclusion criteria and IHD definitions, and failure of multiple studies to adjust for regression dilution bias.

All six studies investigated the association of fasting glucose with IHD.^{6, 102-105, 142} In the Korean Cancer Prevention Study (KCPS) of almost 1.2 million men and women, individuals with known CVD or diabetes at baseline were excluded, and data on fatal and non-fatal events—58 148 IHD and 16 659 MI—were obtained from health insurance and national mortality data.¹⁰² After adjusting for confounders, there was a J-shaped association between baseline FPG and both IHD and MI risk, with the lowest risk at a FPG of ~5.0 mmol/L. Exclusion of events during the first five years of follow-up did not markedly alter the association. Compared with the reference group (4.8-5.5 mmol/L), a FPG level of 5.6-6.0 mmol/L was associated with an adjusted IHD HR of 1.04 (95% CI 1.01-1.08) and 1.00 (0.95-1.05) in men and women, respectively, which increased to 1.14 (1.10-1.19) and 1.15 (1.08-1.23) at 6.1-6.9 mmol/L, and increased further above the diabetic threshold (7.0-7.7 mmol/L: men 1.27, 1.18-1.37 and women 1.36, 1.23-1.51; ≥ 7.8 mmol/L: men 1.79, 1.71-1.87 and women 2.21, 1.98-2.26). The association with MI was moderately stronger. Strengths of association for a given increment in FPG were not reported, nor were analyses

adjusted for regression dilution bias. Whilst potential inaccuracies in outcome data, derived from routine data, are acknowledged by the study's authors, a validation study comparing MI diagnoses with medical record review found 73% accuracy, although the validity of broader IHD diagnoses, arguably more susceptible to misclassification, was not described.

Two individual participant data meta-analyses examined the shape and strength of association of FBG levels in individuals without known diabetes with IHD risk.^{6, 103} The previously described ERFC included 51 cohort studies with a total of ~260 000 individuals (~13 000 IHD events) without a prior history of diabetes, stroke, MI or angina.⁶ There was a J-shaped association, after extensive adjustment for potential confounders, with no marked association at FBG levels between 3.9 and 5.6 mmol/L, but above this the IHD risk increased progressively with increasing FBG, with a 12% (95% CI 8-15%) increased IHD risk per 1 mmol/L higher baseline FBG. The main analyses were not adjusted for regression dilution bias, but a regression dilution ratio of 0.69 was reported. After adjustment for this, each 1 mmol/L higher FBG would be associated with a HR of 1.18 for IHD. The Asia Pacific Cohort Studies Collaboration (APCSC) included ~240 000 individuals from 13 prospective studies in the Asia Pacific region and recorded 816 IHD events during an average 5-year follow-up.¹⁰³ After adjustment for age, sex and cohort, a positive log-linear association was demonstrated between usual FBG levels and IHD risk, extending down to at least 4.9 mmol/L, with a 23% (19-30%) lower IHD risk per 1 mmol/L lower usual FBG. This association was modestly stronger than that of usual FBG in the ERFC study, possibly reflecting incomplete adjustment for confounding and inclusion of individuals with known diabetes in the APCSC. Additional adjustment for SBP, total cholesterol, smoking and BMI or exclusion of individuals with known diabetes were reported to not substantially change the results, although these data were not presented. Furthermore, inclusion of individuals with prior IHD and the relatively short follow-up period may have inflated risk estimates due to reverse causality. The APCSC also examined the association of non-fasting blood

glucose with IHD based on ~30 000 participants (240 IHD events). The association was weaker than that of FBG, although risk estimates were not presented.

A pooled analysis of tabular data (ie, summary data from both published and unpublished analyses¹⁴³) from the ERFC, the APCSC and the Diabetes Epidemiology: Collaborative analysis of Diagnostic criteria in Europe (DECODE) Study (comprising data from 13 European cohorts) included approximately 372 000 individuals from 116 cohorts (with some overlap in the cohorts included across different meta-analyses).¹⁴² The shape of the association between FPG and IHD risk was not described, but there was a significant association between usual FPG levels (ie, adjusted for regression dilution bias) and fatal or non-fatal IHD, which appeared to be stronger at younger than older ages (HR 1.21, 95% CI 1.02-1.43 per 1 mmol/L higher usual FPG at 35-44 years versus 1.14, 1.07-1.22 at ≥85 years). An overall risk estimate across all age groups was not presented.

Einarson et al. conducted a published data meta-analysis of 33 prospective studies (24 of which were included in other meta-analyses in this review^{6, 103, 105}) investigating the association of blood glucose levels with CVD in individuals without known diabetes.¹⁰⁴ Five included studies, comprising ~14 500 participants and ~500 events, reported on the association of FBG levels with fatal IHD. Based on the (untested) assumption of a continuous relationship, it was estimated that each 1 mmol/L increase in FBG was associated with a HR of 1.16 (95% CI 1.04-1.30) for fatal IHD. The extent of adjustment for confounding and for regression dilution bias is unclear. Although the number of IHD events was small, the risk estimates were reasonably consistent with those from the ERFC⁶ and APCSC.¹⁰³ Another published data meta-analysis by Sarwar et al included 22 studies from Western Europe, North America and Australia,¹⁰⁵ all but three of which (including ~2800 IHD events) were included in other meta-analyses.^{6, 103, 104} This study, which included ~250 000 individuals without diabetes and almost 11 000 IHD cases, reported somewhat less extreme risk estimates; overall, after adjusting for age, sex, smoking, SBP, lipids and BMI, but not regression dilution bias, each 1 mmol/L higher FBG (from a random effects model)

was associated with a pooled estimated HR of 1.06 (1.00-1.12) for IHD.¹⁰⁵ The weaker association may reflect exclusion of individuals with a FBG level in the diabetic range. Due to reliance on published data only, this study¹⁰⁵ and that by Einarson et al¹⁰⁴ are susceptible to publication bias, most likely over-estimating blood glucose-associated risks of IHD. Although formal statistical testing for publication bias in Sarwar et al's study found no significant evidence of this,¹⁰⁵ such tests have limited power.

The two published data meta-analyses investigated the association of other blood glucose measures with IHD risk.^{104, 105} Associations of post-load blood glucose levels with fatal IHD in the Einarson meta-analysis were modestly weaker than that of FBG, with HRs of 1.10 (95% CI 1.06-1.15) per 1 mmol/L higher 1-hour post-load blood glucose, based on ~1700 IHD events in ~39 000 participants, and 1.11 (1.03-1.20) per 1 mmol/L higher 2-hour post-load blood glucose levels in ~33 000 participants and ~4500 IHD deaths.¹⁰⁴ The narrower confidence limits for these estimates, compared with those for FBG, reflect larger sample sizes. In contrast, there was no association between random blood glucose (RBG) and IHD risk (1.02, 0.98-1.07), although these analyses included only ~9000 participants and 314 IHD events.¹⁰⁴ As described, the extent of adjustment for regression dilution bias and confounding is unclear. Sarwar et al. combined results from 15 studies (six of which were also included in Einarson et al.'s study¹⁰⁴), including 12 652 IHD events, investigating any form of post-load blood glucose.¹⁰⁵ After full adjustment, each 1 mmol/L higher post-load blood glucose was associated with an IHD HR of 1.05 (1.03-1.07).

Limitations of the evidence and implications for future study

The positive associations of blood glucose with IHD risk could reflect increased future risk of diabetes, and associated increased IHD risk, in individuals with higher blood glucose levels, but included studies did not address this. Since most included studies were meta-analyses, variation in blood glucose measurement methods, including assays used, between included studies could reduce the ability to identify true associations. Finally, there was investigation of effect modification by age^{103, 142} and sex^{102, 103} only.

This review demonstrates continuing controversy regarding the nature of the association of blood glucose levels with IHD risk. Large, well-powered studies are needed to reliably characterise the association, including the relevance of future diabetes risk, and more extensively investigate effect modification. In addition, despite the relevance and practical nature of RPG as a glycaemic measure, it has been minimally investigated. Uncertainties in the relationship of blood glucose with IHD thus remain in more widely-studied Western populations, but even more so in Chinese populations.

Stroke

Seven studies investigating the association of blood glucose levels with incident stroke risk were identified, which included two individual cohort studies^{102, 144} and five meta-analyses^{6, 83, 103, 104, 142} (Tables 2.15 and 2.16). None of the identified studies reported on the association in a Chinese population. Although most studies reported positive associations between blood glucose and TS and IS risk, the shapes and strengths of the associations differed. Despite including only the largest studies, the inconsistency of the findings may reflect inadequate power, due to the relatively weak association of blood glucose with stroke risk, combined investigation of different stroke subtypes in many studies, differences in glycaemic measures investigated and variable adjustment for regression dilution bias.

Most studies investigating the relationship of blood glucose levels with stroke risk investigated TS,^{83, 102-104, 144} potentially masking associations with stroke subtypes. The Aerobics Center Longitudinal Study included almost 44 000 men aged 20-87 years recruited from a Dallas preventive medicine clinic.¹⁴⁴ Individuals with prevalent MI, stroke or diabetes, or a FPG level <4.4 mmol/L at baseline were excluded, and 612 TS events were recorded during more than 700 000 person years of follow-up. Non-fatal events were ascertained through self-report, with a 70% response rate (baseline characteristics were reported to be similar in responders and non-responders). There was no association of FPG levels below the IFG range with age-adjusted TS incidence. There was, however, a suggestion of a positive log-linear association above this level (p for trend=0.01), with an

adjusted 6% (0-12%) increased risk of TS per 0.56 mmol/L higher FPG. The authors reported reasonable validity of stroke diagnoses, although CT/MR scanning was likely limited in this study, established in 1971.

Einarson et al's published data meta-analysis also reported on the association of blood glucose levels with TS.¹⁰⁴ Based on almost 16 500 individuals and 257 TS deaths, and assuming a continuous relationship, they found an 11% (95% CI 6-17%) increased risk of TS mortality per 1 mmol/L higher 1-hour post-load blood glucose. Although similar strengths of association with TS mortality were found for 2-hour post-load blood glucose and RBG, these were not significant. Based on ~29 000 participants and 2387 TS events, each 1 mmol/L higher 2-hour post-load blood glucose was associated with a HR of 1.10 (0.94-1.29), and each 1 mmol/L higher RBG was associated with a HR of 1.11 (0.95-1.31) based on 544 TS deaths in 3140 participants. The limitations of this study have been described, but, in particular, the relatively small number of stroke events, and investigation of TS outcomes, may have precluded detection of associations.

The DECODE Study included 21 906 adults aged 25-90 years, and investigated the relationships of both fasting and 2-hour post-load plasma glucose with TS mortality.⁸³ After excluding ~1000 individuals with known diabetes, just 255 stroke deaths were recorded. Compared with participants with a plasma glucose level below the pre-diabetic range, there was an elevated risk of fatal TS amongst those with screen-detected diabetes (HR 2.31, 95% CI 1.46-3.67 based on 2-hour post-load plasma glucose and 1.83, 1.16-2.89 based on FPG). A non-significantly elevated risk was found among those with pre-diabetes (2-hour post-load plasma glucose: 1.30, 0.92-1.84; FPG: 1.19, 0.87-1.64), probably reflecting inadequate statistical power. Assuming a continuous association (which was not formally tested), each 1 SD higher FPG (1.3 mmol/L) was associated with a 16% (0-33%) increased TS mortality risk, and each 1 SD higher 2-hour post-load plasma glucose (2.7 mmol/L) was associated with a 25% (12-40%) increased TS mortality risk. Again, these analyses did not adjust for regression dilution bias. A complication in assimilating evidence from different

studies is the use of different glucose measures. The DECODE Study, however, investigated the relative value of different measures for predicting TS mortality. When 2-hour post-load plasma glucose was added to a model containing FPG, log likelihood ratio testing showed an improved ability to predict TS mortality in men ($\chi^2=10.12$, $p=0.001$), but not in women ($\chi^2=0.01$, $p=0.94$). However, adding FPG to a model containing 2-hour post-load plasma glucose significantly improved the model's ability to predict TS mortality in women ($\chi^2=4.08$, $p=0.04$), but not in men ($\chi^2=3.29$, $p=0.07$). The inconsistent findings between sexes are perhaps due to the small sample size.

The APCSC also included 1661 fatal or non-fatal TS events, and, in keeping with the findings for IHD, there was a positive log-linear association between TS and usual FBG that continued down to at least 4.9 mmol/L, with a 21% (95% CI 18-24%) greater TS risk per 1 mmol/L higher usual FBG.¹⁰³ It was reported that more comprehensive adjustment for potential confounders, and exclusion of individuals with known diabetes at baseline did not substantially alter the association, although these data were not presented. The strength of association appeared to be similar between sexes, age groups and Asian and Australasian cohorts, and for fatal and non-fatal events. Again, individuals with prior stroke were not excluded from these analyses. A few studies included in the APCSC had non-fasting blood glucose data, and analyses showed a less extreme association with TS, although, again, no specific risk estimates were reported. The previously described meta-analysis of tabular data from the APCSC, ERFC and DECODE Study also reported age-specific associations of FPG with TS risk, and found evidence of a stronger association at younger ages.¹⁴²

Just three studies investigated the association of blood glucose with IS risk.^{6, 102, 103} The large KCPS recorded over 45 000 hospitalisations or deaths due to IS amongst individuals without a prior history of diabetes, with 83% of diagnoses validated in a sub-study.¹⁰² After adjusting for most relevant confounders, there was a J-shaped association between baseline FPG and IS risk in both men and women, continuing down to approximately 5.0

mmol/L, below which there was a non-significantly elevated risk of IS. Above 5.0 mmol/L, the IS risk increased progressively with higher FPG levels, to a maximum of 2.20 (95% CI 2.10-2.31) and 2.29 (2.15-2.44) at FPG levels ≥ 7.8 mmol/L in men and women, respectively, compared with the reference group (4.8-5.5 mmol/L). Sensitivity analyses excluding the first five years of follow-up did not alter the results, but HRs per unit change in FPG were not presented. In the ERFC, an individual participant data meta-analysis with a much smaller number of IS events ($n=1754$ before exclusion of individuals with known diabetes), there was no significant association between baseline FBG levels and IS risk below the diabetes threshold after adjustment for age, sex, smoking, SBP and BMI.⁶ In the APCSC, with an unspecified number of IS events, there was a non-significant 17% (-1-32%) lower risk of IS per 1 mmol/L lower FBG.¹⁰³

Two of the largest studies reporting on the association of fasting glucose levels with stroke risk thus described different shapes of association;^{102, 103} in contrast to the J-shaped relationship in the KCPS,¹⁰² the APCSC analyses found no threshold in the association.¹⁰³ This could reflect inadequately detailed examination of stroke risk at the lowest FBG levels in the APCSC,¹⁰³ or possibly a spuriously elevated risk in this range in the KCPS,¹⁰² or true differences between study populations.

Only one study (KCPS) included a large number of HS events ($n=19\,407$).¹⁰² This showed a much flatter association between FPG and HS than was seen for ischaemic CVDs. There was a suggestion of a J-shaped association, but HS risk was clearly elevated only amongst individuals in the highest FPG categories when compared with the reference group of 4.8-5.5 mmol/L (7.0-7.7 mmol/L: HR 1.15, 95% CI 1.01-1.30 and 1.12, 0.93-1.33 in men and women, respectively; ≥ 7.8 mmol/L: 1.24, 1.13-1.35 and 1.14, 1.00-1.30 in men and women, respectively). HRs per unit increase in FPG were again not presented. The number of HS events included in the APCSC was not specified, but a 1 mmol/L increment in usual FBG in that study was estimated to be associated with a 20% (0-36%) increased risk of HS.¹⁰³

Both studies included ICH and SAH in their definitions of HS, which may mask true associations, and the relative proportions of these subtypes were not described.^{102, 103}

Limitations of the evidence and implications for future study

This review further highlights the need for large-scale prospective studies investigating the association of usual blood glucose levels, and more specifically of usual RBG levels, with stroke, both generally and in Chinese populations, in order to robustly characterise the association. Inadequate investigation of stroke subtypes and of the impact of future diabetes risk in studies to-date highlights the importance of accurate phenotyping of a range of incident disease events. Investigation in large studies would additionally enable investigation of the associations of blood glucose levels in relevant population subgroups.

2.4.3 Mendelian randomisation studies assessing the causal relevance of blood glucose to cardiovascular disease

Observational epidemiological studies, as described in this discussion of existing literature, have various limitations inherent in their design. Key among these are susceptibility to confounding, measurement error, biases and reverse causality. As a result of these limitations it is not possible to draw conclusions about the causal nature of associations between blood glucose and CVD risk. One approach which can overcome many of these limitations is Mendelian randomisation. In this study design, a genetic variant associated with an environmental exposure of interest, but not with other risk factors for the disease of interest and only with the disease through the exposure of interest, can be used as a proxy for the exposure.¹⁴⁵ Since an individual's genotype is determined at conception, and alleles are inherited independently of each other (unless in linkage disequilibrium [LD]) and unrelated to the factors which frequently confound observational epidemiological studies, using a genetic variant as a proxy for an exposure can overcome issues of residual confounding and reverse causality, and can provide an indicator of lifelong exposure levels.^{146, 147} As such, if the genetic variant is associated with the disease of interest, it

follows that it must be through the exposure of interest, inferring a causal relationship between the exposure and the disease.¹⁴⁵

Three Mendelian randomisation studies examined the association of blood glucose with CVD¹⁴⁸⁻¹⁵⁰ (Tables 2.17 and 2.18). All identified studies used a genetic risk score (GRS) — a single variable summarising multiple genetic variants¹⁵¹—rather than investigating multiple individual variants. Although this provides greater power to detect a causal association and reduces the risk of weak instrument bias, as described below,¹⁵¹ one included study still lacked adequate power to demonstrate an effect. Overall, however, the studies were consistent with a causal relationship between blood glucose and ischaemic CVD.^{148, 149} No studies were identified in Chinese populations.

Analyses based on data from the Atherosclerosis Risk in Communities (ARIC) Study included 7260 Caucasian men and women aged 45-64 years and without prior CVD or diabetes, based on self-report or blood glucose measurement.¹⁴⁹ An internally-weighted FPG GRS was developed based on five single nucleotide polymorphisms (SNPs) located in regions of the genome significantly associated with FPG in a GWAS within the same study. After adjusting for age, sex and study centre, each 1 SD increment in the GRS was associated with a 0.0851 mmol/L higher FPG ($p=1.07 \times 10^{-54}$) and a 0.0048mm increase in carotid intima media thickness (CIMT) ($p=0.009$), which persisted, but was attenuated (0.0032mm, $p=0.09$), after addition of FPG to the model, suggesting FPG lies on the pathway between the GRS and CIMT. In the same population, each 1 SD (0.5 mmol/L) higher measured FPG was associated with a 0.0103mm increase in CIMT ($p=8.01 \times 10^{-8}$). The association of the GRS with CIMT was thus stronger than would be predicted based on the observational data (0.0018mm per 0.0851 mmol/L higher FPG), possibly reflecting the lifelong effects of genetic variants on FPG levels. One of the underlying assumptions of Mendelian randomisation is that the genetic variants are related to the outcome through the exposure of interest only. However, the associations of the genetic variants included in the FPG GRS with other cardiovascular risk factors were not investigated. It is possible,

therefore, that the association of the GRS with CIMT reflects pleiotropic effects of the included SNPs, that is, that the SNPs are associated with CIMT through an association with one or more other atherosclerotic risk factors, or through a direct effect on atherosclerosis. As a result of such pleiotropic effects, the conclusion of a causal association between FPG and CIMT may be inappropriate. Use of a GRS should, however, have reduced the risk of unidentified pleiotropy, since all genetic variants included in the GRS would be unlikely to have the same pleiotropic effects.

Similar analyses examining the association of random plasma glucose (RPG) with MI and IHD risk were undertaken based on combined data from three Danish studies, including individuals with known or newly diagnosed diabetes.¹⁴⁸ Among 16 568 participants without prior IHD in the Copenhagen City Heart Study (CCHS), after adjustment for confounders and for regression dilution bias, each 1 mmol/L higher RPG was associated with an 18% (95%CI 15-22%) and 9% (7-11%) increased risk of IHD and MI, respectively. Five SNPs were selected, based on a previous GWAS of FPG, for inclusion in an unweighted allele score. Among ~65 000 individuals in the CCHS and the Copenhagen General Population Study, a higher allele score was found to be associated with higher RPG levels and, in data on ~80 000 participants in these two studies and the Copenhagen Ischemic Heart Disease Study, with higher risks of MI and IHD. Using a two-stage least-squares approach, each 1 mmol/L increase in genetically-determined RPG was associated with an OR of 1.69 (1.28-2.23) for MI and 1.25 (1.03-1.52) for IHD, suggesting a causal relationship. There was no evidence of pleiotropic effects of the included SNPs on other measured cardiovascular risk factors.

The final identified study used data from CARDIOGRAMplusC4D—a consortium of GWAS data from approximately 64 000 IHD cases and 131 000 controls—, from the Meta-Analyses of Glucose and Insulin-Related Traits Consortium (MAGIC)—a large GWAS of over 133 000 individuals—and from the ERFC study to examine the association of FBG with IHD.¹⁵⁰ This study investigated 30 FBG-associated SNPs (identified in MAGIC) and found a non-

significant association between these SNPs and IHD risk (OR 1.18, 95% CI 0.97-1.42 per 1 mmol/L higher FBG), which was modestly attenuated (1.16, 0.94-1.42) after adjusting for the associations of the SNPs with other cardiovascular risk factors (LDL-cholesterol, HDL-cholesterol, total cholesterol, triglycerides and BMI). Based on observational data from the ERFC, each 1 mmol/L higher FBG was estimated to be associated with an IHD HR of 1.02 (95% CI 1.02-1.03). Although z-testing showed no difference between genetic and observational effect estimates (p for difference 0.143), a power calculation suggested the study had only 5.6% power to detect a genetic association between FBG and IHD, such that limited conclusions can be drawn from the study. Through adjusting the genetic FBG-associated IHD risk for the associations of the included SNPs with potential confounders, possible pleiotropic effects of the variants through these mechanisms were addressed. However, some potentially important cardiovascular risk factors, including blood pressure, were not adjusted for, since relevant data were not available. All included data were again derived from Caucasian populations.

Limitations of the evidence and implications for future study

The genetic variants investigated in these studies explained only a small proportion of variation in blood glucose levels.^{148, 149} For example, in the ARIC study the GRS explained just 2.9% of variance in FPG, which could result in weak instrument bias and possible distortion of the genetic risk estimates towards the confounded observational effect estimate, and an inappropriate conclusion of a causal association.¹⁵² The F statistic measures the strength of association between genetic variant and exposure and can be used to test for weak instruments.¹⁵² Despite the limited influence of the genetic variants on RPG, analyses in the Danish study suggested the allele score was a good instrument (F score >10).¹⁴⁸ Such analyses were not, however, undertaken in the ARIC study.¹⁴⁹ Furthermore, selection of genetic variants in the ARIC study based on a GWAS in the same study may have resulted in the “winner’s curse” effect. In the context of GWAS this may result in underpowered replication studies, and in Mendelian randomisation it may result in

over-estimates of the effect of a genetic variant, since those variants with the strongest effects in the GWAS are those most likely to have an exaggerated effect size.¹⁵³ In this study, however, repeat analyses, including 16 SNPs selected from a large, separate GWAS, only slightly attenuated the association of the GRS with CIMT.¹⁴⁹

Although studies to-date are suggestive of a causal relationship between blood glucose and CVD, no studies have been undertaken in Chinese populations and none has investigated manifest cerebrovascular diseases.

2.5 Summary of findings and main research questions

This review of the available literature demonstrates considerable gaps in understanding of the associations of diabetes, and of blood glucose levels in individuals without diabetes, with IHD and stroke. Although the relationship of diabetes with IHD and TS have been fairly extensively investigated in Western populations, much less is known about these in Chinese populations, in whom previous studies have been generally small, conducted several decades ago (hence unable to reflect more recent secular trends), or restricted to population subsets. These limitations, and the considerably stronger diabetes-associated odds of IHD in Chinese populations than in a combination of over 50 other populations in the INTERHEART case-control study demonstrate the need for large, population-based, prospective studies in China, to enable a robust and reliable understanding of the association of diabetes with IHD and stroke, including stroke subtypes (which have been less well-characterised in Western studies).

Several large individual prospective studies, and meta-analyses of these, have investigated the associations of blood glucose levels in the general population with the risk of IHD and stroke, showing considerable variation in both the shape (where investigated) and strength of association. Studies to-date have rarely adjusted for the effects of regression dilution bias, and no large studies have explored the possibility that a positive association could reflect the development of diabetes during follow-up amongst participants with higher blood

glucose levels. The associations with stroke subtypes have, again, been poorly characterised. No sizeable studies have investigated the associations of RBG, a more practical blood glucose measure with potentially greater real-world relevance. Furthermore, no prospective studies examining the association of blood glucose levels and IHD or stroke (even TS) in Chinese populations were identified. There has also been limited investigation of effect modification of associations of both diabetes and blood glucose.

Mendelian randomisation studies play an important role in determining causal associations between exposures and diseases. Limited data from such studies are consistent with a causal association between higher blood glucose levels and CVD. However, all three studies identified were undertaken in Caucasian populations, and it is unclear whether their findings are generalisable to populations with different genetic architecture, including Chinese populations.

2.6 Aims and objectives of thesis

In light of the presented evidence, and, more importantly, the gaps in this evidence, the aim of this thesis is to better understand the distribution of blood glucose levels and diabetes in Chinese adults, their associations with other cardiovascular risk factors and with prevalent IHD and stroke/TIA, and their roles in determining IHD and stroke incidence, using both conventional and genetic techniques. The work is based on the prospective CKB study of 0.5 million adults recruited from 10 diverse areas of China during 2004-2008 and followed up until 1st January 2014 for cause-specific mortality and hospitalisations.

The specific objectives of this thesis are to:

1. Examine the prevalence of diabetes and the distribution of plasma glucose levels within the CKB population, overall and in certain population subgroups (eg, defined by age, study area and sex)
2. Understand the association of diabetes and of plasma glucose with other cardiovascular risk factors

3. Characterise the association of diabetes and of plasma glucose with prevalent IHD and stroke/TIA
4. Examine how the associations of diabetes and of plasma glucose with prevalent IHD and stroke/TIA are modified by other factors eg, age, sex
5. Characterise the association of diabetes and of usual plasma glucose, after adjustment for within-person variation, with the risk of incident IHD and stroke, including haemorrhagic and ischaemic stroke subtypes
6. Examine how associations of diabetes and of usual plasma glucose with incident IHD and stroke are modified by other factors eg, age, sex
7. Examine the association of blood glucose-related genetic variants with IHD and stroke, and use a Mendelian randomisation approach to investigate the likely causal nature of any association

Table 2.1. Literature review search terms

Search terms		
Stroke	Stroke stroke.mp cerebrovascular disease.mp cerebrovascular disease/ cerebrovascular disorder.mp Cerebrovascular Disorders cerebrovascular accident.mp cerebrovascular accident CVA.mp cerebral vascular disease.mp cerebr\$.mp cerebell\$.mp Brain brain.mp Infarction	infarct.mp Ischemia isch?emi\$.mp ischemia Thromboembolism Thrombosis thrombo\$.mp Embolism emboli\$.mp intracerebral.mp h?emorrhage.mp Hemorrhage Cerebral H?emorrhage.mp Cerebral Hemorrhage Bleed
Ischaemic heart disease	isch?emic heart disease.mp ischemic heart disease myocardial infarction.mp Myocardial Infarction myocardial isch?emia.mp Myocardial Ischemia heart infarction acute coronary syndrome.mp acute coronary syndrome	coronary heart disease.mp coronary disease.mp Coronary Disease acute coronary syndrome.mp Acute Coronary Syndrome/ coronary artery disease.mp coronary artery disease heart muscle ischemia
Cardiovascular disease	cardiovascular disease.mp Cardiovascular Diseases	
Diabetes	diabetes mellitus.mp. Diabetes Mellitus Diabetes Mellitus, Type 1 T1DM.mp Diabetes Mellitus, Type 2 T2DM.mp NIDDM.mp IDDM.mp	diabet\$.mp type 1 diabetes mellitus.mp type 2 diabetes mellitus.mp non insulin dependent diabetes mellitus insulin dependent diabetes mellitus/
Blood glucose	blood glucose.mp glucose blood level Blood Glucose blood sugar.mp serum glucose.mp plasma glucose.mp	glyc?emia.mp glucose tolerance.mp glucose tolerance dysglyc?emia.mp dysglycemia
Prospective studies	cohort stud\$.mp Cohort Studies case cohort stud\$.mp nested case control study.mp cohort anal\$.mp follow up stud\$.mp Longitudinal Studies	longitudinal.mp prospective.mp Prospective Studies cohort analysis longitudinal study prospective study
Cross-sectional and case-control studies	case control study.mp case control study Case-Control Studies Prevalence	prevalent.mp Cross-Sectional Studies cross-sectional.mp
Mendelian randomisation studies	Mendelian Randomization Analysis	Mendelian randomisation.mp

Searches undertaken in MEDLINE (1946 to October Week 1 2013) and Embase (1974 to 14th October 2013) and repeated weekly until 31/08/2015.

Table 2.2. Selection of articles for inclusion in the literature review

Literature search	Number of studies				
	Initial search		Identified in reference lists	Identified in weekly repeats of search	Included in literature review
	Identified	Remaining after exclusions			
Diabetes and cardiovascular disease					
Cross-sectional and case-control ^a	13 469	17	0	0	17
Prospective ^b	13 716	8	1	5	14
Blood glucose and cardiovascular disease					
Cross-sectional and case-control ^c	4147	3	0	0	3
Prospective ^d	3342	4	2	2	8
Mendelian randomisation studies of blood glucose and cardiovascular disease	14	2	0	1	3

^aExclusion criteria: <1000 ischaemic heart disease (IHD) events; <100 stroke events; participants selected on the basis of pre-existing disease (with the exception of diabetes mellitus); all participants with a known diagnosis of diabetes mellitus; investigation of composite cardiovascular disease (CVD) only; individual studies included in meta-analyses selected for inclusion in the literature review.

^bExclusion criteria: <1500 IHD events; <500 stroke events; participants selected on the basis of pre-existing disease (with the exception of diabetes); all participants with a known diagnosis of diabetes mellitus; investigation of composite CVD only; individual studies included in meta-analyses selected for inclusion in the literature review.

^cExclusion criteria: <100 IHD events; <100 stroke events; participants selected on the basis of pre-existing disease (with the exception of diabetes mellitus); all participants with a known diagnosis of diabetes mellitus; investigation of composite CVD only; individual studies included in meta-analyses selected for inclusion in the literature review.

^dExclusion criteria: <500 IHD events; <200 stroke events; participants selected on the basis of pre-existing disease (with the exception of diabetes mellitus); all participants with a known diagnosis of diabetes mellitus; investigation of composite CVD only; individual studies included in meta-analyses selected for inclusion in the literature review.

CVD event number criteria were not applied to studies in Chinese populations.

Where more than one article reported on the same study, the article with the longest follow-up period was included.

Table 2.3. Characteristics of studies examining the relationship between diabetes and prevalent ischaemic heart disease

Study (year established)	No. of participants (% female)	Geographical location	Mean age (years)	Definition of diabetes	No. of events
Cross-sectional					
Care Data Warehouse in Östergötland (1999) ¹¹¹	141 400 (50)	Sweden	N/S	Symptoms BG measurement	11 311
China National Diabetes and Metabolic Disorders Study (2007) ¹¹³	46 239 (60)	China	N/S	Self-report BG measurement	313
Rural Beijing study (2008) ¹¹⁰	58 308 (65)	China	56.3	Self-report on medication	N/S
Jordan Hyperlipidemia and Related Targets Study (N/S) ¹¹²	5000 (34)	Jordan	M: 50.1 F: 54.1	Self-report Medication use	1534
Case-control					
Stockholm Heart Epidemiology Program (1992) ¹¹⁶	5452 (33)	Sweden	N/S	Self-report Medication use BG measurement	2246
INTERHEART (1999) ¹¹⁵	6086 (N/S)	China	N/S	Self-report	3030
INTERHEART (1999) ¹¹⁷	27 098 (75)	Varied	58	Self-report	12 461
Japanese Acute Coronary Syndrome Study (2002) ¹¹⁴	4204 (30)	Japan	Cases: 67.7 Controls: N/S	BG measurement	1925
Hamad Medical Corporation (2006) ¹¹⁸	894 (22)	Qatar	N/S	Self-report	512
IPD meta-analysis					
Italian hospital-based studies (N/A) ¹¹⁹	4054 (40)	Italy	N/S	Self-report	1737

BG, blood glucose; F, female; IPD, individual participant data; M, male; N/A, not applicable; N/S, not stated.

Table 2.4. Key findings of studies examining the relationship between diabetes and prevalent ischaemic heart disease

Study	OR (95%CI)		Adjustment/ matching	Variables investigated for effect modification
	IHD	MI		
Cross-sectional				
Care Data Warehouse in Östergötland ¹¹¹	M: 2.65 (2.52-2.79) F: 3.97 (3.69-4.27)		–	Sex
China National Diabetes and Metabolic Disorders Study ¹¹³	2.44 (1.31-4.53)		++++	None
Rural Beijing study ¹¹⁰	2.51 (2.29-2.75)		+	None
Jordan Hyperlipidemia and Related Targets Study ¹¹²	M: 2.93 (2.50-3.44) F: 5.59 (4.33-7.23)		+++	Sex
Case-control				
Stockholm Heart Epidemiology Program ¹¹⁶		M: 2.45 (1.86-3.22) F: 4.36 (2.72-7.01)	++	Sex
INTERHEART ¹¹⁵		5.10 (4.03-6.44)	+++++	Location
INTERHEART ¹¹⁷		M: 2.67 (2.36-3.02) ^a F: 4.26 (3.51-5.18) ^a	++++	Sex, age, region
Japanese Acute Coronary Syndrome Study ¹¹⁴		M: 2.90 (2.00-4.04) F: 6.12 (3.78-12.02)	+	Sex
Hamad Medical Corporation ¹¹⁸		4.01 (2.51-6.42)	++	None
IPD meta-analysis				
Italian hospital-based studies ¹¹⁹		M: 1.4 (1.0-1.9) F: 4.4 (2.9-6.6)	+++++	Sex, diabetes duration, age at diabetes diagnosis

^a99%CI.

CI, confidence interval; F, female; IHD, ischaemic heart disease; IPD, individual participant data; M, male; MI, myocardial infarction; OR, odds ratio.

Adjustment/matching: - not adjusted; + age and sex; ++ plus smoking; +++ plus other standard variable (blood pressure, physical activity); ++++ plus lipids; +++++ plus socioeconomic status.

Table 2.5. Characteristics of studies examining the relationship between diabetes and prevalent stroke

Study (year established)	No. of participants (% female)	Geographical location	Mean age (years)	Definition of diabetes	No. of events
Cross-sectional					
10/66 Population-Based Study (2003) ¹²⁰	2937 (65)	Cuba	74	Self-report	TS: 229
China National Diabetes and Metabolic Disorders Study (2007) ¹¹³	46 239 (60)	China	N/S	Self-report BG measurement	TS: 472
Rural Beijing study (2008) ¹¹⁰	58 308 (65)	China	56.3	Self-report on medication	N/S
Case-control					
Manitoba Health Services Commission (1970) ¹²⁶	1212 (43)	Canada	67	Unclear	IS: 410 HS: 137 US: 59
MRC GP Research Framework (1981) ¹²¹	732 (100)	UK	N/S	Medical record review	TS: 244
Stroke Registry of Dijon (1985) ¹²³	260 (45)	France	64	Self-report Medical record review	HS: 130
Chongqing case control study (2004) ¹²⁴	2068 (N/S)	China	N/S	Self-report	TS: 1034
Hamad Medical Corporation (2006) ¹¹⁸	644 (22)	Qatar	N/S	Self-report	TS: 262
INTERSTROKE (2007) ¹²²	6000 (37)	Varied	Cases: 61.1 Controls: N/S	Self-report	TS: 3000
Indian case control study (N/S) ¹²⁵	332 (68)	India	N/S	BG measurement	HS: 166

BG, blood glucose; HS, haemorrhagic stroke; IS, ischaemic stroke; N/S, not stated; TS, total stroke; US, unknown stroke.

Table 2.6. Key findings of studies examining the relationship between diabetes and prevalent stroke

Study	OR (95%CI)			Adjustment/ matching	Variables investigated for effect modification
	TS	IS	HS		
Cross-sectional					
10/66 Population-Based Study ¹²⁰	2.6 (2.0-3.5)			+++	None
China National Diabetes and Metabolic Disorders Study ¹¹³	0.97 (0.59-1.61)			++++	None
Rural Beijing study ¹¹⁰	2.24 (1.98-2.52)	A		+	None
Case-control					
Manitoba Health Services Commission ¹²⁶		2.5, p<0.0002	1.3, N/S	+	None
MRC GP Research Framework ¹²¹	10.5 (3.97-27.9)			+	None
Stroke Registry of Dijon ¹²³			1.3, N/S	+ ^a	None
Chongqing case control study ¹²⁴	M: 1.81 (1.14-2.88) F: 3.65 (2.34-5.69)			+	Sex
Hamad Medical Corporation ¹¹⁸	5.12 (3.11-8.60)			++	None
INTERSTROKE ¹²²	1.36 (1.10-1.68) ^b	1.60 (1.29-1.99) ^b		+++	None
Indian case control study ¹²⁵			2.39 (0.75-7.59)	++++	None

^aAdjustment unclear; ^b99%CI.

CI, confidence interval; F, female; HS, haemorrhagic stroke; IS, ischaemic stroke; M, male; OR, odds ratio; TS, total stroke. Adjustment/matching as Table 2.4.

Table 2.7. Characteristics of studies examining the relationship between diabetes and incident ischaemic heart disease

Study (year established)	No. of participants (% female)	Geographical location	Mean age at baseline (years)	Definition of diabetes	Mean follow-up (years)	No. of events
Cohort						
Cancer Prevention Study II (1982) ¹²⁷	1 053 831 (56)	USA	N/S	Self-report	26	104 512
USA-People's Republic of China Collaborative Study of Cardiovascular and Cardiopulmonary Epidemiology (1983) ¹³²	9903 (51)	China	46	Self-report on medication	15.1	105
Chinese Multi-provincial Cohort Study (1992) ¹³¹	30 378 (47)	China	46.9	Self-report BG measurement	N/S	N/S
General Practice Research Database (1992) ¹²⁸	235 640 (49)	UK	66.7	Routine data	9	9058
Million Women Study (1996) ¹²⁹	1 242 338 (100)	UK	56	Self-report	6.1	21 928
CALIBER (1998) ⁹⁰	1 921 260 (51)	UK	N/S	Routine data	5.5	15 897
Hong Kong elderly cohort study 2 (1998) ¹³⁰	66 820 (66)	China	N/S	Self-report	10.9	2198
PD meta-analysis						
Varied (N/A) ^{a 133}	858 507 (42)	Varied	N/S	Self-report BG measurement	N/S	28 203
IPD meta-analysis						
Emerging Risk Factors Collaboration (N/A) ⁶	530 083 (42)	Varied	52	Self-report Medication use BG measurement	13.7	26 505

^aCombination of published data and individual participant data meta-analysis.

BG, blood glucose; IPD, individual participant data; N/A, not applicable; N/S, not stated; PD, published data.

Table 2.8. Key findings of studies examining the relationship between diabetes and incident ischaemic heart disease

Study	HR (95%CI)		Adjustment	Variables investigated for effect modification
	IHD	MI		
Cohort				
Cancer Prevention Study II ¹²⁷	M: 2.11 (2.05-2.16) F: 2.46 (2.39-2.53)		+++	Sex
USA-People's Republic of China Collaborative Study of Cardiovascular and Cardiopulmonary Epidemiology ¹³²	M: 1.51, p=0.4318 F: 2.99, p=0.0145		-	Sex
Chinese Multi-provincial Cohort Study ¹³¹	1.81 (1.20-2.71)		++	None
General Practice Research Database ¹²⁸		M: 2.13 (2.01-2.26) F: 2.95 (2.75-3.17)	++++	Age, sex
Million Women Study ¹²⁹	2.44 (2.32-2.57)		++++	None
CALIBER ⁹⁰		1.54 (1.42-1.67)	+++++	Age, sex
Hong Kong elderly cohort study 2 ¹³⁰	M: 2.20 (1.89-2.57) F: 2.25 (1.98-2.55)		+++	Age, sex
PD meta-analysis				
Varied ^{a133}	M: 2.16 (1.82-2.56) F: 2.82 (2.35-3.38)		Varied	Sex
IPD meta-analysis				
Emerging Risk Factors Collaboration ⁶	M:1.89 (1.73-2.06) F: 2.59 (2.29-2.93)	1.82 (1.64-2.03)	++++	Age, sex, smoking, BMI, SBP

^aCombination of published data and individual participant data meta-analysis.

BMI, body mass index; CI, confidence interval; F, female; HR, hazard ratio; IHD, ischaemic heart disease; IPD, individual participant data; M, male; MI, myocardial infarction; PD, published data; SBP, systolic blood pressure. Adjustment as Table 2.4.

Table 2.9. Characteristics of studies examining the relationship between diabetes and incident stroke

Study (year established)	No. of participants (% female)	Geographical location	Mean age at baseline (years)	Definition of diabetes	Mean follow-up (years)	No. of events
Cohort						
Cancer Prevention Study II (1982) ¹²⁷	1 053 831 (56)	USA	N/S	Self-report	26	TS: 34 959
USA-People's Republic of China Collaborative Study of Cardiovascular and Cardiopulmonary Epidemiology (1983) ¹³²	9903 (51)	China	46	Self-report on medication	15.1	IS: 266
Hong Kong elderly cohort study 1 (1986) ¹³⁴	304 (60)	China	N/S	BG measurement	2.5	TS: 7
Chinese Multi-provincial Cohort Study (1992) ¹³¹	30 378 (46.6)	China	46.9	Self-report BG measurement	N/S	N/S
Health Survey for England (1994) ¹³⁵	82 380 (55)	UK	55.4	Self-report	8	TS: 806
Million Women Study (1996) ¹²⁹	1 242 338 (100)	UK	56	Self-report	6.1	TS: 7087 IS: 2572 HS: 2506 US: 2009
Taiwan National Health Insurance (1997) ¹³⁶	1 000 064 (53)	Taiwan	59.7	Routine data	6	IS: 107 475 HS: 14 488
CALIBER (1998) ⁹⁰	1 921 260 (51)	UK	N/S	Routine data	5.5	IS: 5959 HS: 2349
Hong Kong elderly cohort study 2 (1998) ¹³⁰	66 820 (66)	China	N/S	Self-report	10.9	TS: 1945
Newfoundland and Labrador cohort study (1998) ¹³⁷	73 783 (59)	Canada	60.1	Routine data	10	TS: 574

Table 2.9 continued

Study (year established)	No. of participants (% female)	Geographical location	Mean age at baseline (years)	Definition of diabetes	Mean follow-up (years)	No. of events
PD meta-analysis						
Varied (N/A) ^{a 138}	775 385 (41)	Varied	N/S	Self-report BG measurement	N/S	TS: 12 539
IPD meta-analysis						
Emerging Risk Factors Collaboration (N/A) ⁶	530 083 (42)	Varied	52	Self-report Medication use BG measurement	13.7	IS: 3799 HS: 1183 US: 4973

^aCombination of published data and individual participant data meta-analysis.

BG, blood glucose; HS, haemorrhagic stroke; IPD, individual participant data; IS, ischaemic stroke; N/A, not applicable; N/S, not stated; PD, published data; TS, total stroke; US, unknown stroke.

Table 2.10. Key findings of studies examining the relationship between diabetes and incident stroke

Study	HR (95%CI)				Adjustment	Variables investigated for effect modification
	TS	IS	HS	US		
Cohort						
Cancer Prevention Study II ¹²⁷	M: 1.87 (1.77-1.97) F: 1.70 (1.61-1.80)				+++	Sex, follow-up period
USA-People’s Republic of China Collaborative Study of Cardiovascular and Cardiopulmonary Epidemiology ¹³²		M: 1.08, p=0.8351 F: 2.89, p=0.0003			-	Sex
Hong Kong elderly cohort study 1 ¹³⁴	2.0 (0.4-26.3)				-	None
Chinese Multi-provincial Cohort Study ¹³¹		2.06 (1.54-2.76)	0.84 (0.44-1.60)		++	None
Health Survey for England ¹³⁵	1.87 (1.28-2.73)				+	None
Million Women Study ¹²⁹	2.08 (1.88-2.31)	2.55 (2.19-2.96)	0.87 (0.65-1.16)	2.60 (2.19-3.08)	++++	None
Taiwan National Health Insurance ¹³⁶		M: 1.89 (1.85-1.92) F: 2.32 (2.27-2.36)	M: 1.16 (1.11-1.22) F: 1.20 (1.15-1.26)		+	Age, sex
CALIBER ⁹⁰		1.72 (1.52-1.95)	1.28 (1.02-1.62)		+++++	Age, sex
Hong Kong elderly cohort study 2 ¹³⁰	M: 1.53 (1.26-1.85) F: 1.60 (1.38-1.85)				+++	Age, sex
Newfoundland and Labrador cohort study ¹³⁷	1.62 (1.35-1.94)				+	Early and late diabetes diagnosis

Table 2.10 continued

Study	HR (95%CI)				Adjustment	Variables investigated for effect modification
	TS	IS	HS	US		
PD meta-analysis						
Varied ^{a138}	M: 1.83 (1.60-2.08) F: 2.28 (1.93-2.69)				Varied	Sex
IPD meta-analysis						
Emerging Risk Factors Collaboration ⁶		2.27 (1.95-2.65)	1.56 (1.19-2.05)	1.84 (1.59-2.13)	++++	Age, sex, smoking, BMI, SBP

^aCombination of published data and individual participant data meta-analysis.

BMI, body mass index; CI, confidence interval; F, female; HR, hazard ratio; HS, haemorrhagic stroke; IPD, individual participant data; IS, ischaemic stroke; M, male; SBP, systolic blood pressure; TS, total stroke; US, unknown stroke. Adjustment as Table 2.4.

Table 2.11. Characteristics of studies examining the relationship between blood glucose levels and prevalent ischaemic heart disease

Study (year established)	No. of participants (% female)	Geographical location	Mean age (years)	Glycaemic measure	No. of events
Cross-sectional					
Cleveland Clinic Foundation (1996) ¹³⁹	2440 (34)	USA	57.0	FBG	1274
Social Insurance Institution's Coronary Heart Disease Study (N/S) ¹⁴⁰	3351 (0)	Finland	48.5	1h OGTT	Definite MI- 111 Whitehall criteria- 656 Pooling Project Major Abnormalities- 334
Case control					
Bangalore case control study (N/S) ¹⁴¹	600 (93)	India	Cases: 47.3 Controls: 47.0	FPG 2h OGTT	300

1h OGTT, 1 hour post-load glucose in oral glucose tolerance test; 2h OGTT, 2 hour post-load glucose in oral glucose tolerance test; FBG, fasting blood glucose; FPG, fasting plasma glucose; MI, myocardial infarction; N/S, not stated.

Table 2.12. Key findings of studies examining the relationship between blood glucose levels and prevalent ischaemic heart disease

Study	OR (95%CI)		Adjustment/ matching	Correction for regression dilution bias	Variables investigated for effect modification
	IHD	MI			
Cross-sectional					
Cleveland Clinic Foundation ¹³⁹	Q1 Reference Q2 1.17 (0.92-1.50) Q3 1.53 (1.19-1.96) Q4 1.70 (1.32-2.19) Q5 2.50 (1.92-3.25)		-	No	None
Social Insurance Institution's Coronary Heart Disease Study ¹⁴⁰	Per 0.06 mmol/L higher 1h OGTT Whitehall criteria: 1.00 (1.00-1.01) Pooling Project Major Abnormalities: 1.01 (1.00-1.01)	Per 0.06 mmol/L higher 1h: 1.00 (0.99-1.01)	++++	No	None
Case control					
Bangalore case control study ¹⁴¹		2h OGTT per 1SD higher: 2.01 (1.56-2.59) FPG (mmol/L): <4.6 Ref 4.6-5.2 1.2 (0.8-2.0) 5.21-6.3 2.8 (1.7-4.6) >6.3 3.6 (1.4-8.7)	PPPG: +++ FPG: -	No	None

1h OGTT, 1 hour post-load glucose in oral glucose tolerance test; 2h OGTT, 2 hour post-load glucose in oral glucose tolerance test; CI, confidence interval; FPG, fasting plasma glucose; IHD, ischaemic heart disease; MI, myocardial infarction; OR, odds ratio; PPPG, post-prandial plasma glucose; Q, quintile; SBP, systolic blood pressure; SD, standard deviation; SE, standard error. Adjustment/matching as Table 2.4.

Table 2.13. Characteristics of studies examining the relationship between blood glucose levels and incident ischaemic heart disease

Study (year established)	No. of participants (% female)	Geographical location	Mean age at baseline (years)	Glycaemic measure	Mean follow-up (years)	No. of events
Cohort						
Korean Cancer Prevention Study (1992) ¹⁰²	1 197 384 (36)	Korea	46.6	FPG	16	58 148
Tabular/ PD meta-analysis						
Varied (N/A) ¹⁰⁵	303 961 (N/S)	Varied	N/S	FBG OGTT	15.6	16 982
Varied (N/A) ¹⁰⁴	191 249 (27)	Varied	N/S	FBG 1h OGTT 2h OGTT RBG	13.7	FBG: 492 1h OGTT: 1659 2h OGTT: 4458 RBG: 314
Varied (N/A) ¹⁴²	372 000 (N/S)	Varied	N/S	FPG	N/S	N/S
IPD meta-analysis						
Asia Pacific Cohort Studies Collaboration (N/A) ¹⁰³	237 468 (44)	Asia Pacific	47.4	FBG	5	FBG 816 RBG 240
Emerging Risk Factors Collaboration (N/A) ⁶	260 361 (N/S)	Varied	52.0	FBG	10.5	13 248

1h OGTT, 1 hour post-load glucose in oral glucose tolerance test; 2h OGTT, 2 hour post-load glucose in oral glucose tolerance test; FBG, fasting blood glucose; FPG, fasting plasma glucose; IPD, individual participant data; N/A, not applicable; N/S, not stated; PD, published data; RBG, random blood glucose.

Table 2.14. Key findings of studies examining the relationship between blood glucose levels and incident ischaemic heart disease

Study	HR (95%CI)		Adjustment	Correction for regression dilution bias	Variables investigated for effect modification
	Fasting glucose	Random or post-load glucose			
Cohort					
Korean Cancer Prevention Study ¹⁰²	IHD (mmol/l)^a M: <3.9 1.00 (0.96-1.06) 3.9-4.7 1.00 (0.97-1.02) 4.8-5.5 Reference 5.6-6.0 1.04 (1.01-1.08) 6.1-6.9 1.14 (1.10-1.19) 7.0-7.7 1.27 (1.18-1.37) ≥7.8 1.79 (1.71-1.87) F: <3.9 0.95 (0.88-1.02) 3.9-4.7 0.97 (0.94-1.01) 4.8-5.5 Reference 5.6-6.0 1.00 (0.95-1.05) 6.1-6.9 1.15 (1.08-1.23) 7.0-7.7 1.36 (1.23-1.51) ≥7.8 2.12 (1.98-2.26) MI (mmol/l)^a M: <3.9 1.02 (0.94-1.11) 3.9-4.7 1.00 (0.96-1.05) 4.8-5.5 Reference 5.6-6.0 1.02 (0.97-1.08) 6.1-6.9 1.17 (1.09-1.25) 7.0-7.7 1.50 (1.34-1.69) ≥7.8 2.13 (1.97-2.29) F: <3.9 0.97 (0.83-1.14) 3.9-4.7 1.00 (0.92-1.08) 4.8-5.5 Reference 5.6-6.0 1.08 (0.98-1.20) 6.1-6.9 1.37 (1.21-1.54) 7.0-7.7 1.91 (1.59-2.29) ≥7.8 3.18 (2.84-3.57)	N/A	++++	No	Sex

Table 2.14 continued

Study	HR (95%CI)		Adjustment	Correction for regression dilution bias	Variables investigated for effect modification
	Fasting glucose	Random or post-load glucose			
Tabular/ PD meta-analysis					
Varied ¹⁰⁵	IHD Per 1 mmol/L higher: 1.06 (1.00-1.12)	IHD Per 1 mmol/L higher: 1.05 (1.03-1.07)	++++	No	None
Varied ¹⁰⁴	IHD Per 1 mmol/L higher: 1.16 (1.04-1.30)	IHD Per 1 mmol/L higher: 2hr post-load glucose: 1.11 (1.03-1.20) 1hr post-load glucose: 1.10 (1.06-1.15) RBG: 1.02 (0.98-1.07)	N/S	No	None
Varied ¹⁴²	IHD per 1 mmol/L higher: 35-44 years: 1.21 (1.02-1.43) 45-54 years: 1.19 (1.05-1.36) 55-64 years: 1.18 (1.08-1.29) 65-74 years: 1.16 (1.11-1.21) 75-84 years: 1.16 (1.11-1.20) ≥85 years: 1.14 (1.07-1.22)	N/A	+	Yes	Age
IPD meta-analysis					
Asia Pacific Cohort Studies Collaboration ¹⁰³	IHD Per 1 mmol/L higher: 1.23 (1.19-1.30)	N/A	++++	Yes	Age, sex, region
Emerging Risk Factors Collaboration ⁶	IHD Per 1 mmol/L higher ^b : 1.12 (1.08-1.15)	N/A	++++	Yes	None

^aConverted from mg/dL with some overlap in groups due to rounding; ^bAt FBG levels >5.6mmol/L.

1h OGTT, 1 hour post-load glucose in oral glucose tolerance test; 2h OGTT, 2 hour post-load glucose in oral glucose tolerance test; CI, confidence interval; F, female; FBG, fasting blood glucose; HR, hazard ratio; IHD, ischaemic heart disease; IPD, individual participant data; M, male; N/A, not applicable; PD, published data; RBG, random blood glucose. Adjustment as Table 2.4.

Table 2.15. Characteristics of studies examining the relationship between blood glucose levels and incident stroke

Study (year established)	No. of participants (% female)	Geographical location	Mean age at baseline (years)	Glycaemic measure	Mean follow-up (years)	No. of events
Cohort						
Aerobics Centre Longitudinal Study (1971) ¹⁴⁴	43 933 (0)	USA	44.4	FPG	Fatal: 15.3 Non-fatal: 17.7	TS: 612
Korean Cancer Prevention Study (1992) ¹⁰²	1 197 384 (36)	Korea	46.6	FPG	16	TS: 74 631 IS:46 596 HS:19 407
Tabular/ PD meta-analysis						
Varied (N/A) ¹⁰⁴	191 249 (N/S)	Varied	N/S	1h OGTT 2h OGTT RBG	13.7	TS: 2-hr: 2387 1hr: 257 RBG: 544
Varied (N/A) ¹⁴²	372 000 (N/S)	Varied	N/S	FPG	N/S	N/S
IPD meta-analysis						
Asia Pacific Cohort Studies Collaboration (N/A) ¹⁰³	237 468 (44)	Asia Pacific	47.4	FBG	5	TS: 1661
Diabetes epidemiology: collaborative analysis of diagnostic criteria in Europe (N/A) ⁸³	21 906 (45)	Europe	56.8	FPG 2h OGTT	10.5	TS: 255
Emerging Risk Factors Collaboration (N/A) ⁶	260 361 (N/S)	Varied	52.0	FBG	10.5	IS: 1754

1h OGTT, 1 hour post-load glucose in oral glucose tolerance test; 2h OGTT, 2 hour post-load glucose in oral glucose tolerance test; FBG, fasting blood glucose; FPG, fasting plasma glucose; HS, haemorrhagic stroke; IPD, individual participant data; IS, ischaemic stroke; N/A, not applicable; N/S, not stated; PD, published data; RBG, random blood glucose; TS, total stroke.

Table 2.16. Key findings of studies examining the relationship between blood glucose levels and incident stroke

Study	HR (95%CI)		Adjustment	Correction for regression dilution bias	Variables investigated for effect modification
	Fasting glucose	Random or post-load glucose			
Cohort					
Aerobics Centre Longitudinal Study ¹⁴⁴	TS per 0.55 mmol/L higher above 6.1 mmol/L: 1.06 (1.00-1.12)	N/A	++++	No	None

Table 2.16 continued

Study	HR (95%CI)		Adjustment	Correction for regression dilution bias	Variables investigated for effect modification
	Fasting glucose	Random or post-load glucose			
Korean Cancer Prevention Study ¹⁰²	IS (mmol/L): ^a				
	M				
	<3.9	1.06 (1.00-1.12)			
	3.9-4.7	1.00 (0.98-1.04)			
	4.8-5.5	Ref			
	5.6-6.0	1.08 (1.04-1.12)			
	6.1-6.9	1.21 (1.16-1.26)			
	7.0-7.7	1.51 (1.41-1.63)			
	≥7.8	2.20 (2.10-2.31)			
	F				
	<3.9	1.05 (0.98-1.13)			
	3.9-4.7	1.03 (0.99-1.06)			
	4.8-5.5	Ref			
	5.6-6.0	1.09 (1.04-1.14)			
	6.1-6.9	1.23 (1.16-1.30)			
	7.0-7.7	1.62 (1.48-1.78)			
	≥7.8	2.29 (2.15-2.44)			
	HS (mmol/L): ^a				
	M				
	<3.9	1.02 (0.94-1.11)			
	3.9-4.7	1.00 (0.95-1.04)			
	4.8-5.5	Ref			
	5.6-6.0	0.98 (0.93-1.04)			
	6.1-6.9	1.08 (1.02-1.16)			
	7.0-7.7	1.15 (1.01-1.30)			
	≥7.8	1.24 (1.13-1.35)			
	F				
	<3.9	1.16 (1.04-1.29)			
	3.9-4.7	1.04 (0.99-1.10)			
	4.8-5.5	Ref			
	5.6-6.0	0.96 (0.89-1.04)			
	6.1-6.9	0.99 (0.89-1.09)			
	7.0-7.7	1.12 (0.93-1.33)			
	≥7.8	1.14 (1.00-1.30)			

Table 2.16 continued

Study	HR (95%CI)		Adjustment	Correction for regression dilution bias	Variables investigated for effect modification
	Fasting glucose	Random or post-load glucose			
Tabular/ PD meta-analysis					
Varied ¹⁰⁴	N/A	TS per 1 mmol/L higher: 2hr OGTT 1.10 (0.94-1.29) 1hr OGTT 1.11 (1.06-1.17) RBG 1.11 (0.95-1.31)	N/S	No	None
Varied ¹⁴²	TS per 1 mmol/L higher: 35-44 years: 1.19 (0.93-1.51) 45-54 years: 1.16 (0.97-1.39) 55-64 years: 1.14 (1.01-1.29) 65-74 years: 1.14 (1.08-1.20) 75-84 years: 1.10 (1.06-1.15) ≥85 years: 1.06 (0.98-1.16)	N/A	N/S	Yes	Age
IPD meta-analysis					
Asia Pacific Cohort Studies Collaboration ¹⁰³	Per 1 mmol/L lower: TS: 21% (18-24%) lower risk IS: 17% (-1-32%) lower risk HS: 20% (0-36%) lower risk	N/A	++++	Yes	Sex, region, age
Diabetes epidemiology: collaborative analysis of diagnostic criteria in Europe ⁸³	TS per 1SD higher FPG: M 1.02 (0.83-1.25) F 1.52 (1.22-1.88)	TS per 1SD higher: M 1.21 (1.06-1.38) F 1.31 (1.06-1.61)	++++	No	Sex
Emerging Risk Factors Collaboration ⁶	N/A	N/A	++++	Yes	None

^aConverted from mg/dL with some overlap in groups due to rounding.

1h OGTT, 1 hour post-load glucose in oral glucose tolerance test; 2h OGTT, 2 hour post-load glucose in oral glucose tolerance test; BG, blood glucose; CI, confidence interval; F, female; FPG, fasting plasma glucose; HR, hazard ratio; HS, haemorrhagic stroke; HTN, hypertension; IPD, individual participant data; IS, ischaemic stroke; M, male; N/A, not applicable; PD, published data; Ref, reference group; RBG, random blood glucose; SD, standard deviation; TS, total stroke. Adjustment as Table 2.4.

Table 2.17. Characteristics of Mendelian randomisation studies examining the relationship between blood glucose and cardiovascular disease

Study (year established)	No. of participants (% female)	Geographical location	Ethnicity	Mean age (years)	No. of SNPs	Glycaemic measure	Cardiovascular outcome
Copenhagen City Heart Study, Copenhagen General Population Study and Copenhagen Ischemic Heart Disease Study (1976) ¹⁴⁸	80522 (53)	Denmark	Caucasian	56.4	5	RBG	IHD, n=14 155 MI, n=6247
Atherosclerosis Risk in Communities Study (1987) ¹⁴⁹	7260 (56)	USA	Caucasian	53.9	5	FPG	CIMT
Multiple genetic consortia (N/S) ¹⁵⁰	N/A (N/S)	Varied	Caucasian	N/S	30	FPG	IHD

CIMT, carotid intima-media thickness; FPG, fasting plasma glucose; IHD, ischaemic heart disease; MI, myocardial infarction; N/A, not applicable; N/S, not stated; RBG, random blood glucose; SNP, single nucleotide polymorphism.

Table 2.18. Key findings of Mendelian randomisation studies examining the relationship between blood glucose and cardiovascular disease

Study	HR (95%CI)		Correction of observational association for regression dilution bias
	Observational	Genetic	
Copenhagen City Heart Study, Copenhagen General Population Study and Copenhagen Ischemic Heart Disease Study ¹⁴⁸	Per 1 mmol/L higher RBG: IHD: 1.18 (1.15-1.22) MI: 1.09 (1.07-1.11)	Per 1 mmol/L higher RBG: IHD: 1.25 (1.03-1.52) MI: 1.69 (1.28-2.23)	Yes
Atherosclerosis Risk in Communities Study ¹⁴⁹	Per 0.5 mmol/L higher FPG: CIMT ^a : 0.0103mm	Per 0.5 mmol/L higher FPG: CIMT ^a : 0.0281mm	No
Multiple genetic consortia ¹⁵⁰	Per 1 mmol/L higher FPG: 1.02 (1.02-1.03)	Per 1 mmol/L higher FPG: 1.16 (0.94-1.42)	N/S

^aChange in CIMT.

CI, confidence interval; CIMT, carotid intima-media thickness; FPG, fasting plasma glucose; HR, hazard ratio; IHD, ischaemic heart disease; N/S, not stated; RBG, random blood glucose.

Chapter 3. China Kadoorie Biobank methods and study population

3.1 Background

The CKB is the largest ever prospective study of lifestyle, environmental and genetic factors as determinants of chronic diseases in China. Although studies, both retrospective¹⁵⁴⁻¹⁵⁶ and prospective,¹⁵⁷⁻¹⁶³ have been undertaken previously to examine the causes and associations of chronic diseases in China, they have been limited, variably, by their retrospective design,¹⁵⁴⁻¹⁵⁶ inadequate sample sizes, restriction to specific study populations or relatively limited collection of exposure and/or outcome data.¹⁶⁴ The scale of the CKB is such that it enables investigation of the typically relatively modest associations of many risk factors with chronic diseases,¹⁶⁵ which many previous studies have been too small to reliably and precisely characterise. Furthermore, the diversity of the study population and collection of information on a wide range of exposures, including through blood sampling, means the CKB is ideally placed to facilitate a clearer understanding of the interplay of environmental, lifestyle, genetic and other exposures in the determination of chronic diseases. In addition, the extended follow-up for cause-specific morbidity and mortality, with robust phenotyping, enables insight into a wide range of chronic diseases.^{7, 164, 166} The CKB thus provides a valuable resource to investigate some of the current gaps in understanding of the roles of blood glucose, and of diabetes, as CVD risk factors, with particular relevance to China, but also to populations more generally. Many of the proposed differences in diabetes pathophysiology between Chinese and Western populations are relevant to Asian, and particularly East Asian, populations more broadly.^{88, 89} These populations likely also share, to some extent, similarities in demographic, socioeconomic and cultural factors.⁸⁸ Whilst there is a need for improved understanding of the associations of diabetes, and of blood glucose levels throughout the full glycaemic range, with the risk of CVDs in China, this is true also of East Asian populations more widely, and the

understanding of these associations gained from the CKB will be of relevance and importance to these populations. This chapter describes the design and methods of the CKB study, and provides a description of the baseline characteristics of the study population.

3.2 Study design

The CKB is a large, blood-based prospective study of over half a million middle-aged Chinese men and women established in 2004, through collaboration between the Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU) of the University of Oxford and the China National Centre for Disease Control and Prevention (CDC).^{7, 164} The study aims to reliably assess the effects of risk factors, both established and emerging, for a wide range of diseases through collection of questionnaire data, physical measurements and blood samples at recruitment, and subsequent follow-up for cause-specific morbidity and mortality over a planned period of at least two decades.

Ethical approval for the study was obtained, prior to commencement, from Oxford University, the China National CDC and the 10 study areas' local CDCs.

3.3 Study population

The baseline survey took place in 10 localities across China (Figure 3.1)—five urban and five rural, based upon standard governmental definitions—between 2004 and 2008. The study areas were chosen to provide diversity in disease patterns and exposures, additionally taking account of logistical considerations, including population stability, quality of death and disease registries and local capacity and commitment. In each area, all permanent residents aged 35-74 years from 100 to 150 administrative units (rural villages or urban residential committees) were identified through official residential records, then invited to participate by letter (accompanied by study information leaflets) following extensive publicity campaigns. The estimated overall population response rate for the study was approximately 30%, ranging between 26% and 38% in rural areas and between 16%

and 50% in urban areas. Overall, 512 891 individuals participated (including a few slightly outside the 35-74 year age range). The CKB was not designed to be representative of the Chinese adult population, as its main objectives are to assess the determinants of common chronic disease in the Chinese population. The very large sample size and the diversity of the population covered (eg, urban and rural, men and women, young and old, and people of different socioeconomic status) are such that the lack of representativeness would not be expected to bias the estimated effect sizes¹⁶⁷ which would be expected to be generalisable to the Chinese population more generally.

3.4 Baseline data collection

During the baseline survey, study assessment centres were set up in the local community and eligible participants were invited to attend; eligibility was confirmed using participants' unique national identity cards. All participants provided informed, written consent prior to participation, which included permission for long-term storage of blood for anonymised and non-specified medical research, and access to their medical records. Participants moved through 10 stations in the assessment centre, and took approximately 60-75 minutes to complete the entire process. They completed an interviewer-administered, laptop-based electronic questionnaire (with range and logic checks to minimise missing values, errors and inconsistencies), collating information on demographic and socioeconomic characteristics, lifestyle (including diet, smoking, alcohol and physical activity), personal and family medical history (including a history of doctor-diagnosed diabetes, IHD—MI or angina—and stroke/TIA), mental health and, for women, reproductive history. Participants reporting a history of doctor-diagnosed diabetes, IHD, stroke/TIA or hypertension were also asked about their age at diagnosis and current use of medications. Physical measurements were undertaken including height, weight, hip and waist circumferences, bio-impedance, lung function, heart rate and blood pressure using standard, calibrated instruments with standard protocols.¹⁶⁸⁻¹⁷²

A 10ml non-fasting (except in one study area—Zhejiang—where participants were asked to fast to allow for specific, locally arranged tests) blood sample was collected from participants into an ethylenediaminetetraacetic acid (EDTA) vacutainer (BD Hemogard™, BD, Franklin Lakes, New Jersey, USA), with the time passed since participants last ate recorded. Immediate, on-site testing of plasma glucose level was undertaken using the Johnson and Johnson SureStep Plus meter (Lifespan, Milpitas, California, USA). Participants with a glucose level ≥ 7.8 mmol/L and < 11.1 mmol/L were invited to return for a FPG test the following day. The filled EDTA vacutainers were kept at 0-4°C with ice packs for a few hours and were then transported to the local study laboratory where they were immediately centrifuged and aliquoted into four cryovials, including one DNA-containing buffy coat. The cryovials were stored at -40°C for 3-4 months then transported on dry ice to Beijing where they were stored at -80°C. Aliquots of plasma were subsequently transported on dry ice to Oxford for long-term storage in liquid nitrogen tanks, and the buffy coat sample remained in storage in Beijing.

A randomly selected sample of ~3% of participants was invited to participate in a quality control survey within a few weeks of the initial baseline survey. This involved recompletion of the baseline survey and some repeat measurements, but not of plasma glucose. Additional quality control was achieved through regular central monitoring of recruitment, collected data and blood processing times, and through six monthly and annual site visits by regional CDC staff and by CTSU and China CDC staff, respectively.

A total of 354 local staff participated in baseline data collection. 95% of data were collected by 21 staff for initial registration and consenting, 113 for the main questionnaire, 20 for anthropometric measurements, 24 for blood pressure measurement and blood sample collection, and 23 for lung function and carbon monoxide measurement. Extensive training was provided for field survey staff, including instruction on data collection using the electronic questionnaire, on physical measurements using standard protocols and on use and maintenance of equipment.

3.5 DNA extraction and genotyping

DNA was extracted from 800 µl buffy coat from 95 680 randomly selected CKB participants (including 66 outside the 30-79 year age range included in this thesis), using a magnetic bead purification method (KingFisher™ Flex Magnetic Particle Processors; Thermo Fisher Scientific Inc., USA). An array containing a panel of 384 SNPs (GoldenGate Genotyping Assay; Illumina® Inc., USA) was then used to genotype the samples. SNPs were selected for inclusion based on several criteria, specifically: demonstrated associations with the outcomes of interest in Chinese or other East Asian populations, particularly in large-scale GWAS; demonstrated associations with the outcomes of interest in other populations, with conserved LD across populations; functional variants (eg, non-synonymous substitutions); population-specific haplotypes at loci of specific interest; pleiotropic effects on multiple outcomes of interest; and perceived biological significance. The array included 56 type 2 diabetes-related SNPs (Table B.1).

Additional genotyping of ~100 000 participants is currently underway, using a bespoke, genome-wide array developed by the CKB team using a similar approach to that of the UK Biobank, but customised specifically for a Chinese population. These genome-wide data were not available for inclusion in this thesis.

3.6 Resurvey

On a five yearly basis a 5% randomly selected (using cluster random sampling with the administrative unit as the sampling unit) sample of surviving participants from each study area is invited to participate in a resurvey, collecting the same data as at baseline with a small number of additions. These resurveys enable assessment of within-person variation in exposures and of consistency of data, and collection of additional information to enhance and calibrate baseline data collection. Resurveys took place in 2008 and 2013-2014. 24 760 participants were invited to participate in the 2008 resurvey, and a participation rate of 79.9% was achieved (n=19 788). Measurement of plasma glucose in the first resurvey was

undertaken using the same methods as were used in the baseline survey, and participants in Zhejiang were again asked to fast prior to attending the assessment centre. 33 073 participants were invited to participate in the second resurvey, with a 75.6% participation rate (n=25 011). As the SureStep Plus meter was no longer available, the On Call Advanced blood glucose meter (ACON Laboratories Inc., San Diego, California, USA) was used in the second resurvey. Participants in Zhejiang were not asked to fast prior to attending the assessment centre for the second resurvey.

Random measurement error, as well as biological variation in individuals over time, is such that a single exposure measurement at baseline is unlikely to reflect an individual's true or usual level of the exposure.¹⁷³⁻¹⁷⁵ Most prospective studies collect exposure measurements only once, at baseline.¹⁷⁴ However, by relying on single-exposure measurements, these studies do not take account of this described variation. As a result, analyses of disease associations will likely result in under- (or occasionally over-) estimation of the true association of the risk factor with the disease of interest.^{173, 175} Repeat exposure measurements in individuals will tend to fluctuate around their usual exposure level.¹⁷⁵ Thus, through the use of repeat measurements it is possible to quantify and adjust for this intra-individual variation, and thereby estimate disease risks associated with usual, or longer-term average, exposure levels.¹⁷³⁻¹⁷⁵ Repeat exposure measurements on an entire study population are frequently impractical, particularly in large studies. They are therefore more frequently taken on a representative subsample,¹⁷⁴ as in the CKB. Since there will be error in estimates of the extent of intra-individual variation, which will be transferred to estimates of disease risk associated with usual exposure levels, these estimates must be based on an adequately-sized subsample.¹⁷⁵ It has been shown that the extent of intra-individual variation varies over time, becoming more marked with longer intervals between repeat measures.¹⁷³ Because of this, when correcting disease associations for intra-individual variation in exposures, repeat measurements should ideally be taken at a time period approximately equivalent to the mid-point of the follow-up period.¹⁷³ Hence the need,

in studies with prolonged follow-up periods, for multiple resurveys of at least a subset of the population.

3.7 Follow-up and outcome adjudication

Study participants were followed-up for cause specific morbidity and mortality. For incident disease events eight study areas (Harbin, Suzhou, Liuzhou, Sichuan, Gansu, Henan, Zhejiang, and Hunan) had established disease registries from which information was obtained on stroke, IHD, cancer and diabetes. One additional study area (Qingdao) had only a cancer registry. Additionally, established national HI databases in all study areas provided information on International Classification of Diseases (ICD)-10 coded¹⁷⁶ hospitalisations and coded examination and treatment procedures through electronic linkage via unique national ID. More than 98% of CKB participants were covered by this HI system.

Causes of death were derived through China's Disease Surveillance Points¹⁷⁷ which cover all study areas. Death certificates were reviewed, cross-checked with study participant records and ICD-10¹⁷⁶ coded by regional CDC staff, blinded to participant baseline information. Causes of death were supplemented by review of available medical records. In the small proportion of deaths without prior medical attention (<5%), verbal autopsies were conducted with family members using a validated instrument.¹⁷⁸

Validation exercises are underway to examine the accuracy of a number of main disease outcomes of interest within the CKB. Medical records of participants with cancer, IHD, stroke, chronic obstructive pulmonary disease and diabetes events recorded in any of the three follow-up data sources are retrieved by regional CDC staff and copies are made of documentation relating to the relevant disease events, which are then up-loaded to a secure study server. The retrieved information is reviewed by clinicians at the CKB National or International Co-ordinating Centres, who are blinded to baseline and resurvey data, in order to confirm, or otherwise, disease events, and, for some conditions, to enable more detailed

phenotyping. This process is on-going for IHD and stroke events, in particular to enable improved sub-phenotyping, but initial findings of the validation process are discussed in Section 6.4.3.

Residential records in the study areas were reviewed annually which, in addition to visits to local communities, enabled identification of participants who had permanently moved out of study areas.

3.8 Baseline characteristics of China Kadoorie Biobank participants

Overall, 512 891 participants (210 259 men and 302 632 women) were recruited to the study (Table 3.1), with 55.9% recruited from rural areas. Across the 10 study areas the number of recruited participants ranged from 29 689 (Haikou) to 63 357 (Henan). The mean (SD) study area-standardised age was 52.4 (10.9) years in men and 51.0 (10.4) years in women (Table 3.2). 26.1% of men and 17.5% of women were educated to high school level or above and nearly all were married. There was a striking sex-difference in smoking patterns; the age and area-adjusted prevalence of ever regular smoking was 74.2% in men and 3.3% in women ($p<0.001$). There were also large sex differences in alcohol consumption, with 33.4% of men reporting weekly alcohol consumption, compared with just 2.1% of women ($p<0.001$). Men reported mean (SD) daily physical activity—relating to leisure, occupation and commuting—equivalent to 22.0 (14.2) metabolic equivalent of task (MET) hours/day, whilst women reported a level of 20.4 (11.6) MET hours/day ($p<0.001$). Regular consumption (defined as consumption on at least 4-6 days per week) of fresh vegetables was almost universal. The proportion of participants regularly consuming other foodstuffs varied considerably, but with limited differences between sexes.

Hypertension was the most commonly reported doctor-diagnosed, chronic condition (men: 10.8%; women 12.3%, $p<0.001$), with markedly lower (generally $<3\%$) self-reported prevalence of most other conditions at baseline. Mean (SD) age- and study area-adjusted SBP was slightly higher in men than in women (132.3 [18.8] mmHg and 130.5 [20.2] mmHg,

respectively, $p < 0.001$) (Table 3.3). Mean (SD) BMI was similar between the sexes (23.5 [3.1] kg/m² vs. 23.8 [3.3] kg/m², $p < 0.001$). Men had a higher mean (SD) waist circumference than women (82.1 [9.2] cm vs. 79.1 [9.1] cm, $p < 0.001$), whilst mean (SD) hip circumference was slightly higher in women (90.7 [5.9] cm vs. 91.1 [6.3] cm, $p < 0.001$). Consequently, men had a modestly higher average (SD) waist-to-hip ratio (0.90 [0.06] vs. 0.88 [0.07], $p < 0.001$), but a lower average (SD) percentage body fat (22.1 [5.9] % vs. 32.1 [6.9%], $p < 0.001$). The mean (SD) RPG level in the entire population (irrespective of diabetes status) was 6.1 (2.3) mmol/L (men: 6.0 [2.3] mmol/L; women: 6.2 [2.3] mmol/L, $p < 0.001$).

3.9 Baseline characteristics of the China Kadoorie Biobank resurvey populations

Selected baseline characteristics of the 19 788 participants included in the first resurvey and the 25 011 participants included in the second resurvey, and of the total study population, are presented in Table 3.4. Although there were some statistically significant differences between the first and second resurvey populations and the total study population, these were not clinically significant.

3.10 Baseline characteristics of the China Kadoorie Biobank genotyped population

A small number of samples were excluded in initial quality control on the genotyping data due to linkage errors, gender non-concordance, low call rates and high heterozygosity (described more fully in Section 8.2.2). Selected relevant baseline characteristics of the remaining 93 131 participants (after exclusion of 66 outside the 30-79 year age range) are presented in Table 3.5; these are reasonably comparable with those of the entire study population, despite some statistically significant differences which likely reflect the large sample size rather than meaningful, clinically significant differences.

3.11 Discussion

The CKB is one of the largest ever blood-based prospective studies. The uniquely large and diverse study population, rich baseline data, repeat exposure measurements at regular intervals during follow-up, and well-phenotyped disease incidence data on diabetes, IHD and stroke (including stroke subtypes) enable uniquely reliable assessment of relationships of diabetes and of plasma glucose levels with CVDs. Furthermore, the availability of genetic data, currently for a subset of the study population, facilitates more in-depth understanding of these relationships, and, in particular, the likelihood of causal associations.

Figure 3.1. Locations of the China Kadoorie Biobank recruitment centres

Map produced by Dr Hongchao Pan, CTSU University of Oxford.

- Urban
- Rural

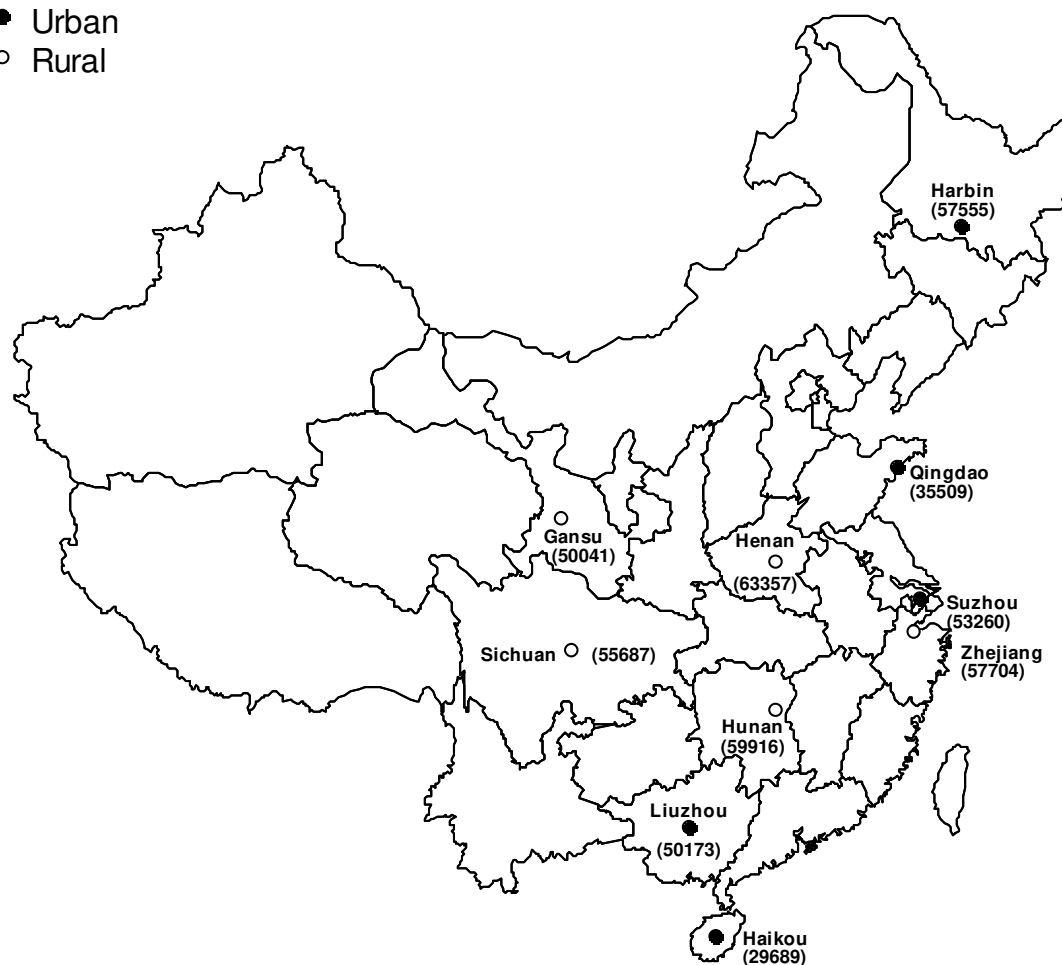


Table 3.1. Proportion of participants recruited in study areas, by sex

Study area	Men (<i>n</i> =210 259)		Women (<i>n</i> =302 632)	
	<i>n</i>	%	<i>n</i>	%
Urban				
Qingdao	15626	7.4	19883	6.6
Harbin	23254	11.1	34301	11.3
Haikou	10792	5.1	18897	6.2
Suzhou	22362	10.6	30898	10.2
Liuzhou	19322	9.2	30851	10.2
Total	91356	43.5	134830	44.6
Rural				
Sichuan	21315	10.1	34372	11.4
Gansu	19345	9.2	30696	10.1
Henan	27841	13.2	35516	11.7
Zhejiang	24027	11.4	33677	11.1
Hunan	26375	12.5	33541	11.1
Total	118903	56.6	167802	55.5

Table 3.2. Baseline demographic and lifestyle characteristics of participants by sex

Characteristic	Men, % (n=210 259)	Women, % (n=302 632)
Age (years)^a		
30-59	72.4	77.9
60-69	19.7	16.6
70-79	7.8	5.5
<i>Mean (SD)</i>	<i>52.4 (10.9)</i>	<i>51.0 (10.4)</i>
Marital status		
Married with spouse	93.3	88.4
Widowed, separated or divorced	5.3	11.4
Never married	1.5	0.2
Highest level of education		
No formal schooling	8.3	26.2
Primary school	32.3	31.6
Middle school	33.2	24.7
High school	18.0	13.1
College or university	8.1	4.4
Annual household income (Yuan/year)		
<2500	2.6	3.4
2500-4999	6.2	7.1
5000-9999	17.1	19.3
10 000-19 999	28.4	29.5
20 000-34 999	25.2	24.3
≥35 000	20.3	16.4
Smoking		
Never regular	14.4	94.8
Occasional	11.4	1.8
Ex-regular	12.9	0.9
Current regular	61.3	2.4
Alcohol consumption		
Never regular	19.9	63.7
Occasional	43.1	33.8
Ex-regular	3.6	0.5
Current regular	33.4	2.1
Regular consumption of foodstuffs		
Fresh fruit	23.4	31.3
Fresh vegetables	98.2	98.4
Meat/ poultry	53.0	43.6
Soybean	10.6	9.4
Fish/ seafood	9.7	8.3
Dairy products	10.8	12.6
Fresh eggs	25.3	23.7
Physical activity (MET hours/day)		
<10	24.7	23.1
10-14.9	14.1	22.0
≥15	61.2	54.9
<i>Mean (SD)</i>	<i>22.0 (14.2)</i>	<i>20.4 (11.6)</i>
Self-reported prevalent disease		
Ischaemic heart disease	2.6	3.3
Stroke/transient ischaemic attack	2.2	1.4
Diabetes	2.9	3.4
Hypertension	10.8	12.3
Chronic respiratory diseases	3.3	2.7
Chronic hepatitis/ cirrhosis	1.8	0.8
Chronic kidney disease	1.3	1.6
Cancer	0.4	0.5

Standardised to the age and study area structure of the study population. ^aStandardised to study area structure only. MET, metabolic equivalent of task.

Table 3.3. Mean values and distributions of physical measurements and plasma glucose by sex

Characteristic	Men, % (n=210 259)	Women, % (n=302 632)
SBP (mmHg)		
<120	26.9	34.8
120-139	43.6	36.5
≥140	29.5	28.6
<i>Mean (SD)</i>	<i>132.3 (18.8)</i>	<i>130.5 (20.2)</i>
Height (cm)		
<155	5.3	55.5
155-159	14.6	28.0
160-164	27.7	12.9
≥165	52.5	3.5
<i>Mean (SD)</i>	<i>165.3 (5.8)</i>	<i>154.1 (5.3)</i>
BMI (kg/m²)^a		
<22.0	35.0	31.2
22.0 to <25.0	33.8	34.4
≥25.0	31.2	34.3
<i>Mean (SD)</i>	<i>23.5 (3.1)</i>	<i>23.8 (3.3)</i>
Waist circumference (cm)		
<80	43.2	55.1
80-89	34.7	31.4
≥90	22.2	13.5
<i>Mean (SD)</i>	<i>82.1 (9.2)</i>	<i>79.1 (9.1)</i>
Hip circumference (cm)		
<88	35.4	32.8
88-92.9	28.7	30.1
≥93	35.9	37.1
<i>Mean (SD)</i>	<i>90.7 (5.9)</i>	<i>91.1 (6.3)</i>
Waist to hip ratio		
<0.85	18.3	38.8
0.85-0.90	26.6	27.5
≥0.90	55.1	33.7
<i>Mean (SD)</i>	<i>0.90 (0.06)</i>	<i>0.88 (0.07)</i>
Percentage body fat^b		
<15	12.6	0.4
15-24	56.2	15.2
25-34	28.8	51.7
≥35	2.4	32.7
<i>Mean (SD)</i>	<i>22.1 (5.9)</i>	<i>32.1 (6.9)</i>
Random plasma glucose (mmol/L)^c		
<5.8	61.4	54.6
5.8-7.7	30.0	36.4
7.8-11.0	5.6	5.9
≥11.1	2.9	3.1
<i>Mean (SD)</i>	<i>6.0 (2.3)</i>	<i>6.2 (2.3)</i>

Standardised to the age and study area structure of the study population.

^aData missing for 2 participants; ^bData missing for 241 participants; ^cData missing for 8341 participants.

BMI, body mass index; SBP, systolic blood pressure.

Table 3.4. Baseline characteristics of the entire study population and of resurvey participants

Characteristic	Total population, % (n=512 891)	First resurvey (n=19 788)		Second resurvey (n=25 011)	
		%	p value ^a	%	p value ^a
Men, %	41.0	39.3	<0.001	38.3	<0.001
Age (years), mean	51.5 (10.7)	51.6 (10.6)	0.2	51.0 (10.1)	<0.001
Living in urban area, %	44.1	41.2	<0.001	43.1	0.002
≥6 years of education, %	49.1	47.3	<0.001	47.3	<0.001
Current smoker, %	26.4	25.2	<0.001	24.6	<0.001
Regular alcohol consumption, %	14.9	14.5	0.2	14.6	0.3
Physical activity (MET hours/day), mean	21.1 (13.9)	20.7 (14.2)	<0.001	21.4 (14.0)	<0.001
SBP (mmHg), mean	131.1 (21.3)	132.2 (21.8)	<0.001	131.0 (20.8)	0.5
BMI (kg/m²), mean	23.7 (3.4)	23.7 (3.4)	1.0	23.7 (3.4)	1.0
Random plasma glucose (mmol/L), mean	6.1 (2.3)	5.9 (2.3)	<0.001	5.9 (2.1)	<0.001
Self-reported prevalent disease, %					
Ischaemic heart disease	3.0	3.4	<0.001	2.8	0.1
Stroke/transient ischaemic attack	1.7	1.6	0.2	2.9	<0.001
Diabetes	3.2	3.1	0.9	2.6	<0.001
Hypertension	11.6	11.3	0.1	10.4	<0.001

^aComparison with total study population.

BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

Table 3.5. Baseline characteristics of the entire study population and genotyped population

Characteristic	Total population, % (n=512 891)	Genotyped population, % (n=93 131)	p value
Men, %	41.0	40.5	0.002
Age (years), mean	51.5 (10.7)	51.4 (10.7)	0.008
Living in urban area, %	44.1	40.8	<0.001
≥6 years of education, %	49.1	47.7	<0.001
Current smoker, %	26.4	26.3	0.5
Regular alcohol consumption, %	14.9	14.7	0.2
Physical activity (MET hours/day), mean	21.1 (13.9)	21.6 (14.0)	<0.001
SBP (mmHg), mean	131.1 (21.3)	131.3 (21.3)	0.008
BMI (kg/m²), mean	23.7 (3.4)	23.6 (3.4)	0.001
Random plasma glucose (mmol/L), mean	6.1 (2.3)	6.1 (2.4)	1.0
Self-reported prevalent disease, %			
Ischaemic heart disease	3.0	2.9	0.1
Stroke/transient ischaemic attack	1.7	1.7	0.7
Diabetes	3.2	3.1	0.7
Hypertension	11.6	11.5	0.1

BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

Chapter 4. Prevalence, variability and correlates of diabetes and plasma glucose

4.1 Background

Several studies have estimated the prevalence of diabetes in China,^{87, 179-181} but none, to my knowledge, has described the distribution of blood glucose levels throughout the glycaemic range. The CKB was not designed to be nationally representative, but characterisation of the distribution of diabetes and of plasma glucose levels in the study population enables relevant comparisons. Furthermore, in view of the complex interplay of both exposures with other cardiovascular risk factors,¹⁸² reliable characterisation of their distribution and relationships in the CKB is essential for assessing, and adjusting for, confounding in analyses of their associations with CVDs. As a result of random measurement error and more systematic changes within individuals over time, it is likely that a single exposure measurement will not accurately reflect an individual's true or usual level of the exposure; failure to account for this in analyses of disease associations will result in under- (typically) or over-estimation of effect sizes.¹⁷³ Through the use of repeat exposure measurements it is possible to quantify and adjust for intra-individual variation, enabling estimation of disease risks associated with usual exposure levels, which more accurately reflect real disease associations.

This chapter addresses these issues through cross-sectional analyses of data from the CKB. The objectives of the chapter were to examine a) the distribution of diabetes and plasma glucose levels in the CKB population; b) the associations of diabetes and plasma glucose with other relevant baseline variables; and c) the extent of within-person variation in RPG levels. These analyses are discussed in the context of previous published literature, and their implications for subsequent analyses in this thesis are considered.

4.2 Methods

4.2.1 Assessment of diabetes status

Participants were classified as having self-reported diabetes if they answered yes to the question “Has a doctor ever told you that you had had diabetes?” in the baseline survey. Screen-detected diabetes was defined as not having self-reported diabetes, but having a plasma glucose level ≥ 7.0 mmol/L if time since last food ≥ 8 hours, or ≥ 11.1 mmol/L if time since last food < 8 hours, or a FPG level ≥ 7.0 mmol/L on subsequent testing. This definition was adapted from American Diabetes Association (ADA) recommendations on the diagnosis of diabetes¹⁸³ in light of the observed relationship between plasma glucose levels and fasting period in the CKB (Section 4.3.3). Total diabetes was defined as self-reported or screen-detected diabetes. No specific information was available on diabetes type (eg, type 1, type 2, monogenic diabetes), but an attempt was made to estimate the relative frequencies of type 1 and type 2 diabetes; participants with self-reported diabetes diagnosed before 35 years of age and treated with insulin were classified as type 1 diabetes¹⁸⁴ and all other self-reported diabetes as type 2.

4.2.2 Assessment of plasma glucose

RPG data were not available for 8160 participants without self-reported diabetes at baseline. Most of these participants were recruited early in the study, prior to commencement of plasma glucose testing. Blood samples were not taken, for technical reasons, from 83.2% (n=99) of the 119 participants recruited after 2004 with no plasma glucose data available. Since recruitment commenced at different times in different study areas, the frequency of missing RPG data varied by study area, ranging from 0 in Liuzhou, where recruitment commenced in October 2004, to 3792 in Zhejiang, commencing recruitment in August 2004 (Table 4.1). Overall, participants with missing RPG data were more likely to be older, female, to live in rural areas, to regularly drink alcohol and to have lower SBP than participants with data available (Table 4.2). These differences largely reflect

the characteristics of study areas recruiting a greater proportion of participants earlier, but the small number of participants with missing RPG data limits within-study area comparisons.

The SureStep Plus meter used in the baseline survey and first resurvey provides plasma glucose measurements up to 27.8 mmol/L, above which a reading of “HI” is obtained.¹⁸⁵ At baseline, 231 participants without self-reported diabetes had RPG levels >27.8 mmol/L, of whom 10 were included in the first resurvey, when their mean (SD) RPG level was 17.4 (8.0) mmol/L, and 8 were included in the second resurvey, with a mean (SD) RPG level of 11.1 (4.9) mmol/L. It was thus likely that baseline plasma glucose levels >27.8 mmol/L were truly high levels; in light of these observations, in analyses examining baseline plasma glucose as a continuous variable, levels >27.8 mmol/L were imputed as 27.8 mmol/L, and in analyses examining RPG as a categorical variable these levels were categorised as ≥ 11.1 mmol/L.

4.2.3 Participant exclusions

Participants were excluded from cross-sectional analyses if they had missing data for BMI or extreme or implausible values for SBP, DBP, height, waist circumference, hip circumference, waist-to-hip ratio or BMI, since these extreme values were likely to impact on some of the investigated associations (Figure 4.1). Following these exclusions 511 810 participants (209 904 men and 301 906 women) remained. Analyses examining associations of plasma glucose additionally excluded participants with a self-reported history of diabetes at baseline and with missing baseline RPG data, leaving 487 552 participants (200 433 men and 287 119 women) for inclusion in these analyses.

4.2.4 Statistical analyses

Initial descriptive analyses

Standard descriptive statistics were employed to summarise exposure and potential confounding variables, and to enable familiarisation with the data. Mean and SDs were

calculated for continuous variables and their distributions were examined using histograms. Categorical variables were examined using counts, percentages and bar charts.

The association between self-reported time since last food (fasting period) and plasma glucose level was examined in participants without self-reported diabetes at baseline. General linear models were used to calculate mean age-, sex- and study area-adjusted plasma glucose levels in 1 hour fasting period categories from 0 to <1 hour to 17 to <18 hours and in a ≥ 18 hours fasting period. Mean plasma glucose levels in 1 hour fasting periods <8 hours were compared with a ≥ 8 hour fasting period, and differences were tested using the student t-test.

Correlates of diabetes and random plasma glucose

Prevalence of self-reported and screen-detected diabetes were compared across levels of other baseline variables standardised to the age, sex, study area and education structure of the whole study population. Age-, sex-, education- and study area-adjusted RPG levels among participants without self-reported diabetes were compared across categories of other baseline variables using general linear models. In order to investigate the independence of the demonstrated associations, the analyses were repeated adjusting for age, sex, study area, education, alcohol consumption, smoking, SBP, physical activity, BMI, waist and hip circumferences and frequency of consumption of food stuffs, specifically fruit, vegetables, meat, soybean, fish, dairy, rice and wheat, and of tea. An intermediate level of adjustment was examined for some variables to understand the nature of the associations independent of other particularly relevant potential confounding factors; in addition to basic adjustment (age, sex, study area and education), associations with smoking were adjusted for alcohol consumption, and vice versa. Similarly, associations with general (eg, BMI, percentage body fat) and more specific (eg, waist and hip circumferences) measures of adiposity were adjusted for each other within or across types of measures, and associations with DBP were adjusted for SBP. These analyses were repeated in the prospective

analyses populations, which, in addition to the exclusions described in Section 4.2.3, further excluded individuals with self-reported IHD or stroke/TIA at baseline (Figure 4.1).

Within-person variation in plasma glucose levels

The regression dilution ratio reflects the degree of within-person variation in exposure variables, and is the ratio of the steepness of the association of baseline or single exposure measurements with disease risk, to the steepness of the association of usual exposure levels with disease risk.¹⁷³ It can be calculated using parametric^{186, 187} or non-parametric¹⁷⁴ methods. The non-parametric, or MacMahon's, method is an assumption-free approach. Individuals are divided into groups based on their initial baseline exposure measurement, and the mean repeat exposure measurement is calculated in these groups, with the regression dilution ratio calculated as the ratio of the range of repeat exposure measurements to the range of initial measurements.¹⁷⁴ Rosner's parametric regression method assumes normally distributed exposure measurements and a linear relationship between baseline and repeat measurements.¹⁸⁶ Repeat measurements are regressed against baseline measurements, and the regression dilution ratio is the coefficient for the baseline measurements.¹⁸⁶ This method has additionally been valuably adapted, for example, to incorporate adjustment of confounding variables for intra-individual variation in multi-variable regression analyses or to enable adjustment for regression dilution bias in individual participant data meta-analyses.^{188, 189} Finally, a self-correlation approach can be employed, which assumes a constant variance (the variances of baseline and resurvey exposure measurements are similar).¹⁸⁷ Using this method, the regression dilution ratio is calculated as the correlation coefficient between baseline and repeat measurements.¹⁸⁷ Because the assumptions for the parametric methods did not appear to be met for RPG in the CKB (Section 4.3.5), MacMahon's non-parametric method was used to estimate regression dilution ratios. However, for comparison, Rosner's and the self-correlation methods were also examined.

Usual RPG levels, that is RPG levels adjusted for regression dilution bias, were calculated as the mean resurvey RPG for individuals in five categories defined by baseline RPG levels (<5.8, 5.8-6.7, 6.8-7.7, 7.8-11.0, ≥ 11.1 mmol/L), and the regression dilution ratio was calculated using MacMahon's approach, as described above.¹⁷⁴ Intra-individual variation in exposures increases with longer periods of follow-up, and repeat measures used to calculate regression dilution ratios for prospective analyses should ideally be obtained at approximately the midpoint of the follow-up period.¹⁷³ The present prospective analyses were based on ~7 years' follow-up; data from the first resurvey (approximately 3 years after baseline) were therefore used for the main regression dilution ratio analyses. These included those participants included in analyses examining the association of RPG with incident CVD risks who also had RPG data available from the first resurvey (n= 17 863). Participants who did not report doctor-diagnosed diabetes at baseline but who did report this at resurvey would likely have made lifestyle changes or commenced treatment for diabetes prior to resurvey, which may have reduced their plasma glucose levels beyond the effects of regression dilution. Analyses were therefore repeated excluding these participants (n=314). Analyses were also repeated stratified by baseline and resurvey fasting periods (Appendix Note C.1).

Although second resurvey data were not used for estimating the main regression dilution ratios, usual RPG levels were additionally calculated based on second resurvey data in order to investigate the effect of different follow-up periods. These analyses included 13 642 participants who participated in both first and second resurveys.

Statistical analyses were conducted using SAS version 9.3 (SAS Institute Inc., Cary, North Carolina, USA). Line graphs and bar plots were produced using R version 2.13.1 (R Foundation for Statistical Computing, Vienna, Austria).

4.3 Results

4.3.1 Prevalence of diabetes

At baseline 2.9% (n=6113) of men and 3.3% (n=10 012) of women reported a history of doctor-diagnosed diabetes. After accounting for study area, education and age, women had a higher overall prevalence of self-reported diabetes (men 2.8%, 95% CI 2.7-2.8%; women 3.5%, 3.4-3.6%), but the direction and magnitude of sex differences varied by age, with a higher prevalence in men <50 years and in women ≥50 years (Figure 4.2). The prevalence also varied markedly by study area (Figure 4.3), and was, on average, approximately twice as high in urban as in rural areas in both men and women (men: urban 4.1%, rural 1.8%; women: urban 4.1%, rural 2.5%).

A further 2.6% (n=5537) of men and 2.8% (n=8545) of women were found to have screen-detected diabetes at baseline, representing a total diabetes prevalence of 5.6% in men (n=11 650) and 6.2% in women (n=18 557). Screen-detected diabetes demonstrated similar age- and sex-patterns as self-reported diabetes (Figure 4.2). Again, prevalence was higher in urban than in rural areas, after taking account of age and education, although the difference was less marked than for self-reported diabetes (men: urban 3.6% versus rural 1.9%; women: urban 3.3% versus rural 2.4%) (Figure 4.3). After adjusting for age and education, a greater proportion of diabetes cases were screen-detected in rural than in urban areas in both men (50.2% vs 45.8%) and women (49.2% vs 44.6%). The proportion of diabetes cases newly detected at baseline was also inversely associated with age and education and was higher in men than in women (48.7% vs. 45.4%).

Participants with a history of diabetes in a first-degree relative were four times more likely to have self-reported diabetes, and approximately twice as likely to have screen-detected diabetes, as those without such a family history, although diabetes family history data were missing for ~25 000 participants.

Among the 14 854 participants included in both first and second resurveys, 354 (2.4%) reported a history of doctor-diagnosed diabetes at baseline, and, of these, 86.7% (n=307) again reported a history of diabetes at first resurvey and 92.9% (n=329) at second resurvey (Table 4.3). After excluding participants with new self-reported diabetes at resurvey (since new cases would be expected) kappa statistics for agreement in self-reported diabetes diagnoses at baseline and resurvey were 0.93 (95% CI 0.91-0.95) for the first resurvey and 0.96 (0.95-0.98) for the second resurvey. Of those with self-reported diabetes at baseline, 82.8% (n=293) reported a history of diabetes at both first and second resurveys (kappa 0.90, 0.88-0.93). Overall, 252 (1.7%) participants included in both resurveys had screen-detected diabetes at baseline, of whom 70.2% (n=177) and 75.8% (n=191) had self-reported or screen-detected diabetes at first and second resurveys, respectively. Thus, of the 606 participants with any form of diabetes at baseline, 83.7% (n=507) were classified as having diabetes at first resurvey, 87.3% (n=529) at second resurvey and 91.4% (n=554) at first or second resurvey.

The mean (SD) age at diagnosis of diabetes among participants with self-reported diabetes at baseline was 53.4 (9.3) years (men 52.9 [9.9] years; women 53.7 [8.9] years; $p < 0.001$). Based on these data, the mean (SD) duration of diagnosed diabetes at baseline was 5.7 (5.3) years (men: 5.5 [5.4] years; women: 5.8 [5.3] years). Of those with self-reported diabetes, 76.9% reported use of diabetes medications, with the majority (61.1% of men and 63.2% of women) taking oral hypoglycaemic agents alone, just over 10% taking insulin alone (13.1% of men and 10.9% of women) and <3% taking both insulin and oral hypoglycaemic agents (2.6% of men and 2.8% of women) (Table 4.4). The mean (SD) age-, area- and education-adjusted RPG levels in men and women with self-reported diabetes at baseline were 11.5 (5.8) mmol/L and 11.2 (5.7) mmol/L, respectively, with no clinically-significant difference between those who did and did not report treatment with insulin or oral hypoglycaemic agents (11.6 [5.5] mmol/L vs 11.3 [5.6] mmol/L, $p = 0.01$). Using data on age at diagnosis of diabetes and insulin use to estimate the relative proportions of type 1

and type 2 diabetes, it was estimated that <1% of self-reported diabetes cases at baseline were type 1.

4.3.2 Cross-sectional associations of diabetes

Tables 4.5 and 4.6 and Figures 4.4 to 4.6 show the cross-sectional associations of self-reported and screen-detected diabetes with other baseline variables. There was a modest positive association between education and self-reported diabetes prevalence after adjusting for age, sex and study area, ranging from 3.4% (95%CI 2.9-3.9%) among participants with no formal schooling, to 3.9% (3.1-4.7%) among those with college or university education (Table 4.5). In contrast, the prevalence of screen-detected diabetes decreased with increasing education, with prevalences of 3.9% (3.4-4.4%) and 2.3% (1.8-2.6%) in the same two groups. These differing associations could be explained, in part, by the inverse association between education and the proportion of diabetes cases that were screen-detected. Total diabetes prevalence was higher in individuals with no formal schooling (7.3, 6.7-7.9%) than in other education categories. After further adjustment for all other potential confounding factors considered (eg, adiposity, SBP, physical activity, smoking, alcohol consumption and diet), no clear associations of education with diabetes prevalence remained.

Ex-regular smokers had the highest prevalence of self-reported and screen-detected diabetes compared with other smoking categories after adjustment for age, sex, study area, education and alcohol (Table 4.5). Self-reported diabetes prevalence was also highest among ex-regular drinkers, but no clear association with alcohol consumption was seen for screen-detected diabetes. There was a strong inverse association between physical activity and self-reported diabetes prevalence ($p < 0.001$), and a similar, but weaker, association with screen-detected diabetes ($p < 0.001$). Both associations persisted after adjustment for other potential confounding factors, including adiposity. Individuals who reported eating fresh fruit rarely or never had higher self-reported diabetes prevalence than those who more

regularly consumed fresh fruit (Table 4.6). The prevalence of self-reported diabetes was positively associated with soybean and dairy product consumption (both $p < 0.001$), which persisted after full adjustment. These associations were not apparent with screen-detected diabetes. There was a weak positive association between frequency of tea consumption and screen-detected diabetes prevalence ($p < 0.001$), but this was attenuated after further adjustment.

Both self-reported and screen-detected diabetes prevalence were strongly positively associated with measures of adiposity (Figures 4.4 and 4.5). For both BMI and percentage body fat, the positive associations with self-reported diabetes prevalence when adjusted for age, sex, study area and education, were moderately attenuated by further adjustment for other adiposity measures (p for trend < 0.001). By contrast, for waist circumference, the positive association with diabetes prevalence became stronger after additionally adjusting for BMI and hip circumference (p for trend < 0.001). Hip circumference was positively associated with diabetes prevalence when adjusted only for age, sex, study area and education, but the association was reversed after additional adjustment for waist circumference and BMI (p for trend < 0.001). When combined as waist-to-hip ratio, there was a positive association with diabetes prevalence after minimal adjustment, which remained after additional adjustment for other adiposity measures (p for trend < 0.001) and for all potential confounders considered. Diabetes prevalence was positively associated with SBP (p for trend < 0.001), which remained after adjustment for other potential confounding factors including adiposity (Figure 4.6). There was a weaker, but qualitatively similar, association with DBP (p for trend < 0.001).

Analyses examining the cross-sectional associations of diabetes among those without prior CVD ($n = 488\,760$) demonstrated similar results and are therefore not presented.

4.3.3 Plasma glucose distribution

RPG levels were approximately normally distributed, although slightly right-skewed (Figure 4.7) among the 487 552 participants without self-reported diabetes and with RPG data who were included in the cross-sectional analyses. Mean (SD) RPG levels were 5.8 (1.9) mmol/L and 6.0 (1.9) mmol/L in men and women, respectively.

Most participants (78.9%) reported having eaten within the previous eight hours. At baseline, participants in Zhejiang were asked to fast prior to attending the assessment centre and, consequently, 70.7% of participants in Zhejiang reported a fasting period ≥ 8 hours, compared with between 1.9% and 35.1% of participants in the remaining nine study areas. At fasting periods < 6 hours there was an inverse association between fasting period and mean age-, sex-, study area- and education-adjusted plasma glucose, which flattened off at longer fasting periods (Figure 4.8). Based on the observed association of plasma glucose levels with fasting period, and on guidance recommending at least 8 hours fasting prior to measurement of fasting plasma glucose levels,¹⁸³ an 8 hour cut-point was selected for comparison of fasting and non-fasting plasma glucose levels. Compared with individuals with a ≥ 8 hour fasting period, those with fasting periods < 7 hours had significantly higher plasma glucose levels ($p < 0.001$) (Table 4.7), although some absolute differences were minimal. Despite this association, the proportion of variation in plasma glucose levels explained by fasting period was small (< 8 hours fasting time $r^2 = 0.01$; ≥ 8 hours fasting time $r^2 = 0.001$).

Study participants with a baseline RPG level of 7.8-11.0 mmol/L were invited back the following morning to undergo FPG testing. 25 428 participants without self-reported diabetes met these criteria and 12 236 (48.1%) underwent FPG testing. Among these 12 236 participants mean (SD) FPG was 5.8 (1.1) mmol/L in men and 6.0 (1.1) mmol/L in women. Although most participants had a FPG level within the “normoglycaemic” range (≤ 6.0 mmol/L), over 10% (men: 11.8%; women: 14.7%) had a FPG level consistent with a

diagnosis of diabetes (≥ 7.0 mmol/L) and a further quarter (men: 23.9%; women: 28.3%) had a level consistent with a diagnosis of IFG (6.1-6.9 mmol/L), based on WHO criteria (Table 4.8).⁷⁸

There was a positive association of mean study area- and education-adjusted RPG levels with age (p for trend $p < 0.001$), with higher mean RPG in women, than in men, at all ages ($p < 0.001$) (Figure 4.9). There was modest variation in RPG levels across the ten study areas (Table 4.9) and overall RPG was modestly higher in urban than in rural areas ($p < 0.001$). This difference persisted after excluding Zhejiang (where participants were asked to fast) (6.0 mmol/L vs. 5.8 mmol/L, $p < 0.001$).

4.3.4 Cross-sectional associations of plasma glucose

RPG levels in individuals without self-reported diabetes at baseline varied little by socioeconomic or lifestyle factors (Tables 4.10 and 4.11). As for diabetes, there was a positive association of mean age-, sex-, study area- and education-adjusted RPG with various measures of adiposity. For BMI, the strength of the association was attenuated by additional adjustment for waist and hip circumference (p for trend < 0.001) (Figure 4.10). The association of RPG with waist circumference was modestly stronger than with BMI, and was not much attenuated after additional adjustment for BMI and hip circumference (p for trend < 0.001). Hip circumference was weakly positively associated with RPG after minimal adjustment, but additional adjustment for BMI and waist circumference resulted in an inverse association (p for trend < 0.001). Waist-to-hip ratio was positively associated with RPG after minimal adjustment, and further adjustment for BMI had little effect on the association (p for trend < 0.001). There was a weak positive association of percentage body fat with age-, sex-, study area- and education-adjusted RPG, which was reversed with additional adjustment for BMI and waist and hip circumference (p for trend $p < 0.001$). Age-, sex-, study area- and education-adjusted RPG increased modestly with increasing SBP

(p for trend <0.001) and showed a very modest positive association with DBP (p for trend <0.001) (Figure 4.11).

Repetition of these analyses among those without prior CVDs (ie, the study population for prospective analyses of RPG) showed very similar results which are therefore not presented separately.

4.3.5 Within-person variation in plasma glucose

After dividing participants into five baseline RPG groups, mean baseline RPG ranged from 4.9 mmol/L in the lowest category, to 15.8 mmol/L in the highest. At first resurvey, on average 2.6 years later, mean RPG for the same baseline categories were 5.9 mmol/L and 11.9 mmol/L, respectively (Figure 4.12). The range of mean RPG between top and bottom categories decreased from 10.9 mmol/L at baseline to 6.1 mmol/L at first resurvey, with an overall regression dilution ratio of 0.56. Table 4.12 shows age- and sex-specific regression dilution ratios. Overall, there was a lower regression dilution ratio in men than in women (0.51 vs. 0.58). There was no consistent trend with age, but the regression dilution ratio was lower in men aged 30-59 years (0.47) than in other age- and sex-defined subgroups.

Additional analyses were undertaken using second resurvey data to assess the impact of the time interval between initial and repeat measurements on intra-individual variation in RPG. Based on the 13 642 participants included in the main regression dilution bias analyses and in the second resurvey, mean second resurvey RPG levels in the five baseline RPG categories ranged from 5.5 mmol/L in the lowest category to 10.1 mmol/L in the highest, with a range of 4.6 mmol/L (Figure 4.13). Thus, after a mean 8.1 year interval between baseline and second resurvey, the regression dilution ratio was 0.42. The regression dilution ratio derived from first and second resurvey data, representing a mean 5.5 year interval, was 0.77.

Although the linearity assumption underlying Rosner's regression method for calculating regression dilution ratios (Figure 4.14) and the equal variances assumption underlying the

self-correlation approach (3.4 mmol/L and 4.2 mmol/L at baseline and resurvey, respectively) were not met, these methods were investigated for comparison. The age- and sex-specific regression dilution ratios calculated using Rosner's method were generally modestly lower than those calculated using MacMahon's approach (overall regression dilution ratio 0.50 versus 0.56, Table 4.12). Regression dilution ratios calculated using the self-correlation method were consistently, and markedly, lower (overall regression dilution ratio 0.26) than those calculated using both alternative approaches. The self-correlation approach is particularly susceptible to a non-normal distribution, but log-transformation of RPG did not substantially alter these findings. A key advantage of both Rosner's approach¹⁹⁰ and the self-correlation approach is the ability to calculate regression dilution ratios adjusted for factors potentially confounding the exposure-disease association under investigation; age- and sex-adjusted regression dilution ratios closely resembled the unadjusted ratios.

Regression dilution ratios calculated using MacMahon's method were moderately higher in analyses additionally excluding individuals with self-reported diabetes at resurvey (Table C.1), but exclusion of these individuals made little difference to regression dilution ratios calculated using Rosner's or self-correlation approaches.

4.4 Discussion

These large-scale cross-sectional analyses of data from approximately 0.5 million middle-aged Chinese men and women in the CKB provide extensive and detailed quantification of the distribution and correlates of diabetes and RPG, and of intra-individual variation in RPG levels.

4.4.1 Prevalence and correlates of diabetes

Although the CKB was not designed to be nationally representative, the prevalence of diabetes, based on self-report and on plasma glucose measurement, was broadly similar to some other contemporaneous Chinese studies. In the InterASIA Study, a nationally

representative survey of >15 000 Chinese adults aged 35-74 years undertaken in 2000-2001, the overall diabetes prevalence, based on self-report or FPG measurement, was 5.5%.¹⁸⁰ The China National Nutrition and Health Survey, conducted one year later and including >45 000 adults aged ≥ 20 years found a lower diabetes prevalence, ascertained in the same manner, of 2.7%.¹⁸¹ Two more recent nationally representative surveys reported higher diabetes prevalence. The CNDMDS, conducted in 2007-2008 and including >45 000 men and women aged ≥ 20 years from 14 provinces and municipalities, found an overall diabetes prevalence of 9.7%.¹⁷⁹ In the most recent 2010 survey, which included almost 100 000 men and women aged ≥ 18 years, the reported diabetes prevalence was 11.6%.⁸⁷ In these latter two studies, the prevalence of self-reported diabetes was similar (4.1%¹⁷⁹ and 3.6%⁸⁷ in men, and 3.5%¹⁷⁹ and 3.4%⁸⁷ in women), and modestly higher than the 2.9% and 3.3% prevalence in men and women, respectively, in the CKB. It has been estimated that worldwide, about half of diabetes cases remain undiagnosed;⁶⁵ this is very similar to the proportion (47%) of diabetes cases that were screen-detected at baseline in the CKB, but much lower than that in the two more recent surveys in China. In the CNDMDS screen-detected diabetes prevalence was 6.5% in men and 5.2% in women, accounting for 60% and 58% of male and female diabetes cases, respectively.¹⁷⁹ In the 2010 survey, screen-detected diabetes prevalence was 8.5% in men and 7.7% in women, accounting for 70% of both male and female diabetes cases.⁸⁷ Multiple factors likely contribute to differences between studies, including secular trends reflecting the growing diabetes epidemic in China,⁸⁹ and differing study populations, particularly in terms of age, and sampling schemes. Furthermore, a range of glycaemic measures were used. The 8.1% screen-detected diabetes prevalence in the 2010 survey was based on OGTT and measurement of FPG and HbA1c,⁷⁸ which at least partly explains the higher screen-detected diabetes prevalence in this study.

As in previous studies,^{87, 179-181} diabetes prevalence was higher in urban than rural areas in the CKB, which has been attributed to lifestyle factors, including higher levels of physical

inactivity, poor diet and adiposity in urban areas.⁸⁹ Increasing diabetes prevalence with increasing age is universally acknowledged,^{65, 87, 179-181} reflecting increasing incidence with age, due to increased insulin resistance and reduced pancreatic beta-cell responsiveness,¹⁹¹ and the chronicity of the condition. We found higher diabetes prevalence in men than in women at ages <50 years, but higher prevalence in women than in men ≥50 years. This has been seen in some,^{87, 192} but not all,¹⁹³ previous studies in China. It has been postulated that greater exposure of men, than women, to smoking, high fat diets and occupational stress may account for higher prevalence in younger men, but that the menopause negates this female advantage at older ages.¹⁹² The greater proportion of previously undiagnosed diabetes cases in rural, compared with urban, areas and in younger age groups is consistent with previous reports,^{87, 179} and may be accounted for by differences in education and access to health care between urban and rural areas and differing levels of awareness.

A history of diabetes in first-degree relatives was strongly associated with diabetes in the CKB, likely reflecting the influence of both genetic and shared environmental factors within families. Reverse causality and behaviour changes resulting from diagnosis of diabetes probably explain several observed associations, including the higher prevalence in ex-regular smokers—which is consistent with the lower diabetes prevalence described in current smokers in another large cross-sectional survey⁸⁷—and consumers of alcohol. For screen-detected diabetes there was less bias due to reverse causation, and thus less markedly elevated prevalence in ex-regular smokers and drinkers. The substantially higher prevalence of diabetes in participants who reported rarely or never consuming fresh fruit may reflect reverse causation, ie, reduced consumption in individuals with known diabetes due to concerns regarding high sugar content. However, despite the cross-sectional nature of the analyses, many demonstrated associations between diabetes prevalence and other baseline variables are consistent with previously reported risk factors for, or correlates of, diabetes, and with current understanding of the pathophysiology of diabetes. These include

the positive associations with adiposity measures and the inverse association with physical activity,^{180, 194, 195} postulated to reflect the relationship with adiposity and an independent insulin-sensitising effect of physical activity.⁷⁵

Adiposity is the strongest known modifiable risk factor for type 2 diabetes¹⁹⁶ mediated predominantly through effects on insulin resistance.¹⁹⁷ The strong association of central adiposity (eg, waist circumference) with diabetes prevalence in the CKB, and its persistence after additional adjustment for BMI (a general measure of adiposity), is consistent with the proposed causal role of visceral adiposity in type 2 diabetes.¹⁹⁷⁻²⁰⁰ Fatty acids and cytokines produced by lipolysis in visceral adipose tissue induce insulin resistance,¹⁹⁶ and this effect is enhanced by their transport, via the portal circulation, from visceral adipose tissue to the liver, an important site of action of insulin.^{197, 201} Asian populations, including Chinese, are more susceptible to visceral adiposity than, for example, Caucasian populations,²⁰² and are thus thought to be at greater risk of diabetes at a given BMI.²⁰³ This is a particularly pertinent point given increasing levels of adiposity in Chinese populations over recent decades.⁸⁹ The apparent inverse association between hip circumference and diabetes prevalence, after accounting for central and overall adiposity, likely reflects a relative protective effect of gluteofemoral fat against diabetes. Gluteofemoral fat is thought to be associated with lower free fatty acid production,¹⁹⁷ greater fatty acid storage²⁰⁴ and production of a more favourable adipokine profile, resulting in relatively lower insulin resistance.²⁰⁴ The positive association of diabetes with both SBP and DBP may reflect a direct association²⁰⁵ or shared risk factors, particularly adiposity.^{199, 205, 206}

4.4.2 Validity of diabetes diagnoses

Self-reporting of diabetes, which is widely used in epidemiological studies, is prone to misclassification. However, evidence from various studies suggests it is a valid indicator of established diabetes diagnoses, with specificity and sensitivity estimates ranging from 97% to 100% and 66% to 84%, respectively.²⁰⁷⁻²¹¹ Furthermore, the high level of agreement

between baseline and resurvey reporting of diabetes diagnoses in the CKB supports the validity of this approach.

ADA⁶⁴ and WHO^{78, 107} guidelines recommend the use of FPG, OGTT or HbA1c for diabetes diagnosis in a clinical setting. Diagnosis based on RPG measurement is recommended only in the presence of classic symptoms of hyperglycaemia.⁶⁴ Priorities and objectives differ, however, in the context of large-scale epidemiological studies, such as the CKB, when certainly OGTT, but also FPG testing,²¹² are less feasible. Similarly, it was not possible to measure HbA1c at the time of recruitment in the CKB due to cost and logistic constraints, although HbA1c may be measured later on the stored buffy coat.²¹³ RPG was therefore selected as the main glycaemic measure. Findings of several studies investigating the use of RPG as a screening test for diabetes²¹⁴⁻²¹⁷ suggest the screen-detected diabetes definition used in the CKB would have relatively low sensitivity, but high specificity. That is, there would be relatively few false positive, but more false negative, screen-detected diabetes cases in the CKB. Importantly, however, use of fasting period data should have reduced misclassification, and the relatively high proportion of individuals with screen-detected diabetes at baseline who had self-reported or screen-detected diabetes at resurvey further supports the validity of the diagnosis.

4.4.3 Distribution and correlates of random plasma glucose

The distribution of RPG in individuals without diabetes has not been well described in Chinese, or other, populations, and most previous studies have examined FPG.^{79, 218, 219} The slightly right-skewed RPG distribution in the CKB is consistent with previous reports.²¹⁸ Although log transformation of RPG produced a more normal distribution, this was not used routinely in the presented analyses due to the minimal effects of transformation on the distribution, potential difficulties in interpreting the results, lower clinical relevance and applicability, and greater complexity of comparisons with the majority of previous studies, which have investigated non-transformed glucose levels.

Correlates of RPG in individuals without diabetes are not well described in Chinese populations. Studies in other populations have shown positive associations of glucose levels (variably fasting, post-load and random glucose) with age^{82, 83, 218, 220-222} and higher levels in women than in men,^{220, 221} although this sex difference is inconsistent between studies.^{219, 222} The positive associations of RPG levels with measures of adiposity and with SBP were qualitatively similar to associations of diabetes prevalence, likely reflecting similar underlying pathophysiologies, and consistent with findings from other study populations.^{82-84, 103, 223}

4.4.4 Intra-individual variation in random plasma glucose levels

The availability of repeat exposure measurements on a random sample of the CKB study population provides a unique opportunity to investigate the nature and extent of within-person variation in RPG. Based on data for ~18 000 individuals without self-reported diabetes at baseline, we estimated an overall RPG regression dilution ratio of 0.56, not inconsistent with the ratio of 0.46 (95% CI 0.33-0.59) from the ERFC based on repeat RPG measures, on average, 4 years after baseline.²²⁴ These suggest considerable variation in RPG levels within individuals, likely resulting from technical measurement error, short-term within-person variation, for example, due to variable fasting times and composition of recent dietary intake, and longer-term variation, such as that resulting from ageing.^{173, 225} For FBG, regression dilution ratios were 0.63¹⁰³ and 0.69⁶ in the APCSC and ERFC, respectively. The lower RPG regression dilution ratios suggest greater intra-individual variation in RPG than FPG levels. The lower regression dilution ratio after a longer interval between repeat RPG measurements is in keeping with patterns seen for other exposures,¹⁷³ and demonstrates the increasing extent and importance of intra-individual variation in RPG over time.¹⁷³

The assumptions underlying Rosner's¹⁸⁶ and self-correlation¹⁸⁷ methods for regression dilution ratio calculation were not met in the CKB. MacMahon's non-parametric approach¹⁷⁴

was therefore used. This consistently produced more conservative RPG regression dilution ratio estimates than the parametric approaches, and it imposes some limitations. It is not possible, using MacMahon's method, to calculate regression dilution ratios adjusted for other covariates, and it has been observed for other exposures that unadjusted regression dilution ratios are generally less extreme than adjusted ratios.²²⁵ Whilst acknowledging potential limitations of the parametric approaches in this specific context, age- and sex-adjusted regression dilution ratios were only modestly lower than unadjusted ratios in the CKB. As described, variable fasting times prior to plasma glucose measurement likely contributed to intra-individual variation in RPG levels. Accounting for this when calculating regression dilution ratios may have improved the accuracy of adjustment for regression dilution bias, but the effect would be expected to be minimal given the small proportion of variation in RPG levels explained by fasting time. Furthermore, there is no agreed method for calculating confidence limits for ratios calculated using MacMahon's approach. A further limitation of MacMahon's method is that it uses only a subset of the data available,¹⁷⁴ specifically data from individuals in the top and bottom baseline RPG categories. This likely explains why excluding individuals with self-reported diabetes at resurvey had the greatest impact on regression dilution ratios calculated using MacMahon's method.

4.4.5 Random plasma glucose as a glycaemic measure

RPG was the main glycaemic measure in the CKB for practical and logistical reasons. One of the key problems with using RPG as a marker of an individual's glycaemic status is the influence of recent dietary intake, even in individuals without diabetes.^{215, 217, 226} Consistent with previous literature,²²⁶ there was a negative association of plasma glucose levels with fasting period in the CKB, but, importantly, fasting period explained only a small proportion of variation in RPG. Although the nature of recently consumed food was not accounted for, it is likely that other factors, including age and adiposity had a greater impact on RPG level variation between individuals. Furthermore, in the context of investigating the association of plasma glucose with long-term risk of CVD, it has been suggested that, as we spend

most of our time in a non-fasting state, it is equally, or more, appropriate to study the relevance of non-fasting, than fasting, plasma glucose.¹⁴⁸ Thus, despite concerns,²²⁷ in the context of large-scale population-based epidemiological studies, such as the CKB, RPG would appear to be a sufficiently reliable glycaemic indicator.²²⁶

There is little, if any, published evidence on the appropriateness of glucometers for measurement of plasma glucose levels in large-scale, epidemiological studies. Previous studies comparing the use of glucometers with laboratory-based blood glucose measurement techniques typically compare capillary plasma glucose levels measured using a glucometer with venous plasma glucose levels measured using laboratory techniques,²²⁸ or are based on specific population groups,^{229, 230} which cannot be directly extrapolated to the current setting. Nevertheless, there are some well-recognised limitations of glucometers. Glucose levels measured on a glucometer may be less accurate than laboratory-based measurements,²²⁸ particularly at extremes of the glycaemic range,^{231, 232} and are susceptible to operator error and errors due to environmental factors (eg, temperature and humidity), to participant factors (eg, contamination or biological variation, including severe hypertriglyceridaemia or hyperuricaemia), and, rarely, to pharmacological factors.²³¹⁻²³³ Furthermore, although the SureStep Plus meter used a glucose oxidase-based technique, with good specificity for glucose,^{185, 232} the On Call Advanced meter, used in the second resurvey, uses a non-specific pyroloquinoline-quinone-dependent glucose dehydrogenase technique²³⁴ which may provide falsely elevated glucose levels due to the presence of other substances, such as maltose, xylose and galactose, in participants' blood,²³² although the effects of this at a population level would be expected to be small. As described, both glucometers used in the study were unable to provide readings at very high plasma glucose concentrations,^{185, 234} but sensitivity analyses excluding individuals with a recorded blood glucose level >27.8 mmol/L at baseline did not alter findings.

Whilst acknowledging these limitations, the decision to use a point-of-care glucometer in the CKB took account of various practical considerations; use of a glucometer is less

resource intensive, avoids falsely low glucose levels due to ex-vivo glycolysis²³¹ and provides an immediate result, enabling feedback to study participants. A range of measures were taken to minimise error in plasma glucose measurement, including staff training, use of standard protocols^{169, 172} and regular calibration of glucometers using manufacturer control solutions. A quality control process was undertaken to examine the precision of the On Call Advanced glucometer (although not the SureStep Plus meter), which, using two levels of the CTSU Wolfson Laboratories routine quality control material, demonstrated reasonable coefficients of variation (3.25% and 2.82%).²³⁵ The accuracy of the On Call Advanced glucometer could not be assessed, since the laboratories did not support an accredited test for blood glucose levels at this time.

4.4.6 Conclusion

In this large Chinese adult population 3.2% reported a history of doctor-diagnosed diabetes at baseline, and a further 2.8% had screen-detected diabetes. There were strong positive associations of both diabetes and RPG with age, adiposity and blood pressure. Diabetes prevalence was higher in ex-regular smokers and in the less physically active, and self-reported diabetes prevalence was also higher in ex-regular alcohol drinkers, but RPG varied little by socioeconomic and lifestyle characteristics. The cross-sectional nature of the presented analyses and potential biases from reverse causality prevent any inferences regarding the determinants of diabetes or RPG levels. The analyses do, however, provide valuable information on the inter-relationships of both exposures with other relevant baseline variables in the CKB population, enabling informed and robust adjustment for confounding in subsequent analyses of their associations with CVDs. Furthermore, understanding of variation in RPG levels within individuals in the CKB provides essential information for avoidance of regression dilution bias in analyses of the associations of RPG with CVDs.

Table 4.1. Frequency of missing baseline random plasma glucose levels and random plasma glucose levels >27.8 mmol/L by study area in individuals without self-reported diabetes at baseline

Study area	Population	Date of commencement of recruitment	Random plasma glucose (%)			
			Missing		>27.8 mmol/L	
Urban						
Qingdao	33602	July 2004	438	(1.3)	20	(0.1)
Harbin	54013	September 2004	64	(0.1)	32	(0.1)
Haikou	28529	April 2005	7	(0.0)	8	(0.0)
Suzhou	51838	November 2004	5	(0.0)	48	(0.1)
Liuzhou	47670	October 2004	0	(0.0)	36	(0.1)
Total	215652	July 2004	514	(0.2)	144	(0.1)
Rural						
Sichuan	54755	June 2004	3669	(6.7)	34	(0.1)
Gansu	49421	October 2004	5	(0.0)	20	(0.0)
Henan	61867	October 2004	1	(0.0)	26	(0.0)
Zhejiang	56257	August 2004	3792	(6.7)	3	(0.0)
Hunan	58777	September 2004	179	(0.3)	4	(0.0)
Total	281077	June 2004	7646	(2.7)	87	(0.0)
Total	496729	June 2004	8160	(1.6)	231	(0.0)

Table 4.2. Baseline characteristics of participants by availability of baseline random plasma glucose data in individuals without self-reported diabetes at baseline

Characteristic	Random plasma glucose data ^a		p value
	Available (n= 488 569)	Missing (n=8160)	
Men^b, %	41.0	36.4	0.08
Age^c (years), mean	51.6	63.6	<0.001
Living in urban area, %	44.1	35.6	<0.001
≥6 years of education, %	49.2	46.6	0.001
Current smoker, %	26.4	23.6	0.3
Regular alcohol consumption, %	14.8	17.2	<0.001
Physical activity (MET hours/day), mean	25.6	26.2	0.0002
SBP (mmHg), mean	131.2	127.3	<0.001
BMI (kg/m²), mean	23.7	23.5	<0.001
Self-reported prevalent disease, %			
Ischaemic heart disease	3.0	4.4	<0.001
Stroke/transient ischaemic attack	1.7	0.5	0.8
Hypertension	11.6	12.0	<0.001

^aAdjusted for age, sex and study area; ^bAdjusted for age and study area; ^cAdjusted for study area and sex.

BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

Figure 4.1. Participant exclusions for presented analyses

*Some participants had missing or outlying data for >1 baseline variable.

BMI, body mass index; CKB, China Kadoorie Biobank; DBP, diastolic blood pressure; IHD, ischaemic heart disease; RPG, random plasma glucose; SBP, systolic blood pressure; TIA, transient ischaemic attack.

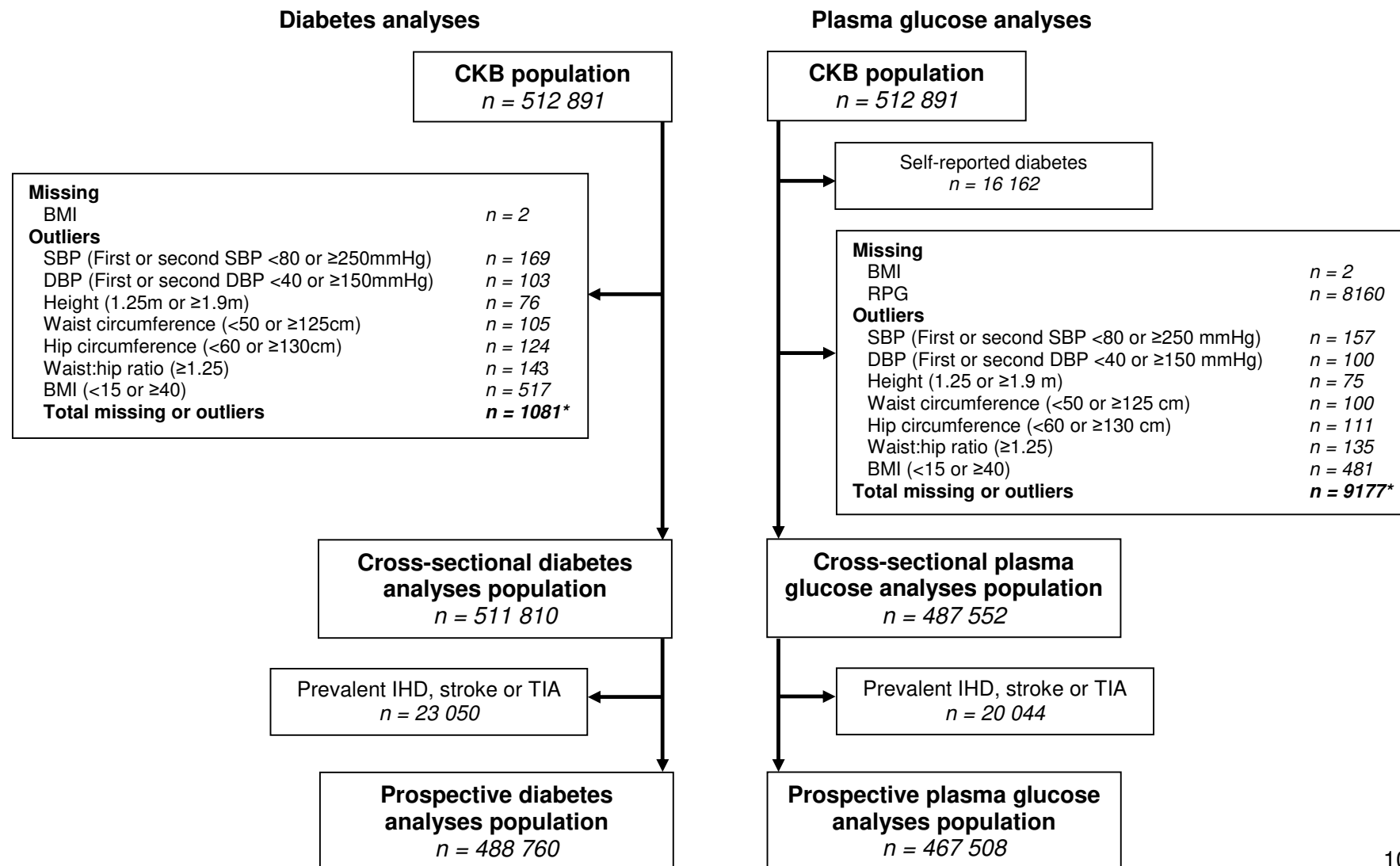


Figure 4.2. Baseline prevalence of self-reported and screen-detected diabetes by age

Adjusted for study area and education.

Squares represent diabetes prevalence and vertical lines represent corresponding 95% confidence intervals.

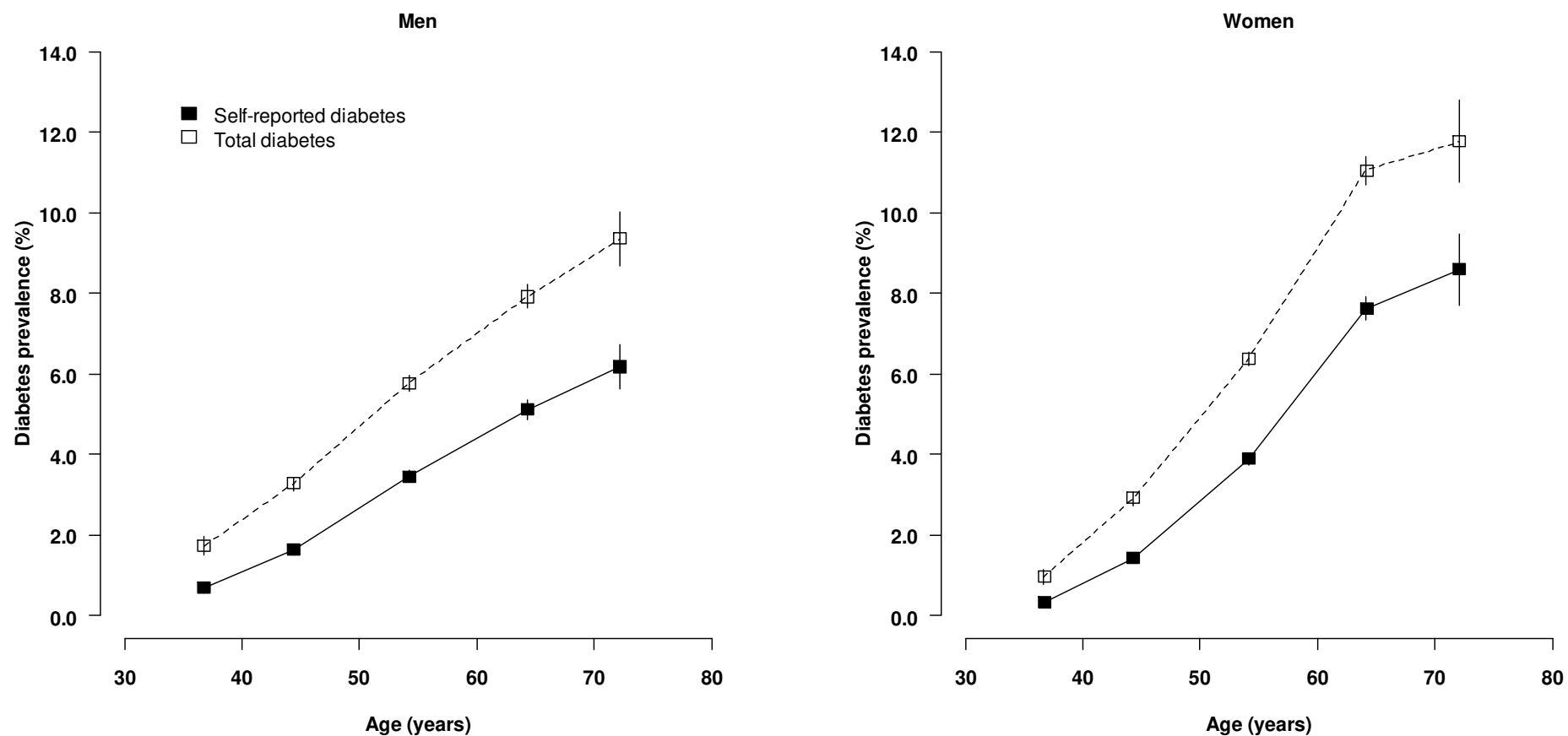


Figure 4.3. Baseline prevalence of self-reported and screen-detected diabetes by study area

Adjusted for age and education.

Error bars represent 95% confidence intervals. Percentages above bars represent the percentage of total diabetes cases that were screen-detected.

R, rural; U, urban.

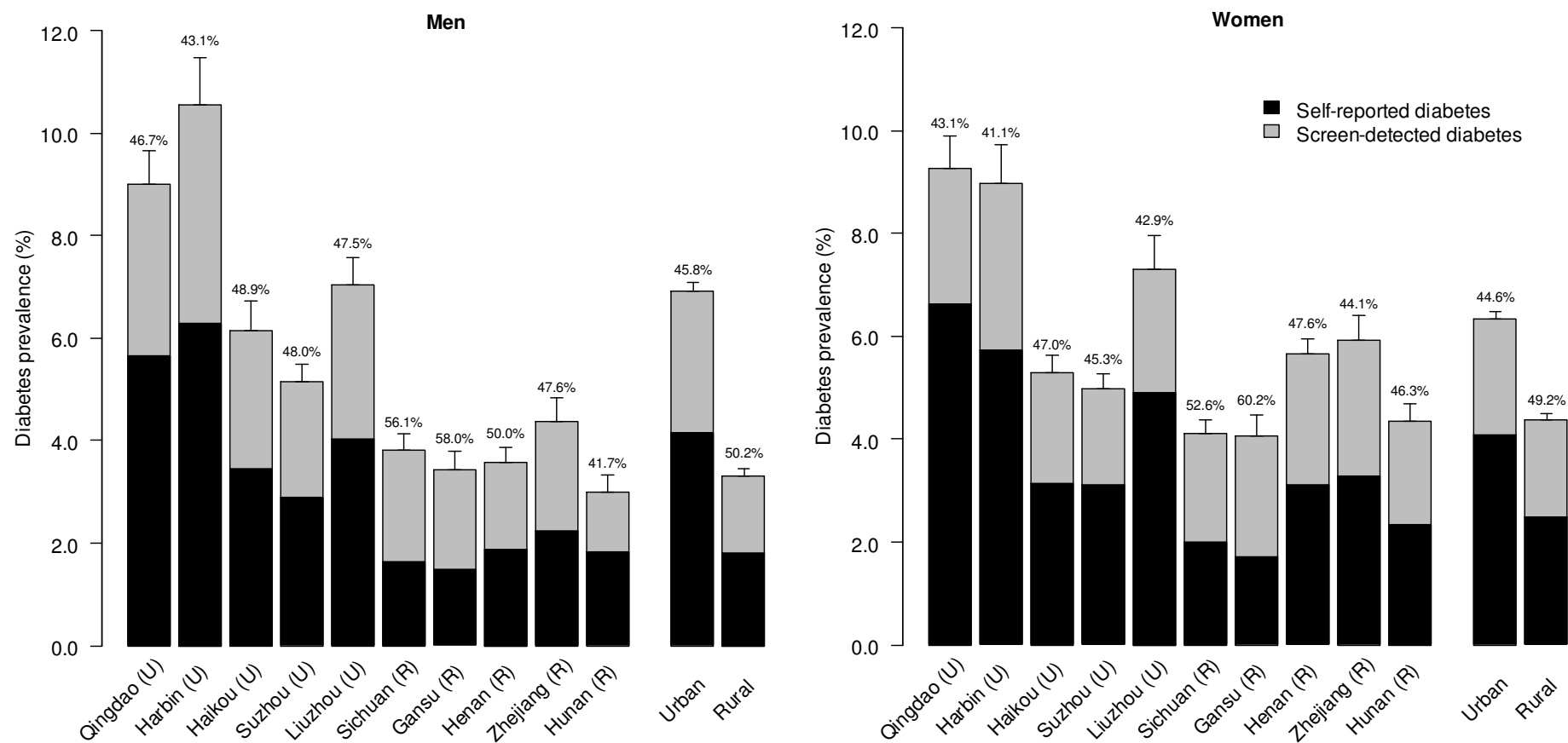


Table 4.3. Diabetes status at baseline and resurvey among participants included in first and second resurveys (n=14 854)

Baseline diabetes status	First resurvey diabetes status								Second resurvey diabetes status							
	None		Screen-detected		Self-reported		Total		None		Screen-detected		Self-reported		Total	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
None (n=14 248)	13765	96.6	353	2.5	130	0.9	483	3.4	13361	93.8	322	2.3	565	4.0	887	6.2
Screen-detected (n=252)	75	29.8	77	30.6	100	39.7	177	70.2	61	24.2	21	8.3	170	67.5	191	75.8
Self-reported (n=354)	24	6.8	23	6.5	307	86.7	330	93.2	16	4.5	9	2.5	329	92.9	338	95.5
Total (n=606)	99	16.3	100	16.5	407	67.2	507	83.7	77	12.7	30	5.0	499	82.3	529	87.3

Kappa for self-reported diabetes based on baseline and resurvey excluding participants with self-reported diabetes at resurvey who did not have self-reported diabetes at baseline: first resurvey 0.93 (95% CI 0.91-0.95); second resurvey 0.96 (95% CI 0.95-0.98); both first and second resurveys: 0.90 (95% CI 0.88-0.93).

Table 4.4. Self-reported diabetes medication use in men and women with self-reported diabetes at baseline

Medication	Men^a, % (n=5052)	Women^b, % (n=8600)	p value
Oral hypoglycaemic agents only	61.1	63.2	0.01
Insulin only	13.1	10.9	<0.001
Insulin and oral hypoglycaemic agents	2.6	2.8	0.5
Insulin or oral hypoglycaemic agents	76.9	76.9	1.0
Not on treatment	23.1	23.1	1.0

^aData missing for 1061 participants; ^bData missing for 1412 participants.

Table 4.5. Baseline diabetes prevalence by lifestyle characteristics

Lifestyle characteristic	No. of participants	Self-reported diabetes prevalence (%)			Screen-detected diabetes prevalence (%)		
		Minimal adjustment ^a	Intermediate adjustment	Full adjustment ^b	Minimal adjustment ^a	Intermediate adjustment	Full adjustment ^b
Education^c							
No formal schooling	94871	3.4	-	3.3	3.9	-	3.3
Primary school	164875	3.3	-	3.0	3.1	-	2.8
Middle school	144705	3.6	-	2.8	2.7	-	2.4
High school	77402	3.8	-	3.0	2.6	-	2.5
College or university	29957	3.9	-	4.2	2.3	-	2.5
Smoking^d							
Never	316870	3.3	3.2	3.3	2.8	2.7	2.8
Occasional	29121	3.5	3.7	2.8	2.9	3.1	2.5
Ex-regular	30499	5.5	5.3	5.7	4.1	4.3	3.6
Current regular	135320	3.2	3.3	2.3	3.4	3.4	2.5
Alcohol^e							
Never regular	234498	4.0	4.0	3.8	2.9	2.9	2.9
Occasional	162976	2.5	2.5	2.4	2.5	2.5	2.4
Ex-regular	9236	7.6	8.0	6.8	3.0	3.1	3.3
Reduced intake	11678	4.4	4.0	5.9	3.0	3.0	3.1
Monthly	17369	2.2	2.3	1.8	2.7	2.8	2.1
Weekly	76053	1.4	1.4	1.9	2.6	2.6	2.9
Physical activity (MET hours/day)							
<10	38603	5.1	-	6.4	3.7	-	4.2
10-19.9	179340	3.8	-	4.6	2.9	-	3.4
20-29.9	127420	2.8	-	2.4	2.6	-	2.4
30-39.9	81988	2.0	-	1.3	2.6	-	2.0
≥40	84459	1.6	-	1.0	2.2	-	1.8

^aAdjusted for age, sex, education and study area; ^bAdditionally adjusted for alcohol, smoking, systolic blood pressure, physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, fresh vegetable consumption, meat consumption, soybean product consumption, fish consumption, dairy product consumption, rice consumption, wheat consumption, tea consumption; ^cAdjusted for variables as above excluding education; ^dIntermediate adjustment: adjusted for age, sex, education, study area and alcohol consumption; ^eIntermediate adjustment: adjusted for age, sex, education, study area and smoking.
MET, metabolic equivalent of task.

Table 4.6. Baseline diabetes prevalence by dietary consumption

Dietary consumption	No. of participants	Self-reported diabetes prevalence (%)		Screen-detected diabetes prevalence (%)	
		Minimal adjustment ^a	Full adjustment ^b	Minimal adjustment ^a	Full adjustment ^b
Fresh fruit					
Rarely/never	32661	11.8	8.7	2.9	2.9
Monthly	173775	4.5	2.6	2.8	2.5
1-3 days/week	161052	2.9	2.6	2.8	2.7
4-6 days/week	47902	2.3	2.2	2.9	2.8
Daily	96420	2.8	3.4	2.8	3.2
Fresh vegetables					
Rarely/never	127	3.3	3.0	6.9	5.2
Monthly	1395	0.6	0.8	1.6	1.6
1-3 days/week	7169	2.7	2.2	2.9	2.2
4-6 days/week	18015	2.5	2.0	3.3	2.8
Daily	485104	3.2	3.2	2.8	2.8
Soybean products					
Rarely/never	56001	3.4	2.3	2.4	2.4
Monthly	147836	2.9	2.3	2.7	2.5
1-3 days/week	257504	3.2	3.6	2.8	2.8
4-6 days/week	31013	4.0	3.9	2.7	2.7
Daily	19456	5.1	6.6	2.8	3.4
Dairy products					
Rarely/never	350001	2.6	2.2	2.8	2.5
Monthly	56967	3.5	3.3	3.0	2.9
1-3 days/week	43905	4.3	4.2	3.0	3.3
4-6 days/week	10587	4.7	4.8	3.1	3.2
Daily	50350	6.8	8.0	2.8	3.4
Fish/ seafood					
Rarely/never	163697	3.2	2.1	2.9	2.4
Monthly	108481	3.5	3.3	2.7	2.8
1-3 days/week	194181	3.2	3.7	2.7	2.9
4-6 days/week	31334	3.3	3.5	3.5	3.0
Daily	14117	4.1	4.4	3.4	3.3
Meat					
Rarely/never	24443	3.4	2.8	2.6	2.8
Monthly	64109	3.2	2.4	2.7	2.5
1-3 days/week	181826	3.0	2.8	2.7	2.6
4-6 days/week	91669	3.1	2.6	2.8	2.6
Daily	149763	3.6	4.1	2.8	3.1
Rice					
Rarely/never	16582	4.1	2.3	5.7	2.3
Monthly	48486	4.9	2.0	2.9	2.3
1-3 days/week	80134	4.6	3.4	2.6	3.1
4-6 days/week	10437	3.3	4.3	2.4	3.3
Daily	356171	3.4	3.2	2.6	2.7

Table 4.6. continued

Dietary consumption	No. of participants	Self-reported diabetes prevalence (%)		Screen-detected diabetes prevalence (%)	
		Minimal adjustment ^a	Full adjustment ^b	Minimal adjustment ^a	Full adjustment ^b
Wheat					
Rarely/never	50612	3.6	1.9	2.6	2.4
Monthly	100512	4.3	2.1	2.4	2.4
1-3 days/week	118829	3.7	3.4	2.8	2.7
4-6 days/week	37319	3.9	3.9	2.8	2.7
Daily	204538	4.2	3.6	3.1	3.0
Tea					
Never/almost never	181723	3.6	3.7	2.6	2.9
Occasionally	125101	3.2	2.8	2.7	2.5
Only at certain seasons	14559	3.1	2.3	2.8	2.5
Monthly but less than weekly	19833	3.1	2.9	3.0	2.9
At least once a week	170594	3.2	2.8	3.2	2.7

^aAdjusted for age, sex, education and study area; ^bAdditionally adjusted for alcohol, smoking, systolic blood pressure, physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, fresh vegetable consumption, meat consumption, soybean product consumption, fish consumption, dairy product consumption, rice consumption, wheat consumption, tea consumption.

Figure 4.4. Baseline self-reported diabetes prevalence by adiposity

Minimal adjustment: adjusted for age, sex, study area and education; Intermediate adjustment: BMI additionally adjusted for waist circumference and hip circumference; waist circumference additionally adjusted for hip circumference and BMI; hip circumference additionally adjusted for waist circumference and BMI; waist-to-hip ratio additionally adjusted for BMI; body fat percent additionally adjusted for BMI, waist circumference and hip circumference.

Squares represent diabetes prevalence and vertical lines represent corresponding 95% confidence intervals.

Data for body fat percent missing for 236 participants. BMI, body mass index.

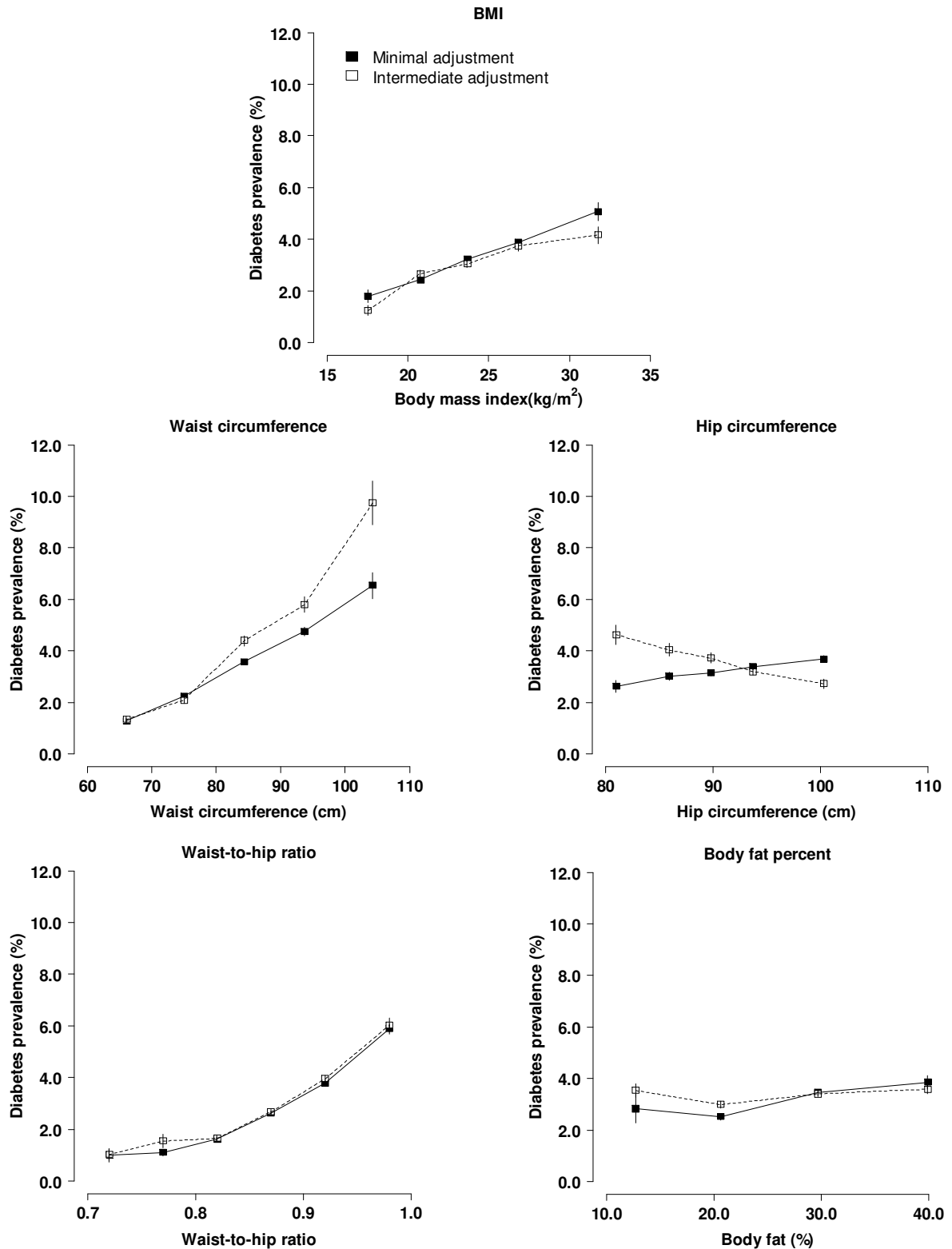


Figure 4.5. Baseline screen-detected diabetes prevalence by adiposity

Minimal adjustment: adjusted for age, sex, study area and education; Intermediate adjustment: BMI additionally adjusted for waist circumference and hip circumference; waist circumference additionally adjusted for hip circumference and BMI; hip circumference additionally adjusted for waist circumference and BMI; waist-to-hip ratio additionally adjusted for BMI; body fat percent additionally adjusted for BMI, waist circumference and hip circumference.

Squares represent diabetes prevalence and vertical lines represent corresponding 95% confidence intervals.

Data for body fat percent missing for 236 participants. BMI, body mass index.

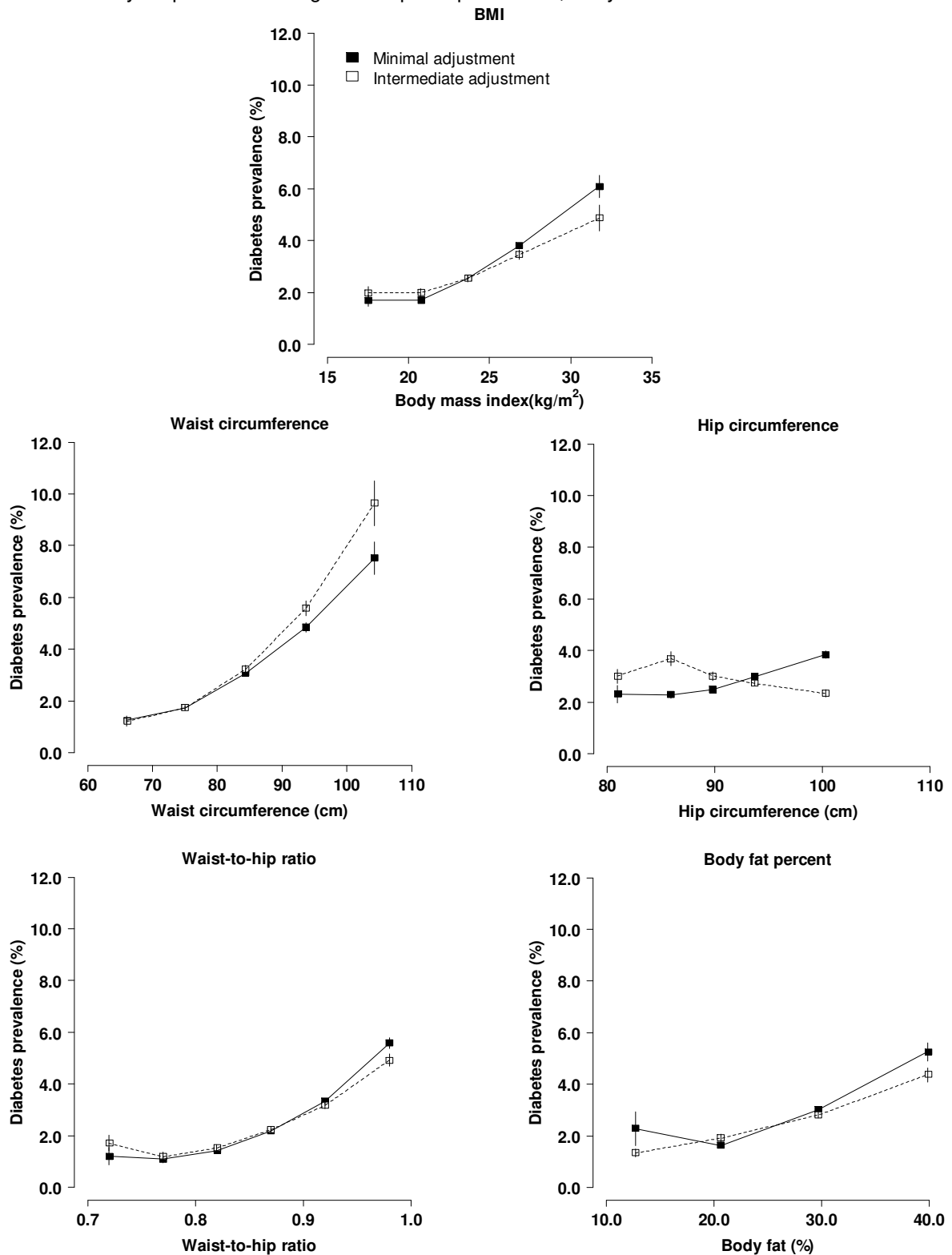


Figure 4.6. Baseline diabetes prevalence by blood pressure

Adjusted for age, sex, study area and education.

Squares represent diabetes prevalence and vertical lines represent corresponding 95% confidence intervals.

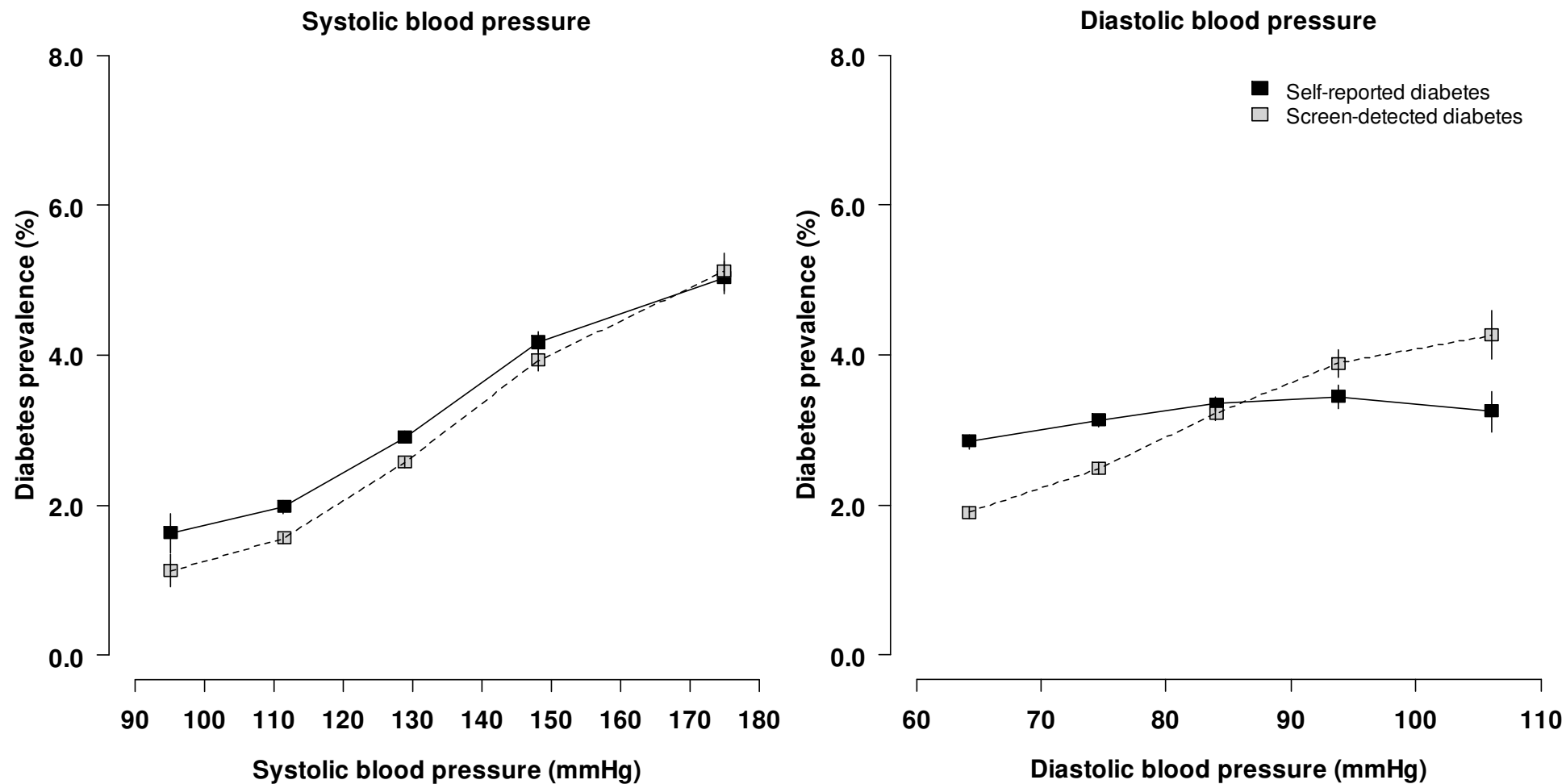


Figure 4.7. Frequency distribution of baseline random plasma glucose in individuals without self-reported diabetes at baseline

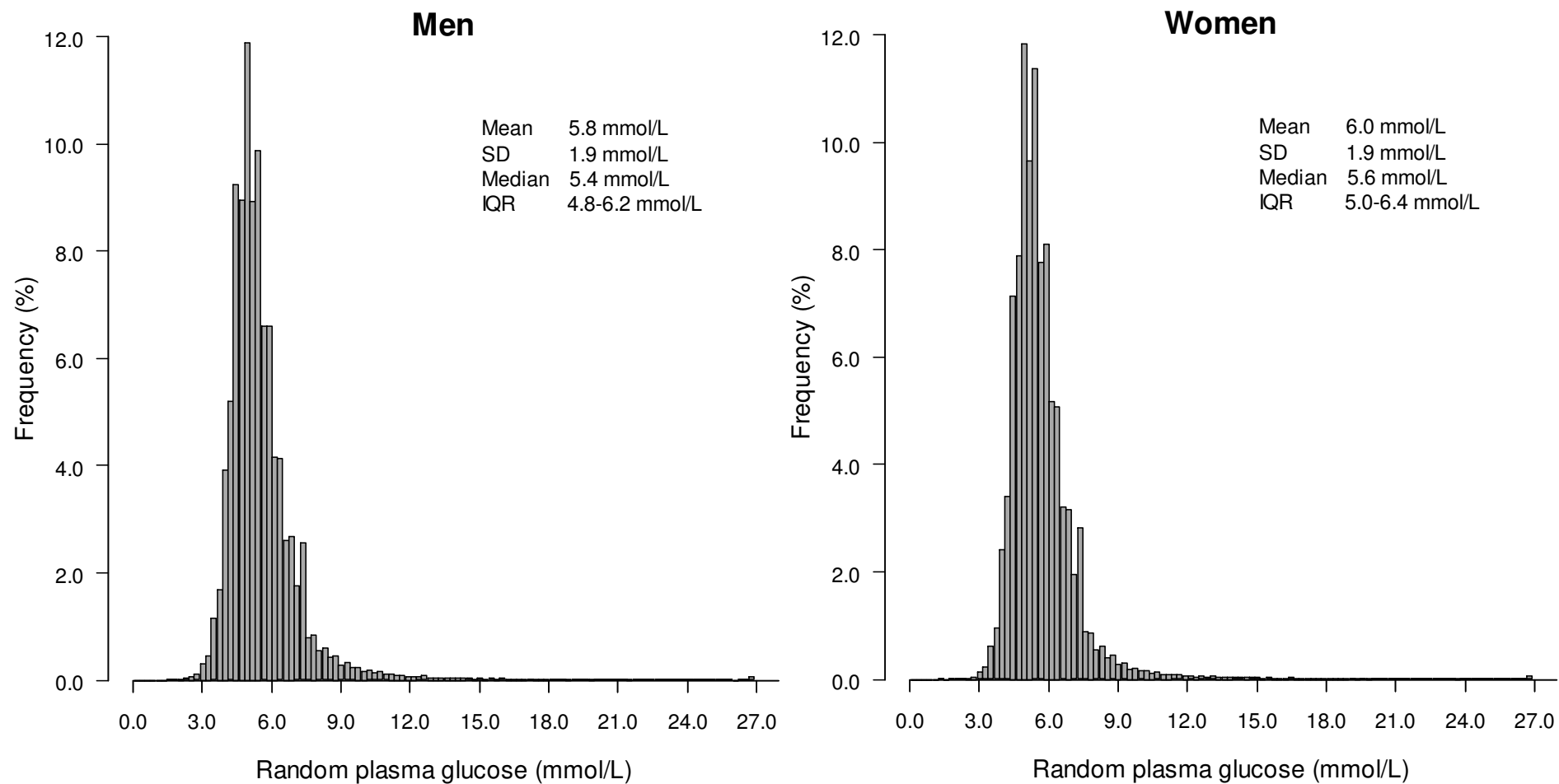


Figure 4.8. Mean baseline plasma glucose level by self-reported fasting time in individuals without self-reported diabetes at baseline

Adjusted for age, sex, study area and education.
Squares represent mean plasma glucose and vertical lines represent corresponding 95% confidence intervals.

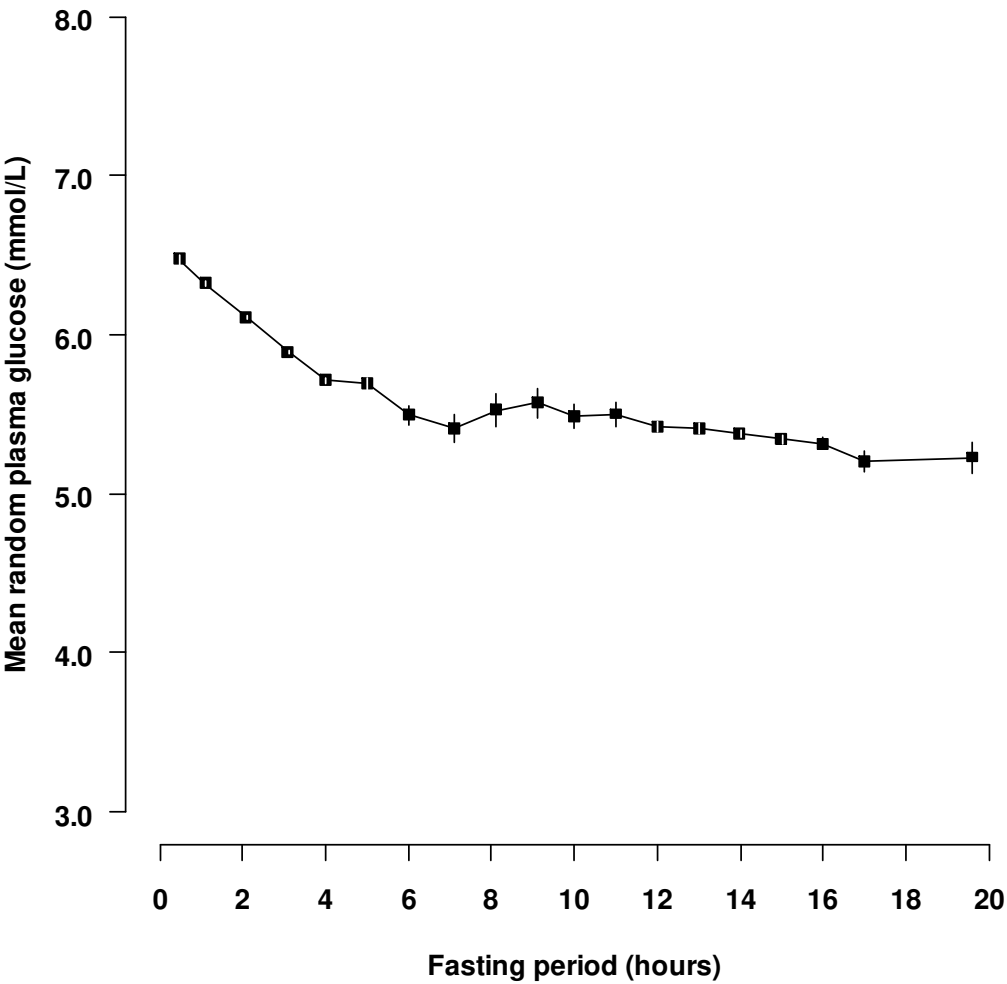


Table 4.7. Mean adjusted difference in baseline plasma glucose level between individuals with fasting period ≥ 8 hours and < 8 hours among individuals without self-reported diabetes at baseline

	Fasting period (hrs)							
	0 to <1 (<i>n</i> =13 258)	1 to <2 (<i>n</i> =74 666)	2 to <3 (<i>n</i> =128 635)	3 to <4 (<i>n</i> =107 823)	4 to <5 (<i>n</i> =44 407)	5 to <6 (<i>n</i> =10 288)	6 to <7 (<i>n</i> =3708)	7 to <8 (<i>n</i> =1824)
Difference in plasma glucose^a, mmol/L	1.1 (1.0-1.1) ^b	1.0 (0.9-1.0) ^b	0.7 (0.7-0.8) ^b	0.5 (0.5-0.5) ^b	0.3 (0.3-0.3) ^b	0.3 (0.2-0.3) ^b	0.1 (0.1-0.1) ^b	0.0 (0.0-0.0)

^aMean (95% confidence interval) difference from fasting (≥ 8 hr) plasma glucose (non-fasting minus fasting) adjusted for age, sex and study area; ^b $P < 0.001$.

Table 4.8. Mean baseline fasting plasma glucose and glycaemic categories among individuals without self-reported diabetes at baseline

	Men, n (%) (n=4837)	Women, n (%) (n=7399)	p value
Mean (SD) fasting plasma glucose	5.8 (1.1)	6.0 (1.1)	<0.001
Normoglycaemic^a	3112 (64.3)	4219 (57.0)	<0.001
Impaired fasting glucose^b	1155 (23.9)	2090 (28.3)	<0.001
Diabetes^c	570 (11.8)	1090 (14.7)	<0.001

13 192 (51.9%) participants who met the criteria for fasting plasma glucose testing (no self-reported diabetes at baseline with random plasma glucose ≥ 7.8 mmol/L and < 11.1 mmol/L) had no fasting plasma glucose data available.

^aFasting plasma glucose ≤ 6.0 mmol/L; ^bFasting plasma glucose 6.1-6.9 mmol/L; ^cFasting plasma glucose ≥ 7.0 mmol/L.⁷⁸

Figure 4.9. Baseline random plasma glucose by age in participants without self-reported diabetes at baseline

Adjusted for study area and education.
Squares represent mean random plasma glucose and vertical lines represent corresponding 95% confidence intervals.

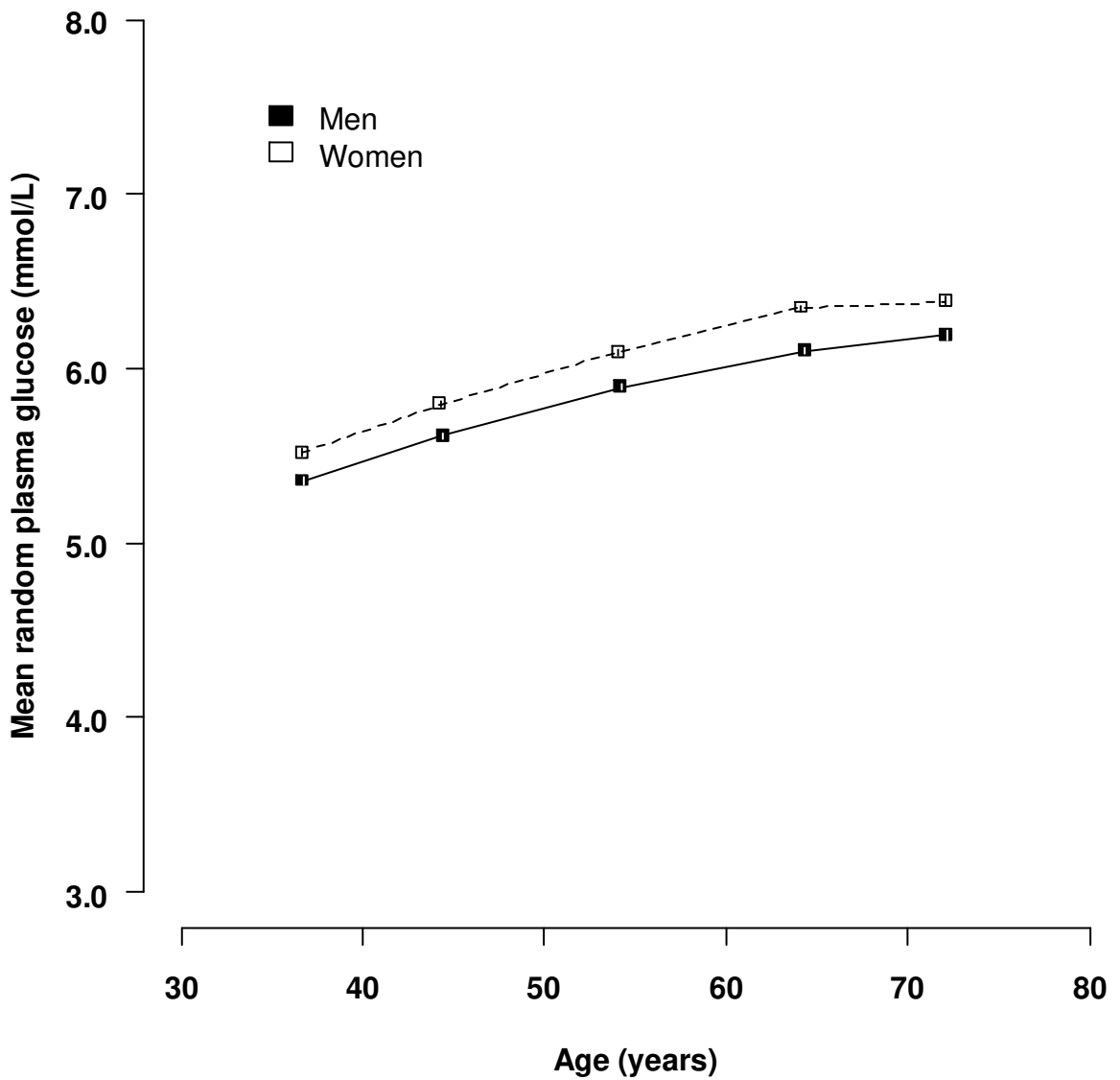


Table 4.9. Baseline random plasma glucose by study area in participants without self-reported diabetes at baseline

Characteristic	Mean (SE) random plasma glucose ^a	
	Men (<i>n</i> =200 433)	Women (<i>n</i> =287 119)
Urban		
Qingdao	6.1 (0.02)	6.1 (0.01)
Harbin	6.1 (0.01)	6.1 (0.01)
Haikou	5.9 (0.02)	6.0 (0.01)
Suzhou	5.9 (0.01)	6.0 (0.01)
Liuzhou	6.0 (0.01)	6.2 (0.01)
<i>Total</i>	<i>6.0 (0.01)</i>	<i>6.1 (0.01)</i>
Rural		
Sichuan	6.0 (0.01)	6.1 (0.01)
Gansu	5.6 (0.01)	5.9 (0.01)
Henan	5.6 (0.01)	6.0 (0.01)
Zhejiang	5.3 (0.01)	5.5 (0.01)
Hunan	5.6 (0.01)	5.8 (0.01)
<i>Total</i>	<i>5.6 (0.01)</i>	<i>5.9 (0.00)</i>
Total	5.8 (0.00)	6.0 (0.00)

^aAdjusted for age and education.

Table 4.10. Baseline random plasma glucose by lifestyle characteristics in participants without self-reported diabetes at baseline

Lifestyle characteristic	No. of participants	Mean random plasma glucose		
		Minimal adjustment ^a	Intermediate adjustment	Full adjustment ^b
Education^c				
No formal schooling	88676	6.0	-	5.9
Primary school	157131	6.0	-	5.9
Middle school	138713	5.9	-	5.9
High school	74560	5.8	-	5.9
College or university	28472	5.8	-	5.8
Smoking^d				
Never	301896	5.9	5.9	5.9
Occasional	27734	5.9	5.9	5.9
Ex-regular	28059	6.0	6.0	5.9
Current regular	129863	5.8	5.8	5.9
Alcohol^e				
Never regular	221895	5.9	5.9	5.9
Occasional	157044	5.8	5.8	5.9
Ex-regular	8335	6.0	5.9	5.9
Reduced intake	10691	5.9	5.9	5.9
Monthly	16875	5.8	5.8	5.8
Weekly	72712	5.9	5.9	5.9
Physical activity (MET hours/day)				
<10	35717	5.9	-	5.9
10-19.9	168970	5.9	-	5.9
20-29.9	121928	5.9	-	5.9
30-39.9	79424	5.9	-	5.9
≥40	81513	5.9	-	5.9

^aAdjusted for age, sex, education and study area; ^bAdditionally adjusted for alcohol, smoking, systolic blood pressure, physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, fresh vegetable consumption, meat consumption, soybean product consumption, fish consumption, dairy product consumption, rice consumption, wheat consumption, tea consumption; ^cAdjusted for variables as above excluding education; ^dIntermediate adjustment: adjusted for age, sex, education, study area and alcohol consumption; ^eIntermediate adjustment: adjusted for age, sex, education, study area and smoking. MET, metabolic equivalent of task.

Table 4.11. Baseline random plasma glucose by dietary consumption in participants without self-reported diabetes at baseline

Dietary consumption	No. of participants	Mean random plasma glucose	
		Minimal adjustment ^a	Full adjustment ^b
Fresh fruit			
Rarely/never	29274	6.0	5.9
Monthly	167247	5.9	5.9
1-3 days/week	153748	5.9	5.9
4-6 days/week	45641	5.9	5.9
Daily	91642	5.9	5.9
Fresh vegetables			
Rarely/never	117	6.2	6.2
Monthly	1375	6.1	6.0
1-3 days/week	6870	5.9	5.9
4-6 days/week	17236	5.9	5.9
Daily	461954	5.9	5.9
Soybean products			
Rarely/never	53439	5.9	5.9
Monthly	142341	5.9	5.9
1-3 days/week	245329	5.9	5.9
4-6 days/week	28706	5.9	5.9
Daily	17737	5.9	5.9
Dairy products			
Rarely/never	335268	5.9	5.9
Monthly	54644	5.9	6.0
1-3 days/week	41796	5.9	5.9
4-6 days/week	9984	5.9	5.9
Daily	45860	5.9	5.9
Fish/ seafood			
Rarely/never	157158	5.9	5.9
Monthly	102248	5.9	5.9
1-3 days/week	184790	5.9	5.9
4-6 days/week	29953	5.9	5.9
Daily	13403	5.9	5.9
Meat			
Rarely/never	23622	6.0	6.0
Monthly	61845	5.9	6.0
1-3 days/week	173223	5.9	5.9
4-6 days/week	87514	5.9	5.9
Daily	141348	5.9	5.9
Rice			
Rarely/never	16165	5.9	5.8
Monthly	47435	5.8	5.8
1-3 days/week	77171	5.9	5.9
4-6 days/week	9909	5.9	5.9
Daily	336872	5.9	5.9

Table 4.11 continued

Dietary consumption	No. of participants	Mean random plasma glucose	
		Minimal adjustment ^a	Full adjustment ^b
Wheat			
Rarely/never	47214	5.9	5.9
Monthly	95792	5.9	5.9
1-3 days/week	112655	5.9	5.9
4-6 days/week	35421	5.9	5.9
Daily	196470	5.9	5.9
Tea			
Never or almost never	172912	5.9	5.9
Occasionally	119360	5.9	5.9
Only at certain seasons	13599	5.9	5.9
Monthly but less than weekly	18898	5.9	5.9
At least once a week	162783	5.9	5.9

^aAdjusted for age, sex, education and study area; ^bAdditionally adjusted for alcohol, smoking, systolic blood pressure, physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, fresh vegetable consumption, meat consumption, soybean product consumption, fish consumption, dairy product consumption, rice consumption, wheat consumption, tea consumption.

Figure 4.10. Associations of baseline random plasma glucose with adiposity measures in participants without self-reported diabetes at baseline

Minimal adjustment: adjusted for age, sex, study area and education; Intermediate adjustment: BMI additionally adjusted for waist circumference and hip circumference; waist circumference additionally adjusted for hip circumference and BMI; hip circumference additionally adjusted for waist circumference and BMI; waist-to-hip ratio additionally adjusted for BMI; body fat percent additionally adjusted for BMI, waist circumference and hip circumference.

Squares represent mean random plasma glucose and vertical lines represent corresponding 95% confidence intervals.

Data for body fat percent missing for 228 participants. BMI, body mass index.

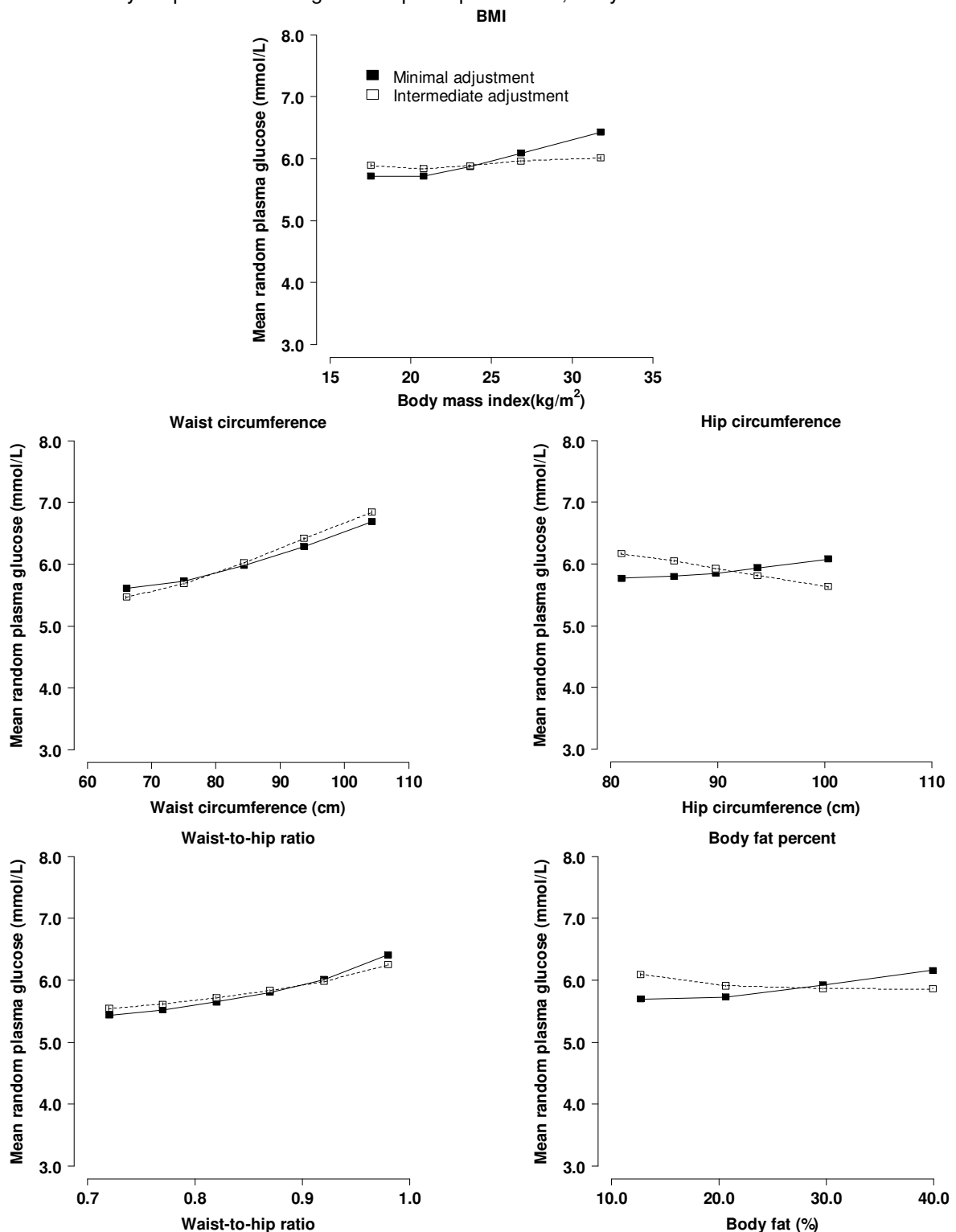


Figure 4.11. Associations of baseline random plasma glucose with blood pressure in participants without self-reported diabetes at baseline

Adjusted for age, sex, study area and education.

Squares represent mean random plasma glucose and vertical lines represent corresponding 95% confidence intervals.

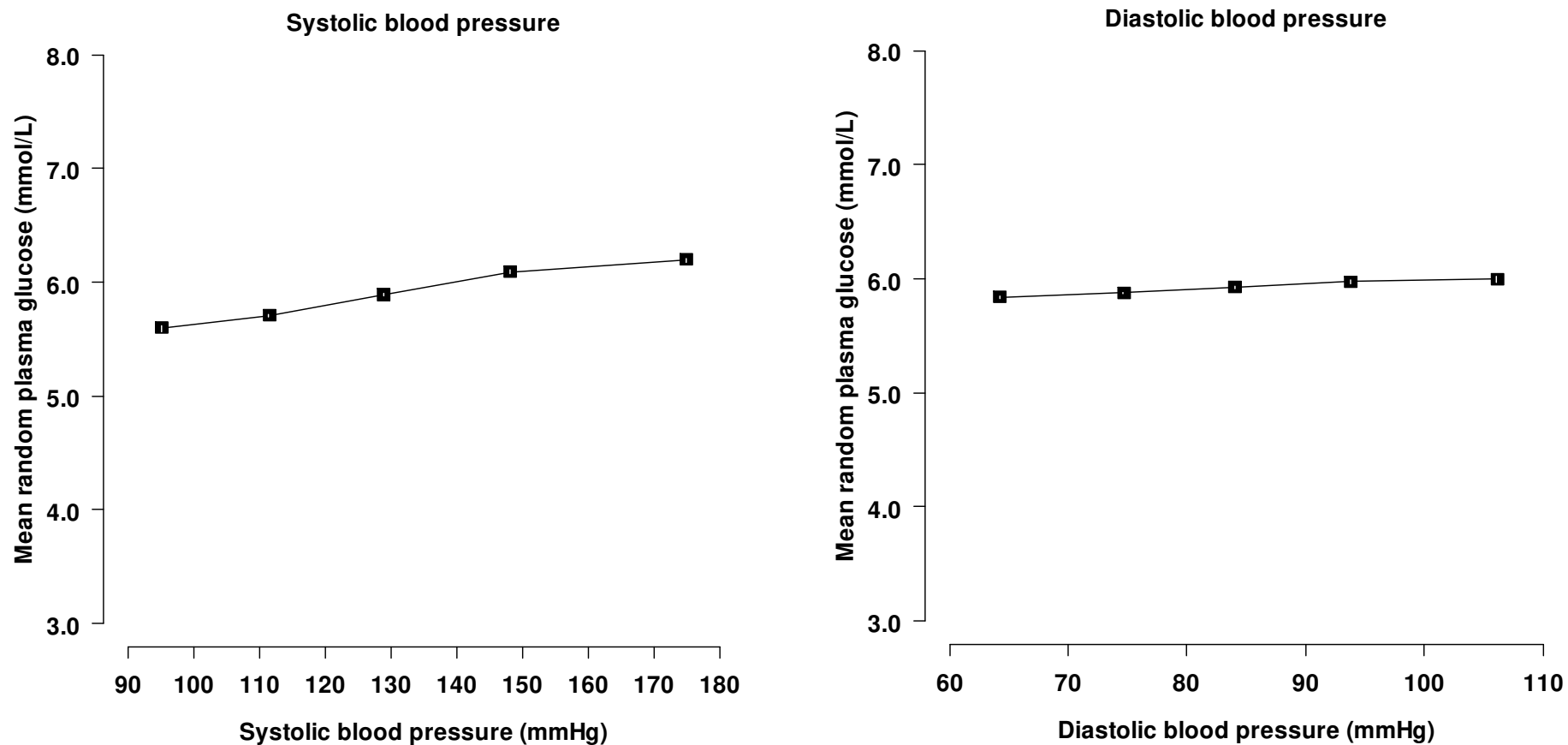


Figure 4.12. Changes in mean random plasma glucose between baseline and first resurvey in five groups defined by baseline random plasma glucose in participants without self-reported diabetes at baseline (n= 17 863)

Baseline random plasma glucose categories: <5.8, 5.8-6.7, 6.8-7.7, 7.8-11.0, ≥11.1 mmol/L.
Numbers represent mean baseline and resurvey random plasma glucose levels.

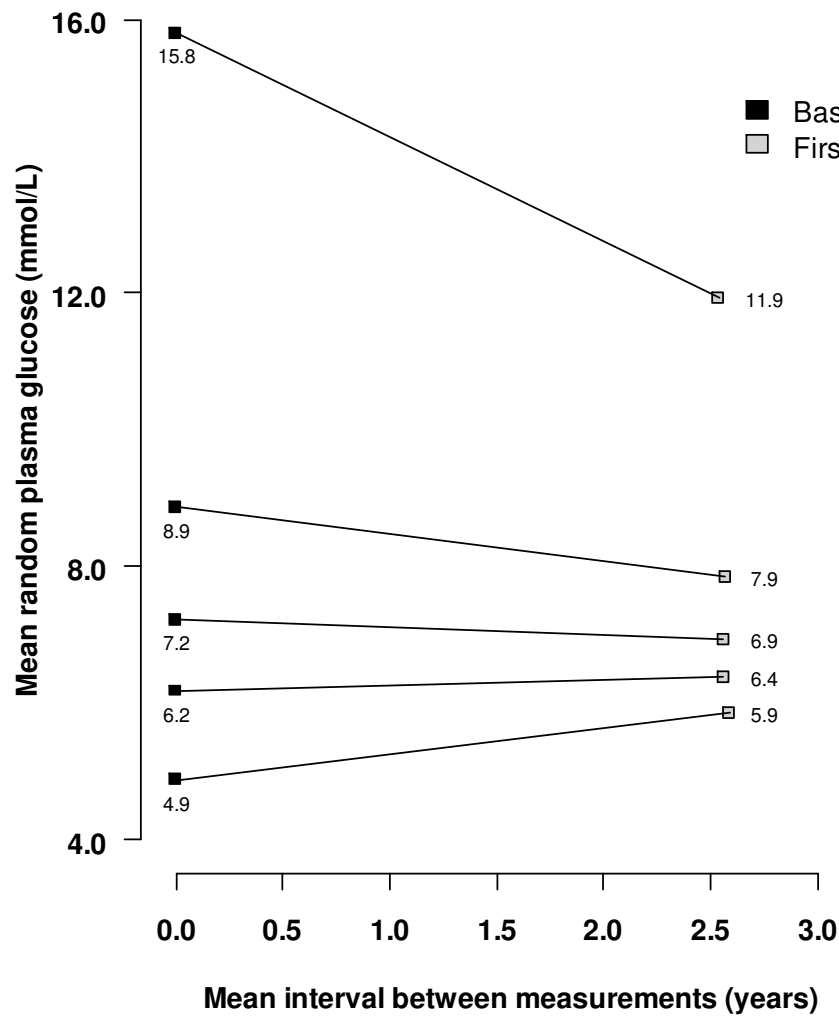


Table 4.12. Age and sex-specific regression dilution ratios based on baseline and first resurvey in participants without self-reported diabetes at baseline, estimated by different methods

Age at baseline, years	No. of participants	Mean baseline to resurvey interval (years)	Regression dilution ratio		
			MacMahon's	Rosner's	Self-correlation ^a
Men					
30-59	5188	2.6	0.47	0.41	0.24
60-69	1369	2.6	0.58	0.53	0.28
70-79	429	2.6	0.53	0.55	0.30
All	6986	2.6	0.51	0.46	0.26
Women					
30-59	8738	2.6	0.59	0.52	0.29
60-69	1680	2.6	0.55	0.48	0.31
70-79	459	2.6	0.56	0.55	0.34
All	10877	2.6	0.58	0.52	0.31
Total- unadjusted	17863	2.6	0.56	0.50	0.29
Total- age- and sex-adjusted	17863	2.6	N/A	0.48	0.26

^aCalculated using Spearman's rank correlation coefficient.

Figure 4.13. Changes in mean random plasma glucose between baseline and first and second resurveys in five groups defined by baseline random plasma glucose in participants without self-reported diabetes at baseline (n= 13 642)

Baseline random plasma glucose categories: <5.8, 5.8-6.7, 6.8-7.7, 7.8-11.0, ≥ 11.1 mmol/L.
Numbers represent mean baseline and resurvey random plasma glucose.

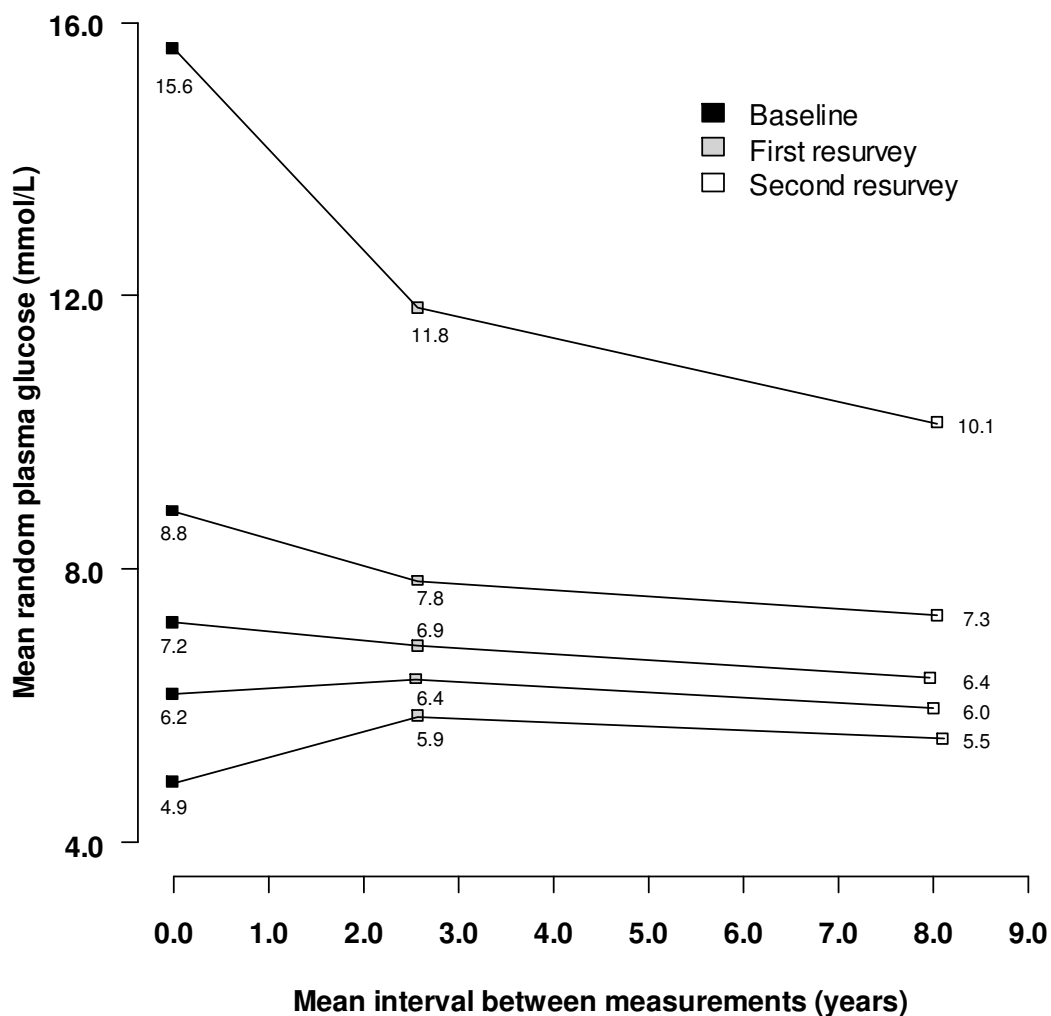
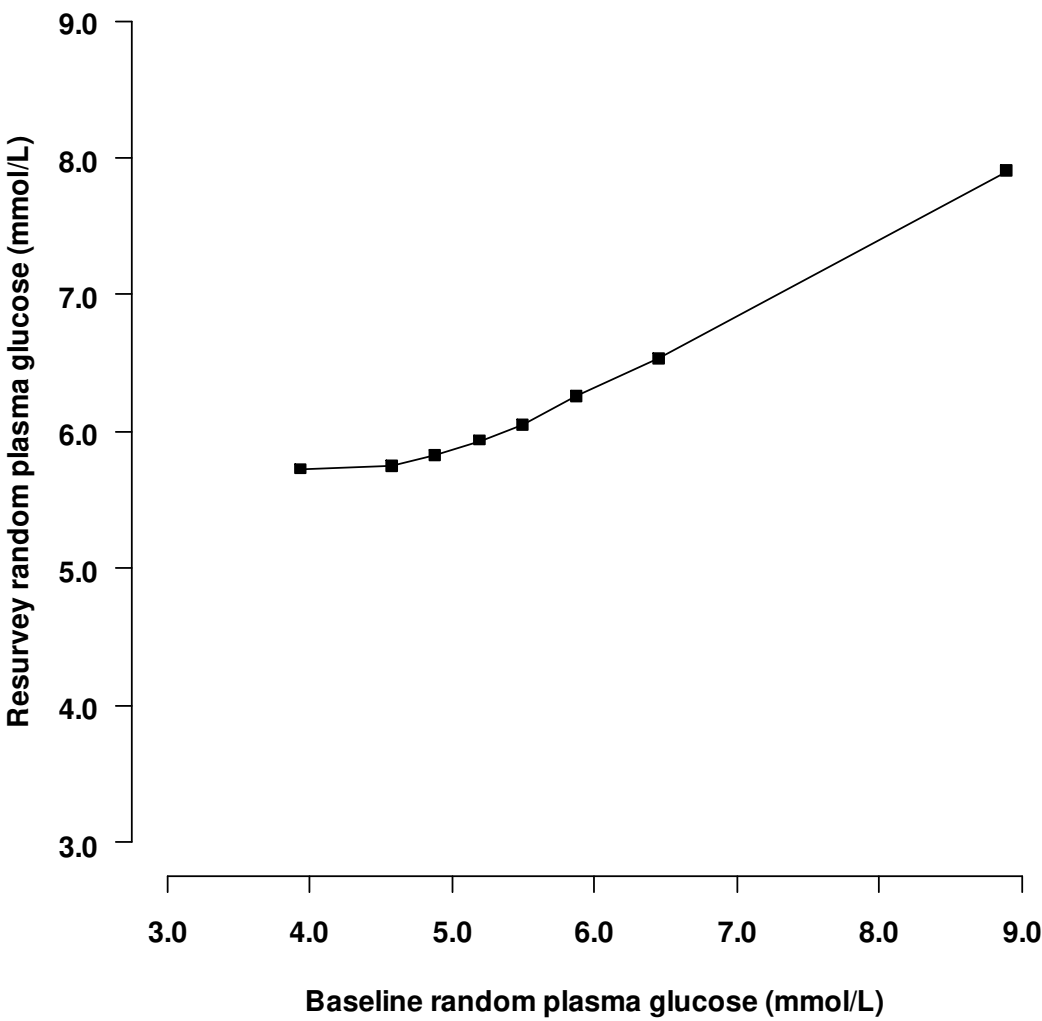


Figure 4.14. Relationship between baseline and first resurvey random plasma glucose levels in participants without self-reported diabetes at baseline



Chapter 5. Diabetes, plasma glucose and prevalent cardiovascular disease

5.1 Background

As described in Chapter 2, the associations of diabetes with prevalent IHD and stroke have been investigated in several previous studies,¹¹⁰⁻¹²⁶ but these have generally been small and of varying quality, with estimated diabetes-associated odds of prevalent IHD ranging from 1.4¹¹⁹ to 6.1.¹¹⁴ For stroke, the results from previous studies have been even more varied, with some studies showing no association,^{113, 123, 125, 126} and others reporting between 1.4-¹²² and 10.5-fold¹²¹ increased odds. Furthermore, previous studies of both prevalent IHD and stroke have been insufficiently detailed to investigate effect modification beyond the effects of sex. The characteristic differences in CVD patterns between China and the West, with typically higher rates of stroke and lower rates of IHD in China,^{236, 237} the rapid increase in diabetes prevalence in China over recent decades⁸⁶ and proposed differences in the pathophysiology of diabetes between China and the West⁸⁸ suggest that findings from Western populations, in which most previous studies have been conducted, may not be readily generalisable to Chinese populations. As described, previous studies have provided very limited evidence on the association of blood glucose levels below the diabetic threshold with prevalent IHD and stroke,^{139-141, 238} with none reporting on this association in Chinese populations. Of note, and as intimated previously, many of the described limitations of previous studies are relevant not only to Chinese populations, but to East Asian populations more generally.

This chapter addresses some of these gaps in our understanding using cross-sectional data from the CKB. The objectives of the chapter were to examine a) the prevalence and correlates of self-reported doctor-diagnosed IHD and stroke/TIA in the CKB; b) the relationship between self-reported doctor-diagnosed diabetes, and screen-detected diabetes, and the prevalence of CVDs; c) the association between RPG levels and the

prevalence of CVDs among people without diabetes; and d) the relevance of age, sex, area, education, smoking status, alcohol consumption, blood pressure, physical activity and adiposity to the observed associations.

5.2 Methods

5.2.1 Exposure definitions

Diabetes and RPG were defined as detailed in Sections 4.2.1 and 4.2.2. Participants were classified as having self-reported diabetes if they answered yes to the question “Has a doctor ever told you that you had had diabetes?” in the baseline survey. Screen-detected diabetes was defined as not having self-reported diabetes, but having a plasma glucose level ≥ 7.0 mmol/L if time since last food ≥ 8 hours, or ≥ 11.1 mmol/L if time since last food < 8 hours, or a FPG level ≥ 7.0 mmol/L on subsequent testing.

5.2.2 Outcome definitions

Self-reported doctor-diagnosed IHD and stroke/TIA were defined as positive responses to the questions “Has a doctor ever told you that you had had coronary heart disease?” and “Has a doctor ever told you that you had had a stroke or transient ischaemic attack?”, respectively.

5.2.3 Selection of confounders

Based upon a priori knowledge of the biological relationships of blood glucose and diabetes with IHD and stroke/TIA, the following variables were identified as potential confounding factors: age, sex, study area, education, household income, smoking, alcohol, family history of diabetes, physical activity, BMI, waist circumference, hip circumference, waist-to-hip ratio, percentage body fat, SBP, DBP and frequency of fresh fruit, meat, dairy products, fresh eggs, wheat and tea consumption. The associations of these variables with IHD and stroke/TIA prevalence were investigated as described below (Section 5.2.5). The ultimate selection of variables to adjust for was based upon demonstration of an association with

exposure and outcome in these analyses, use of causal diagrams, which enable identification of the appropriate variables to adjust for to reduce bias and confounding in effect size estimates^{239, 240} (Figure D.1), and examination of the effect of adjusting for these variables in regression analyses.

5.2.4 Participant exclusions

As described in Section 4.2.3, participants with missing data for BMI or extreme or implausible values for SBP, DBP, height, waist circumference, hip circumference, waist-to-hip ratio or BMI (n=1081) were excluded from analyses examining the associations of diabetes, which included 511 810 participants. For analyses examining the associations of RPG, a further 24 258 participants with self-reported diabetes or missing RPG data at baseline were additionally excluded, leaving 487 552 participants.

5.2.5 Statistical analyses

Correlates of prevalent cardiovascular diseases

The prevalence of self-reported IHD and stroke/TIA were compared across sex, age and study area, standardised (where appropriate) to the age, sex, study area and education structure of the whole study population. Logistic regression analyses were undertaken to examine the association of prevalent IHD and stroke/TIA with potential confounding variables. Where appropriate, the floating absolute risk method was used; this provides estimates of variance across all exposure categories enabling comparisons between any two categories, and not only with the reference group.²⁴¹ All associations were examined adjusted for the effects of age, sex, study area and education, and simultaneously adjusted for all potential confounding variables (age, sex, education, study area, alcohol, smoking, SBP, physical activity, BMI, waist circumference, hip circumference and consumption of fresh fruit, meat, dairy products, eggs, wheat and tea). An intermediate level of adjustment was investigated for some associations to better understand the nature of the associations independent of other relevant potential confounding variables; in addition to adjustment for

age, sex, study area and education, associations with smoking were adjusted for alcohol consumption, and vice versa, associations with measures of adiposity were adjusted for each other (Section 4.2.4) and associations with DBP were adjusted for SBP.

Associations of diabetes and blood glucose levels with prevalent cardiovascular disease

The associations of self-reported and screen-detected diabetes, and of RPG among people without self-reported diabetes, with IHD and stroke/TIA prevalence were examined using multi-variable logistic regression. Individuals with screen-detected diabetes were excluded from self-reported diabetes analyses and vice versa. RPG was categorised into six groups (<4.8, 4.8-5.7, 5.8-6.7, 6.8-7.7, 7.8-11.0, ≥ 11.1 mmol/L); cut-points were selected to include the OGTT 2-hour post glucose-load thresholds for diabetes and IGT,⁷⁸ and at 1 mmol/L intervals below this. ORs for prevalent IHD and stroke/TIA were calculated, initially simultaneously adjusting for age, sex (where appropriate) and study area, with the previously described potential confounders (Section 5.2.3) then progressively added into the model. The main and subsidiary analyses were adjusted for age, sex (where appropriate), study area, education (no formal school, primary school, middle school, high school, college or university), smoking status (never, occasional, ex-regular, current regular), alcohol (never regular, occasional intake, ex-regular, reduced intake, weekly intake), SBP (<100, 100-109, 110-119, 120-129, 130-139, 140-149, 150-159, 160-169, ≥ 170 mmHg) and physical activity (<10, 10-19.9, 20-29.9, 30-39.9, ≥ 40 MET hours/day). Additional analyses were adjusted for waist-to-hip ratio (<0.75, 0.75 to <0.80, 0.80 to <0.85, 0.85 to <0.90, 0.90 to <0.95, ≥ 0.95). Adjustment for SBP, physical activity and waist-to-hip ratio as continuous, rather than categorical, variables did not alter risk estimates. The floating absolute risk method was used in logistic regression analyses of the association of RPG with prevalent IHD and stroke/TIA.²⁴¹ Chi-squared tests for trend in log ORs were conducted, and the estimated ORs of prevalent IHD and stroke/TIA were examined for departure from linearity by testing the type 3 chi-square for categorical RPG, in a model

that contained RPG as a categorical and a continuous variable. ORs for each additional 1 mmol/L of RPG were estimated in participants with a RPG level <11.1 mmol/L.

Adjusted ORs were compared separately across strata of age, sex, area, education, smoking status, alcohol intake, blood pressure, physical activity and adiposity. Self-reported diabetes-associated ORs were also compared across strata of treatment status and duration of diabetes diagnosis. RPG-associated ORs were repeated stratified by fasting period using a cut-point of 8 hours, selected based on fasting duration guidance¹⁸³ and the threshold in the relationship between RPG and fasting time in the CKB (Section 4.3.3). Chi-squared tests for trend and heterogeneity were applied to log ORs and their standard errors (SE) for each variable.²⁴² Global tests of heterogeneity were calculated by summing separate heterogeneity chi-square statistics for each subgroup analysis and their degrees of freedom, to examine overall consistency of ORs across the strata, also taking account of multiple testing.

Statistical analyses were conducted using SAS version 9.3 (SAS Institute Inc., Cary, North Carolina, USA). All plots were produced using R version 2.13.1 (R Foundation for Statistical Computing, Vienna, Austria), and forest plots were produced using a R graphics programme developed within CTSU.

5.3 Results

5.3.1 Prevalence and correlates of cardiovascular disease

Overall, of the 511 810 participants included in analyses of the associations of diabetes, 3.0% (n=15 412) and 1.7% (n=8860) reported a history of doctor-diagnosed IHD and stroke/TIA, respectively, at baseline. The prevalence of both IHD and stroke/TIA increased significantly with age in both men and women after adjusting for study area and education (Figure 5.1). IHD prevalence was higher in women than in men at all ages, with an overall study area- and education-adjusted prevalence of 2.4% in men and 3.5% in women, whilst the opposite was true for stroke/TIA with an overall prevalence of 2.3% and 1.4% in men

and women, respectively. The prevalence of both conditions varied greatly across study areas after adjusting for age, sex and education, and were higher in urban (IHD: 3.6%; stroke/TIA: 2.0%) than in rural areas (IHD: 2.0%; stroke/TIA: 1.4%) (Figures 5.2 and 5.3). Among the 487 552 participants included in analyses of the associations of RPG, the prevalence of IHD and stroke/TIA at baseline were slightly lower (2.7%, n=13 250 and 1.6%, n=7706, respectively), but the trends in prevalence with age and sex and across study areas were comparable.

Among 14 854 participants included in both first and second resurveys 419 (2.8%) reported a history of doctor-diagnosed IHD at baseline. Of them, 48.7% (n=204) again reported a history of IHD at first resurvey, 46.5% (n=195) at second resurvey (Table 5.1), and 68.5% (n=287) reported a history of IHD at either first or second resurvey. After excluding participants with new self-reported IHD at resurvey (considered as new IHD diagnoses) kappa statistics for agreement in self-reported IHD diagnoses at baseline and resurvey were 0.65 (95% CI 0.60-0.69) for the first resurvey and 0.63 (0.58-0.67) for the second resurvey. For participants with a history of stroke/TIA at baseline and included in the two resurveys (n=154), 59.7% (n=92) and 55.8% (n=86) reported a history of stroke/TIA at first and second resurveys, respectively, and 77.3% (n=119) at either first or second resurvey (Table 5.2). After excluding new onset stroke/TIA between baseline and resurvey, kappa statistics for agreement in self-reported diagnoses of stroke/TIA were 0.75 (0.69-0.81) for baseline and first resurvey and 0.71 (0.65-0.78) for baseline and second resurvey.

Detailed discussion of the associations of the main potential confounding variables with prevalent IHD and stroke/TIA is beyond the scope of this thesis. Briefly, there were positive associations of IHD and stroke/TIA prevalence with adiposity and blood pressure, inverse associations with physical activity and higher odds among ex-regular smokers and occasional alcohol drinkers (Figures D.2 to D.5 and Tables D.1 to D.4).

5.3.2 Diabetes and prevalent cardiovascular disease

After adjusting for age, area, education, smoking, alcohol, SBP and physical activity, individuals with self-reported diabetes were more than twice as likely to have IHD (OR 2.16, 95% CI 2.05-2.28) than those without (Table 5.3). Women aged 30-59 years had higher self-reported diabetes-associated odds of IHD than older women, but there was no difference with age in men. The association of self-reported diabetes with prevalent stroke/TIA was modestly weaker than the association with IHD (1.88, 1.75-2.02), and there was a somewhat stronger association of diabetes with stroke/TIA in women (2.09, 1.90-2.31) than in men (1.64, 1.47-1.82) ($p < 0.001$). The youngest women had higher self-reported diabetes-associated odds of prevalent stroke/TIA (2.72, 2.31-3.21) than older women, but this trend was not seen in men. Further adjustment for waist-to-hip ratio moderately attenuated associations with both IHD and stroke/TIA (Table 5.3), but additional adjustment for other potential confounding factors (eg, DBP and dietary intake) had little effect (Figure 5.4). Adjusted ORs did not vary substantially between study areas, although there was a somewhat weaker association of diabetes with prevalent IHD (1.91, 99% CI 1.69-2.15) and stroke/TIA (1.58, 99%CI 1.36-1.84) in one urban region, Harbin (Figure 5.5).

For prevalent IHD, the adjusted self-reported diabetes-associated ORs were similar across strata of other risk factors with the exception of increasing odds with increasing BMI (p for trend < 0.001) (Figure 5.6). For prevalent stroke/TIA, the self-reported diabetes-associated OR increased with increasing levels of physical activity (p for trend = 0.01) and were lower among ex-regular alcohol drinkers (p for heterogeneity < 0.001), but otherwise varied little by other risk factors (Figure 5.7). There was an apparent increase in the odds of prevalent IHD with increasing duration of diabetes diagnosis (p for trend = 0.008; Figure 5.8). After adjusting for other cardiovascular risk factors, the ORs for IHD were significantly higher in participants reporting hypoglycaemic medication use than in those not using such medication (OR 2.23 versus 1.68; p for heterogeneity < 0.001 ; Figure 5.9). For prevalent stroke/TIA, no clear association was seen with use of hypoglycaemic medications (p for

heterogeneity=0.1) (Figure 5.9) or with duration of diabetes diagnosis (p for trend=0.4) (Figure 5.8).

After adjusting for age, area, education, smoking, alcohol, SBP and physical activity, there was no clear difference in the odds of prevalent IHD (OR 1.02, 95% CI 0.94-1.12) or stroke/TIA (0.88, 0.80-0.98) between individuals with screen-detected diabetes and those without any diabetes (Table 5.4).

5.3.3 Blood glucose and prevalent cardiovascular disease

Among participants without self-reported diabetes, there was a positive association between RPG levels and IHD prevalence, with ORs of 1.00, 0.99, 1.06, 1.15, 1.20 and 1.13 at RPG levels of <4.8 (reference), 4.8-5.7, 5.8-6.7, 6.8-7.7, 7.8-11.0 and ≥ 11.1 mmol/L, respectively (p for trend <0.001), after adjusting for age, sex and study area. Further adjustment for education, smoking, alcohol consumption, SBP and physical activity attenuated the association, with ORs of 1.00, 0.98, 1.04, 1.12, 1.15 and 1.05 in the same RPG categories, but a significant trend remained (p for trend <0.001; Figure 5.10). Below 11.1 mmol/L the association did not deviate significantly from log-linearity, with each additional 1 mmol/L of RPG associated with 4% (95% CI 2-5%) higher odds of prevalent IHD (Table 5.5). The prevalence of stroke/TIA also increased with increasing RPG levels, with basically adjusted ORs of 1.00, 1.05, 1.15, 1.26, 1.39 and 1.70 in the same six RPG categories (p for trend <0.001). Further adjustment, as previously, attenuated the associations, with ORs of 1.00, 1.05, 1.11, 1.19, 1.25, 1.35, but the trend remained significant (p for trend <0.001) (Figure 5.10). At RPG levels <11.1 mmol/L there was no significant deviation of the association from log-linearity, and each additional 1 mmol/L of RPG was associated with 5% (3-7%) greater odds of stroke/TIA (Table 5.5). The associations of RPG were similar in men and women (Figure D.6). Additional adjustment for waist-to-hip ratio attenuated the associations of RPG levels with IHD and stroke/TIA (1.02, 1.00-1.03 and 1.04, 1.02-1.05 per 1 mmol/L higher RPG at levels <11.1 mmol/L,

respectively) (Table 5.5). Further adjustment for potential confounding factors, including DBP and dietary variables, did not further attenuate the associations (Figure 5.11). Separate examination of the association of RPG with IHD in the 10 study areas did not show clear evidence of heterogeneity ($p=0.6$), although there was some variation in the association with stroke/TIA ($p=0.003$) (Figure 5.12).

The association of RPG with prevalent IHD appeared to be stronger among participants with lower SBP (p for trend=0.01; Figure 5.13). The association of RPG with prevalent stroke/TIA was stronger in rural than in urban areas ($p<0.001$), but otherwise varied little across strata of other cardiovascular risk factors (Figure 5.14); the significant global heterogeneity test ($p<0.001$) was likely driven largely by the urban/rural difference. There was no clear difference in the association of fasting and non-fasting plasma glucose with prevalent IHD or stroke/TIA (Figure 5.15).

5.4 Discussion

The presented analyses are the largest cross-sectional investigation of the associations of diabetes, and of plasma glucose levels among adults without diabetes, with CVDs in a Chinese population. They showed an independent doubling of the odds of prevalent IHD and stroke/TIA associated with self-reported diabetes. Among individuals without a prior diabetes diagnosis, there was an apparent positive association of RPG with both these conditions.

5.4.1 Diabetes and prevalent cardiovascular disease

A few Chinese studies have previously reported the positive association of diabetes with prevalent IHD.^{110, 113, 115} Two cross-sectional studies showed 2- to 2.5-fold increased odds of prevalent IHD associated with diabetes, similar to that shown in the CKB. In the CNDMDS, a nationally representative survey of >45 000 adults aged ≥ 20 years, diabetes (based on self-report and OGTT) was associated with an OR of 2.44 for IHD.¹¹³ Similar self-reported diabetes-associated odds (OR 2.51) were reported in another study of almost 60

000 adults aged >40 years in rural Beijing.¹¹⁰ In contrast, the Chinese component of the INTERHEART case-control study, including ~3000 AMI cases from 26 centres and a similar number of controls, reported much higher odds of AMI associated with self-reported diabetes (5.10).¹¹⁵ The discrepancy in the strength of the association between INTERHEART-China and other studies, including the CKB, may relate to the nature of IHD studied (MI, a more severe IHD outcome, was examined in the INTERHEART study), the degree of misclassification of IHD (likely greater in the two Chinese cross-sectional studies and in the CKB, where IHD was ascertained through self-report), variable adjustment for confounders, or the effect of additional biases, particularly selection bias, particularly problematic in the case-control study design. In Western population studies, reported quantitative associations of diabetes with prevalent IHD have also varied to some extent. In the largest of these studies, the Swedish Care Data Warehouse in Östergötland study, diabetes (determined through plasma glucose measurement) was associated with age-adjusted prevalence rate ratios for IHD of 2.65 in men and 3.97 in women.¹¹¹ Taking account of the minimal adjustment of these estimates, the findings are reasonably consistent with those of the Chinese cross-sectional studies.^{110, 113}

The diabetes-associated IHD OR estimates in the CKB, although modestly lower, are not inconsistent with previous cross-sectional studies,^{110, 111, 113} but are perhaps more consistent with prospective studies, which have tended to adjust for confounders more comprehensively. Prospective Chinese studies have shown between 1.5- and 2.5-fold increased risks of IHD associated with diabetes, variably defined through self-report^{130, 132} or self-report combined with blood glucose measurement.¹³¹ In a large individual participant data meta-analysis of almost 700 000 people from predominantly Western populations, diabetes (defined by self-report, medication use or FBG levels) was associated with a doubling of risk for IHD in a fully adjusted model.⁶ In another large meta-analysis, combining published and individual participant data from over 850 000 individuals, diabetes (again variably defined) was associated with an IHD HR of 2.16 in men and 2.82 in women, with

variable adjustment for confounding in the included studies.¹³³ A moderately weaker association was found in a study of almost two million adults in the UK, using data from routine health records, which described a 1.5-fold increased risk of MI and 1.6-fold increased risk of angina associated with diabetes.⁹⁰

Evidence on the association of diabetes with stroke in Chinese populations is limited. In the cross-sectional rural Beijing study, including almost 60 000 participants and an unspecified number of stroke events, self-reported diabetes was associated with a two-fold greater prevalence of stroke.¹¹⁰ A significantly elevated self-reported diabetes-associated risk of stroke was also found in a case-control study of ~1000 TS cases and the same number of matched controls in Chongqing, South West China, with an adjusted diabetes-associated OR of 3.65 in women and 1.81 in men.¹²⁴ These findings contrast, however, with those of the previously described CNDMDS in China, which showed no significant association of diabetes with prevalent TS (0.97, 95% CI 0.59-1.61), although the number of stroke cases included was relatively small (n=472).¹¹³ A prospective Chinese study found a 50% increased risk of TS associated with self-reported diabetes,¹³⁰ whilst several large prospective studies of non-Chinese populations,^{129, 136, 243, 244} including meta-analyses,¹³⁸ have shown 1.5- to 3-fold greater risk of TS associated with diabetes, reasonably consistent with the CKB findings.

We found significantly higher odds of prevalent stroke/TIA, but not IHD, among women than men. Similar sex-differences have been reported in some,^{106, 136, 245} though not all,^{90, 130} previous studies, and have been attributed variably to a more adverse diabetes-associated CVD risk profile in women,^{106, 136, 245} possibly due to women developing diabetes at a higher BMI than men,¹⁰⁶ or to treatment differences favouring men.^{136, 245}

As in the CKB, several previous studies have shown non-significantly lower risks of IHD^{6, 246} and stroke^{6, 247} in screen-detected diabetes (or amongst individuals with a comparable glycaemic status) than in self-reported diabetes. The lack of any clear association with

screen-detected diabetes in the CKB may reflect shorter disease duration or less severe glycaemic aberrations in screen-detected diabetes, selective diagnosis of diabetes amongst individuals with IHD or stroke/TIA, or a greater proportion of false-positive diabetes diagnoses in the screen-detected group.

5.4.2 Blood glucose and prevalent cardiovascular disease

Glycaemic thresholds for diabetes may be less relevant for macrovascular, than for microvascular, disease.²⁴⁸ One small cross-sectional US study of 2500 individuals without known diabetes but with hypertension or hyperlipidaemia, found significantly elevated independent odds of self-reported IHD (n=1274) at all FBG levels >4.8 mmol/L compared with <4.4 mmol/L.¹³⁹ Likewise, a small Indian case-control study, including 300 MI cases and matched controls found a positive, independent association between post-challenge plasma glucose and the odds of MI (OR 2.01, 95%CI 1.56-2.59 per 1 SD higher glucose).¹⁴¹

To my knowledge, no studies in mainland China have published data on the cross-sectional or prospective association of blood glucose with IHD or stroke risk among individuals without diabetes. One prospective study involving 16 500 individuals without diabetes in Taiwan showed no statistically significant association of CVD deaths with post-challenge (1.61, 95% CI 0.86-2.99, p for trend=0.05) or fasting (0.84, 0.47-1.51, p for trend=0.77) plasma glucose, comparing the highest and lowest quintiles.²⁴⁹ Several other prospective studies have found an increased risk of IHD only at or above glycaemic thresholds for diabetes or pre-diabetes.^{6, 82, 246} An individual participant data meta-analysis, including over 250 000 participants without known diabetes and ~13 000 IHD events in mainly Western populations, found a significantly elevated risk only at FBG levels >6.1 mmol/L.⁶ Other studies, however, have reported a significant positive association within the “normoglycaemic” range.¹⁰²⁻¹⁰⁴ In the KCPS of almost 1.2 million men and women there was a J-shaped association between FPG and IHD with the lowest risk at a FPG level of ~5.0 mmol/L, increasing to a HR of 1.15 at 6.1-6.9 mmol/L and still further above the

diabetes threshold.¹⁰² Similarly, the APCSC of ~240 000 individuals, found a log-linear association with IHD that extended down to a FBG level of 4.9 mmol/L.¹⁰³

Few studies have examined the relationship with RBG but, in contrast to the findings from the CKB, a published data meta-analysis including almost 10 000 participants and ~300 IHD events showed no significant association with fatal IHD (HR 1.02), although the level of adjustment is unclear, and the study likely lacked adequate statistical power.¹⁰⁴ These apparent inconsistencies between studies may reflect differences in populations studied, sample size, glycaemic measures, adjustment for confounders or for intra-individual variation in blood glucose levels and reverse causality.

Associations between blood glucose levels and stroke have been examined primarily in non-Chinese populations,^{102, 103, 247, 250} but none has investigated the association specifically with prevalent stroke or TIA. Two large prospective studies showed positive associations of fasting glucose levels with stroke risk within the “normoglycaemic” range,^{102, 103} and a study of >15 000 adults in Scotland estimated relative TS mortality rates per 1SD (men: 1.4 mmol/L; women: 1.2 mmol/L) higher RPG of 1.10 (95% CI 0.99–1.23) and 1.21 (1.11–1.33) in men and women, respectively.²⁵⁰ The stronger association of RPG with stroke/TIA than IHD in the CKB contrasts with studies in many Western populations demonstrating no difference,²⁵⁰ or possibly a stronger association with IHD than with stroke.²⁵¹ The role of small vessel pathology in cerebrovascular disease, thought to be particularly prominent in Chinese and other East Asian populations⁵⁶ could explain the stronger association with stroke/TIA.

5.4.3 Limitations

The CKB was not intended to be nationally representative, and the overall response rate for participation at baseline was only ~30%, although this compares favourably with other large nationwide biobank studies, such as the UK Biobank.²⁵² Despite this, the reported prevalence of IHD and stroke in the CKB were broadly similar to those from other

representative surveys in China. In the CNDMDS of approximately 46 000 individuals aged >20 years, self-reported IHD and stroke prevalences were <1%.¹¹³ In contrast, the rural Beijing study, of approximately 58 000 individuals aged >40 years, reported IHD and stroke prevalences of 5.6% and 3.7%, respectively,¹¹⁰ higher than estimates in the present study. Differences between these studies and the CKB probably reflect differences in disease definitions, sampling schemes and populations or temporal trends.

Investigation of composite IHD and cerebrovascular disease outcomes in the presented analyses clearly precludes understanding of the association of diabetes with disease subtypes, of particular relevance to the association with stroke. Moreover, since diabetes has been associated with a higher risk of fatal, than non-fatal, CVD⁶ and with higher CVD mortality rates,²⁵³ the investigation of non-fatal outcomes in these cross-sectional analyses, and potentially less severe forms, could underestimate the strength of association with diabetes. A substantial proportion of participants who reported a history of doctor-diagnosed IHD or stroke/TIA at baseline did not report a history of these conditions at resurvey, with particularly poor repeatability for IHD. Most other studies have not reported on the repeatability of self-reported CVD, but several have assessed the validity of self-reported IHD or stroke in Western populations. These have shown good agreement between self-reported diagnoses and diagnoses in medical records,²⁵⁴⁻²⁵⁷ with, for example, sensitivity >80%^{255, 257} and specificity >99%.^{255, 257} A prospective Japanese study found lower sensitivity of self-reported stroke (73%) than MI (82%), attributed to a higher incidence of stroke and more frequent detection of less severe stroke events (due to more extensive use of CT/MR scans) in Japan;²⁵⁸ this bears some resemblance to the situation in China. The repeatability of self-reported IHD and stroke/TIA were considerably lower than that of diabetes in the CKB, which may reflect the comparative complexity of diagnosing stroke and IHD, and the heterogeneity of these conditions.²⁵⁹ In the absence of a gold-standard data source for prior CVD, it is not possible to assess the validity of self-reported diagnoses in the CKB. Nevertheless, it is clearly important to consider the potential consequences of

misclassification—likely non-differential—of participants, which would be expected to result in underestimation of the strength of association of diabetes with prevalent IHD and stroke/TIA. Conversely, however, a known diagnosis of diabetes could bias self-reporting of CVDs resulting in over-estimation of diabetes-associated risks.

The limitations of the use of RPG levels²¹⁶ measured on a point-of-care glucometer,²³³ as discussed in Section 4.4.5, should also be considered when interpreting the presented findings. Furthermore, the cross-sectional nature of these analyses clearly limits the quality of the evidence provided, in particular due to the risk of reverse causality. It is possible, for example, that diabetes, if present, is more likely to be diagnosed among individuals with CVDs, as a result of more regular contact with healthcare services, than among those without CVDs. This effect may have contributed to the observed association of self-reported diabetes with prevalent IHD and stroke/TIA. The current lack of data on lipid levels for CKB participants may have led to some residual confounding.

5.4.4 Conclusion

The presented analyses provide clear evidence of an independently elevated prevalence of IHD and stroke/TIA among individuals with self-reported diabetes. They also provide supportive evidence of associations of RPG levels below the diabetic threshold with prevalent stroke/TIA and possibly with IHD. Multiple pathophysiological explanations for associations of diabetes and of blood glucose levels with CVD have been proposed,²⁴⁸ and a sound understanding of the role of both exposures in determining IHD and stroke risk is fundamental to effective prevention and control of CVD in Chinese, and other, populations. In light of the limitations of such cross-sectional analyses, however, investigation of the prospective associations of diabetes, and of plasma glucose levels in individuals without known diabetes, with incident CVDs, including IHD and stroke subtypes, is essential.

Figure 5.1. Baseline prevalence of cardiovascular diseases by age and sex

Standardised to study area and education structure of study population.

Squares represent ischaemic heart disease or stroke/TIA prevalence and vertical lines represent corresponding 95% confidence intervals. TIA, transient ischaemic attack.

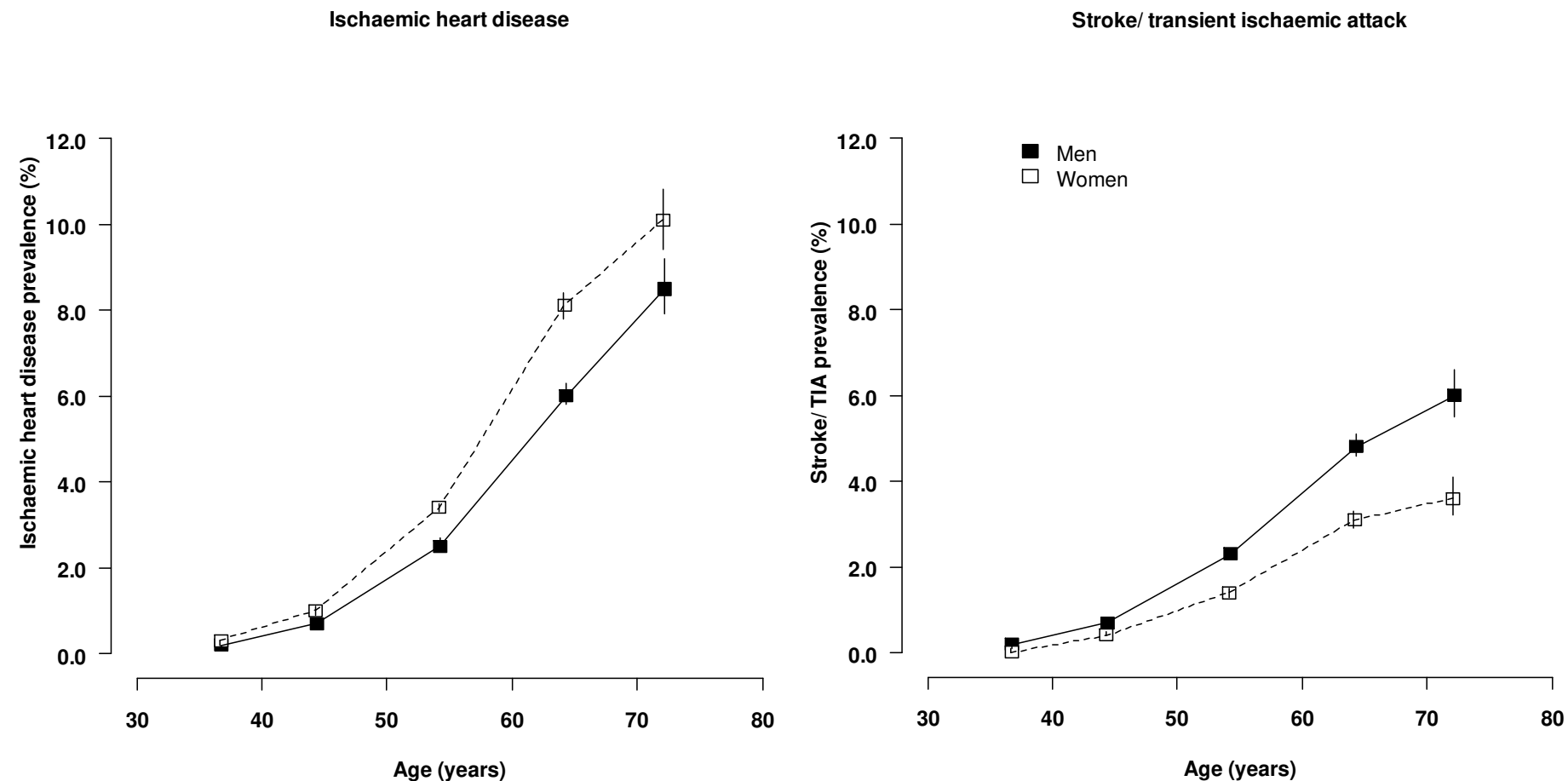


Figure 5.2. Baseline prevalence of ischaemic heart disease by study area

Standardised to age and education structure of study population.

Error bars represent 95% confidence intervals.

R, rural; U, urban.

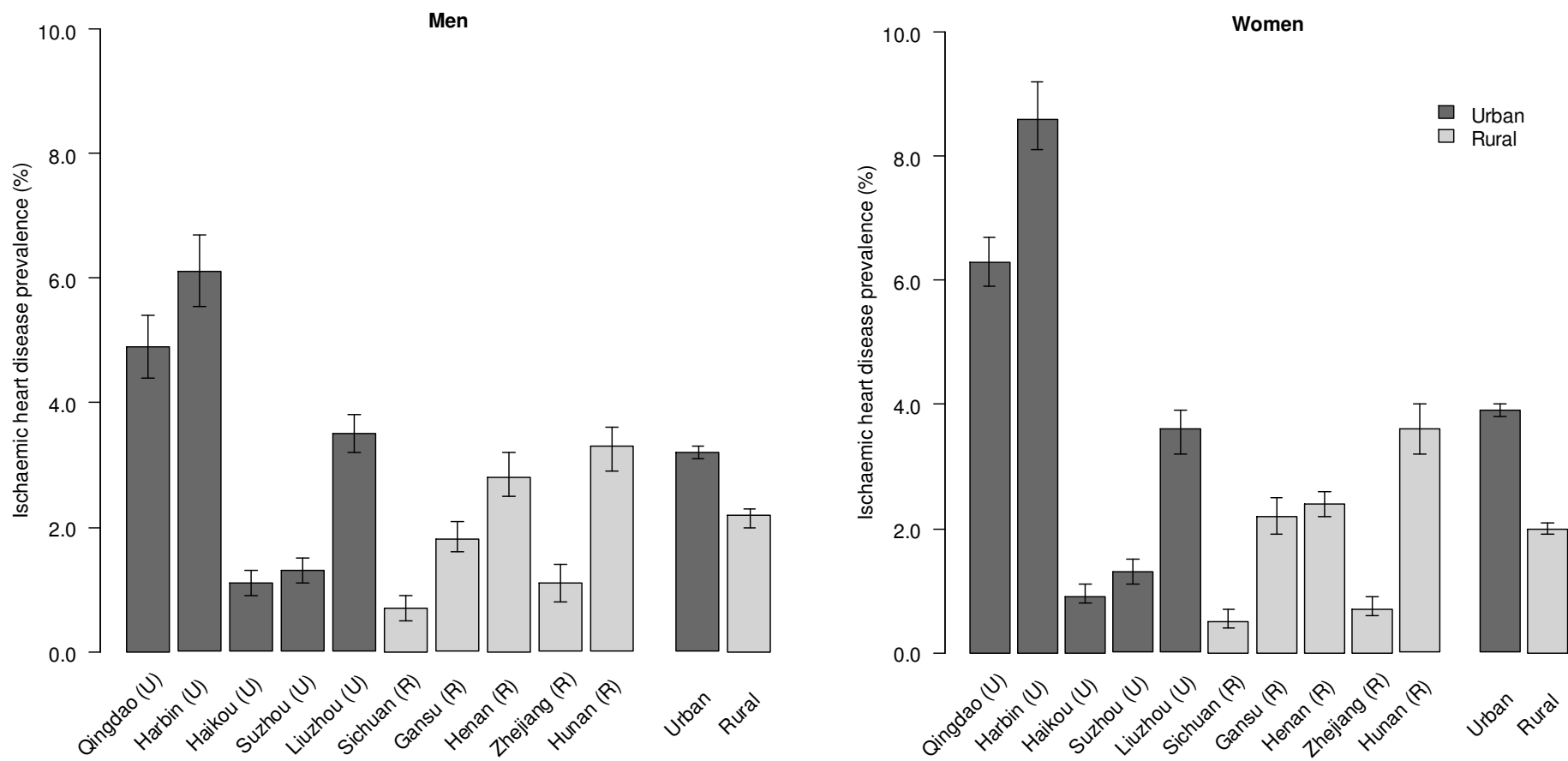


Figure 5.3. Baseline prevalence of stroke/transient ischaemic attack by study area

Standardised to age and education structure of study population.

Error bars represent 95% confidence intervals.

R, rural; TIA, transient ischaemic attack; U, urban.

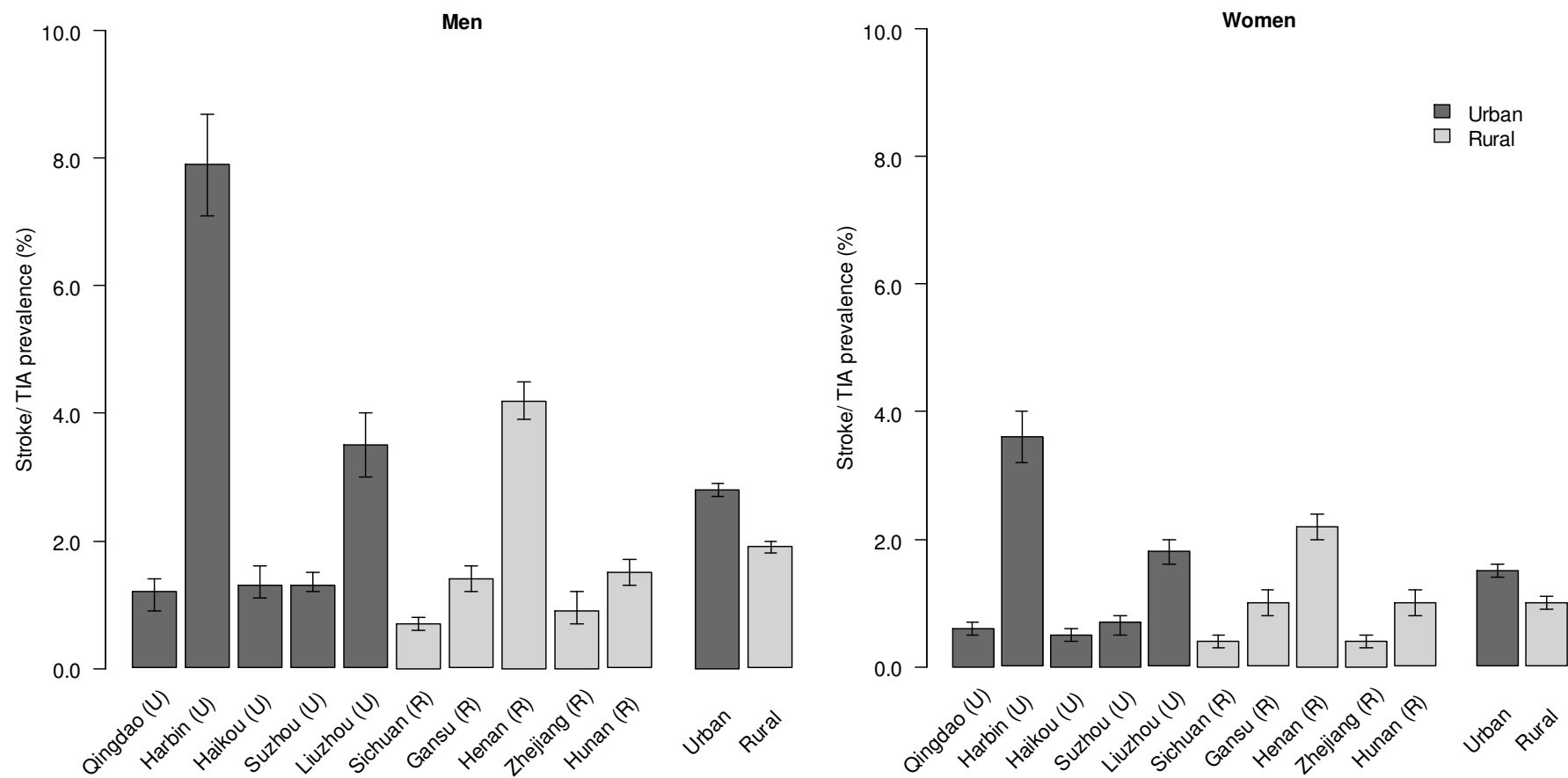


Table 5.1. Self-reported ischaemic heart disease at baseline and resurvey among participants included in both first and second resurveys (n=14 854)

Baseline	First resurvey				Second resurvey			
	No IHD		IHD		No IHD		IHD	
	n	%	n	%	n	%	N	%
No IHD (n=14 435)	14 234	98.6	201	1.4	13 764	95.4	671	4.7
IHD (n=419)	215	51.3	204	48.7	224	53.5	195	46.5

Kappa for self-reported IHD at baseline and resurveys excluding participants with self-reported IHD at resurvey who did not have self-reported IHD at baseline: first resurvey 0.65 (95% CI 0.60-0.69); second resurvey 0.63 (95% CI 0.58-0.67).

CI; confidence interval; IHD, ischaemic heart disease.

Table 5.2. Self-reported stroke/transient ischaemic attack at baseline and resurvey among participants included in both first and second resurveys (n=14 854)

Baseline	First resurvey				Second resurvey			
	No stroke/TIA		Stroke/TIA		No stroke/TIA		Stroke/TIA	
	n	%	n	%	n	%	n	%
No stroke/TIA (n= 14 700)	14 587	99.2	113	0.8	14 282	97.2	418	2.8
Stroke/TIA (n=154)	62	40.3	92	59.7	68	44.2	86	55.8

Kappa for self-reported stroke/TIA at baseline and resurveys excluding participants with self-reported stroke/TIA at resurvey who did not have self-reported stroke/TIA at baseline: first resurvey 0.75 (95% CI 0.69-0.81); second resurvey 0.71 (95% CI 0.65-0.78).
CI; confidence interval; TIA, transient ischaemic attack.

Table 5.3. Adjusted odds ratios for prevalent cardiovascular diseases by self-reported diabetes status in men and women

	Self-reported diabetes		No self-reported diabetes		Model A ^a		Model B ^b		Model C ^c	
	Events	Participants	Events	Participants	OR	95% CI	OR	95% CI	OR	95% CI
Ischaemic heart disease										
<i>Men</i>										
30-59 years	182	3215	1712	145672	3.77	3.21-4.42	2.29	1.93-2.70	2.12	1.79-2.51
60-69 years	297	1950	1972	37895	2.46	2.14-2.82	2.06	1.79-2.36	1.92	1.67-2.21
70-79 years	202	948	1129	14687	2.43	2.04-2.89	2.13	1.78-2.54	2.02	1.68-2.41
Overall	681	6113	4813	198254	2.57	2.35-2.82	2.13	1.94-2.34	1.97	1.80-2.17
<i>Women</i>										
30-59 years	383	4897	3494	224886	4.51	4.03-5.05	2.82	2.51-3.17	2.44	2.17-2.74
60-69 years	678	3778	3160	44071	2.12	1.93-2.34	1.98	1.79-2.18	1.87	1.70-2.07
70-79 years	304	1337	1258	14392	2.17	1.87-2.52	2.02	1.73-2.35	1.94	1.67-2.27
Overall	1365	10012	7912	283349	2.39	2.24-2.56	2.18	2.03-2.33	2.01	1.88-2.15
<i>Men and women</i>										
30-59 years	565	8112	5206	370558	4.20	3.83-4.60	2.59	2.36-2.85	2.30	2.09-2.53
60-69 years	975	5728	5132	81966	2.21	2.05-2.40	1.99	1.84-2.15	1.88	1.74-2.04
70-79 years	506	2285	2387	29079	2.28	2.04-2.55	2.08	1.85-2.34	2.00	1.78-2.25
Overall	2046	16125	12725	481603	2.45	2.32-2.59	2.16	2.05-2.28	2.00	1.89-2.11
Stroke/ transient ischaemic attack										
<i>Men</i>										
30-59 years	141	3215	1614	145672	3.12	2.61-3.74	1.44	1.19-1.75	1.38	1.13-1.67
60-69 years	229	1950	1726	37895	2.29	1.96-2.66	1.68	1.44-1.97	1.61	1.37-1.89
70-79 years	135	948	859	14687	2.17	1.77-2.66	1.77	1.44-2.18	1.68	1.38-2.04
Overall	505	6113	4199	198254	2.34	2.11-2.59	1.64	1.47-1.82	1.54	1.39-1.72
<i>Women</i>										
30-59 years	187	4897	1377	224886	5.67	4.84-6.64	2.72	2.31-3.21	2.42	2.05-2.86
60-69 years	292	3778	1225	44071	2.43	2.12-2.78	1.96	1.70-2.25	1.85	1.61-2.14
70-79 years	121	1337	524	14392	2.07	1.68-2.57	1.76	1.41-2.18	1.64	1.32-2.04
Overall	600	10012	3126	283349	2.67	2.43-2.94	2.09	1.90-2.31	1.94	1.76-2.13
<i>Men and women</i>										
30-59 years	328	8112	2991	370558	4.26	3.79-4.80	2.02	1.78-2.29	1.87	1.64-2.12
60-69 years	521	5728	2951	81966	2.36	2.14-2.62	1.82	1.64-2.02	1.73	1.56-1.93
70-79 years	256	2285	1383	29079	2.15	1.85-2.48	1.79	1.54-2.08	1.66	1.40-1.98
Overall	1105	16125	7325	481603	2.52	2.36-2.71	1.88	1.75-2.02	1.75	1.63-1.88

^aModel A: adjusted for age, sex (where appropriate) and study area; ^bModel B: additionally adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure; ^cModel C: additionally adjusted for waist-to-hip ratio.

CI, confidence interval; OR, odds ratio.

Figure 5.4. Odds ratios for prevalent cardiovascular diseases by self-reported diabetes status, progressively adjusted for potential confounding factors

Squares represent the odds ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs).

BMI, body mass index; DBP, diastolic blood pressure; SBP, systolic blood pressure.

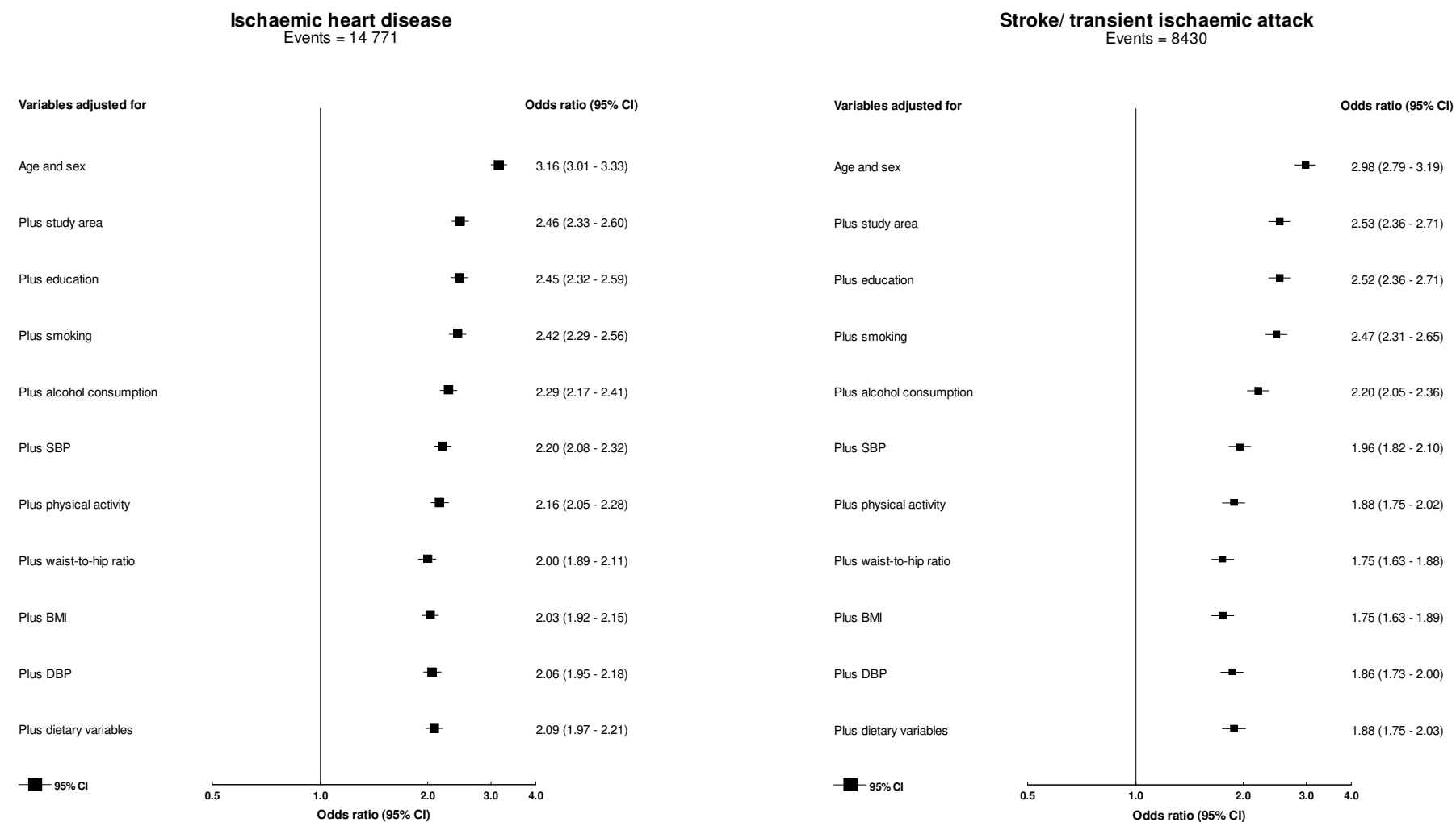


Figure 5.5. Adjusted odds ratios for prevalent cardiovascular diseases by self-reported diabetes status, by study area

Adjusted for age, sex, education, smoking, alcohol consumption, physical activity, systolic blood pressure. Shaded squares represent the odds ratio (OR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall OR. The open diamond represents the overall OR and its 95% CI.

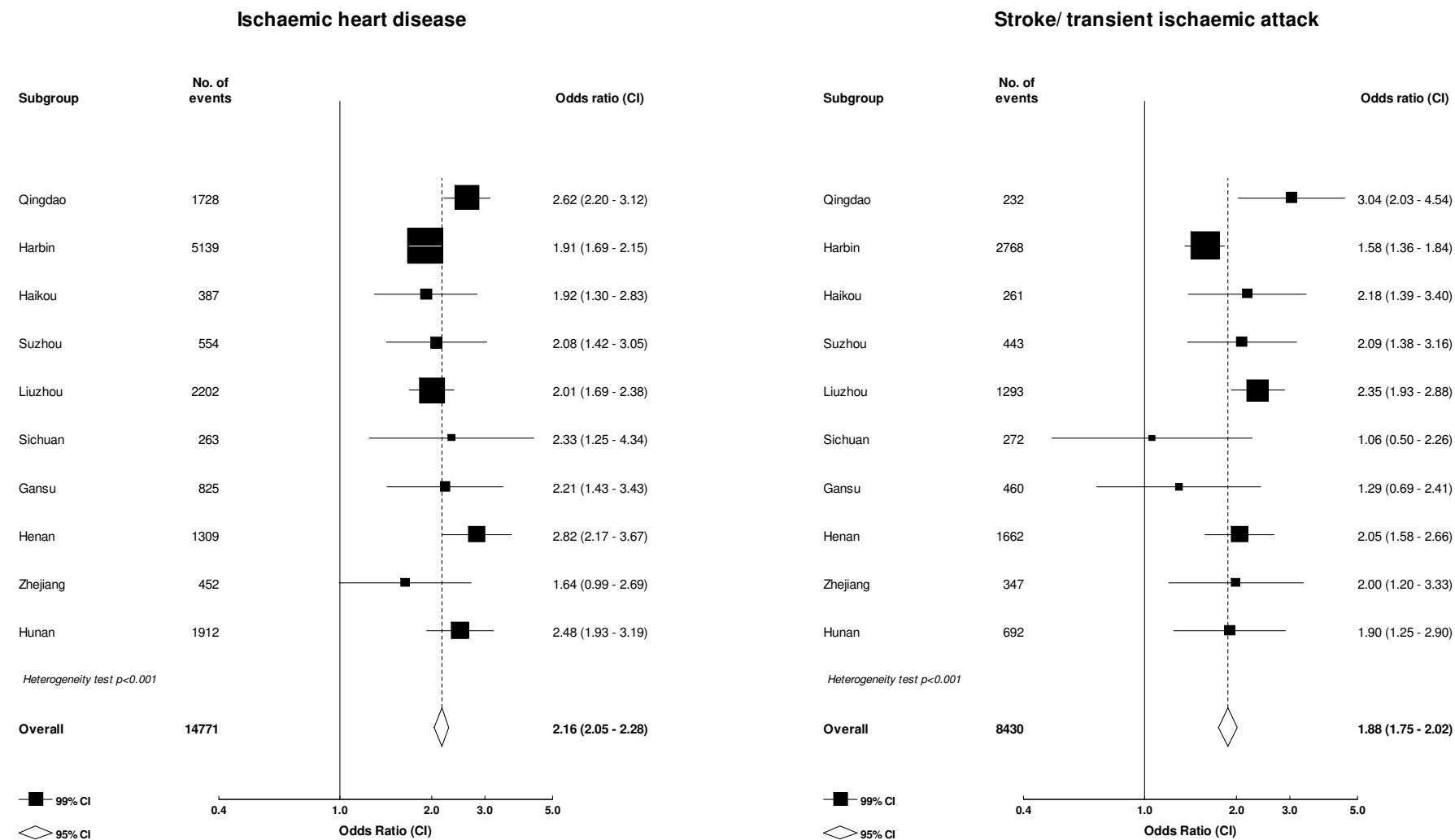


Figure 5.6. Adjusted odds ratios for prevalent ischaemic heart disease by self-reported diabetes status in population subgroups

Adjusted for age, sex, study area, education, smoking, alcohol consumption, physical activity, systolic blood pressure (except where the variable of interest). Shaded squares represent the odds ratios (OR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall OR. The open diamond represents the overall OR and its 95% CI.

BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

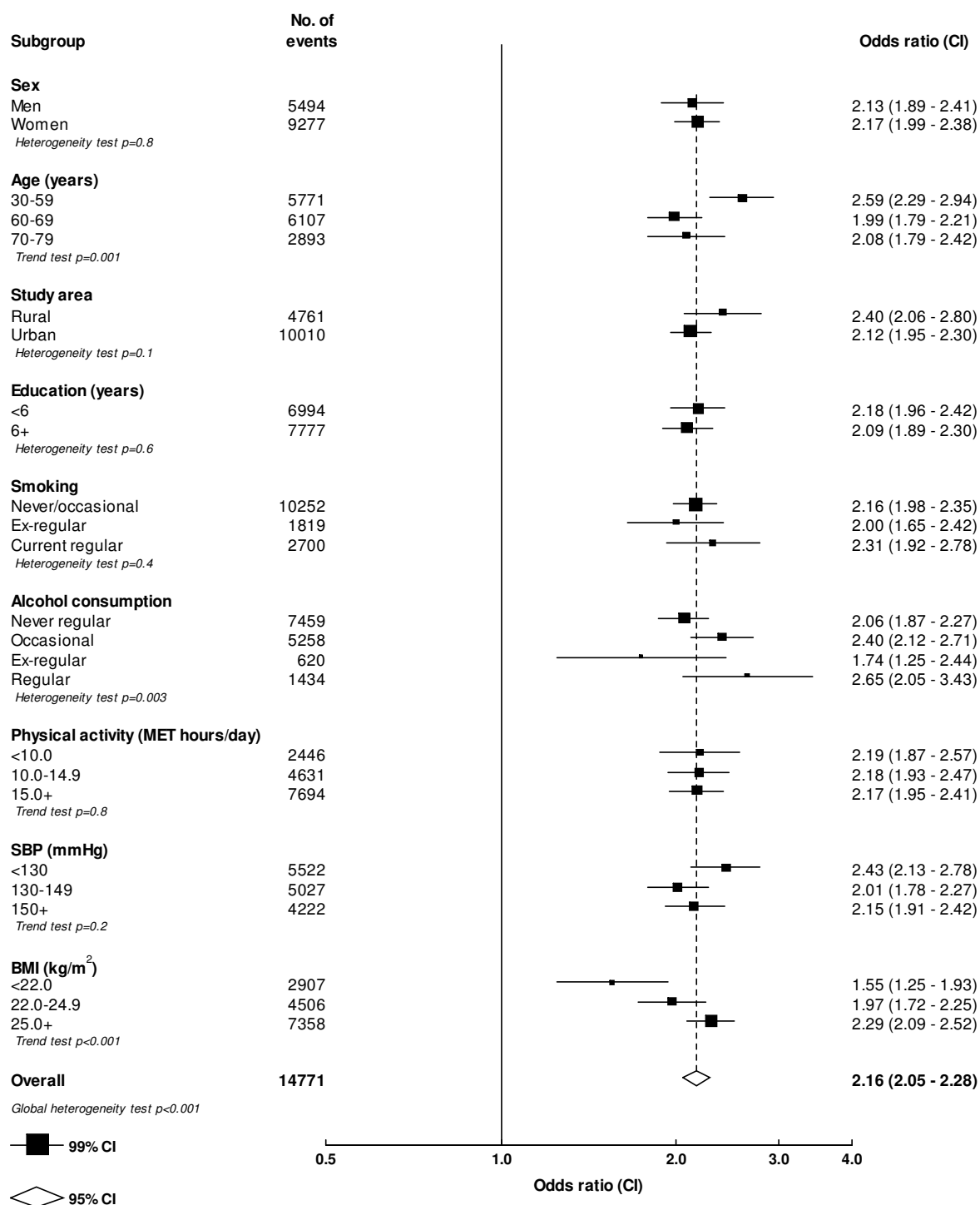


Figure 5.7. Adjusted odds ratios for prevalent stroke/transient ischaemic attack by self-reported diabetes status in population subgroups

Adjusted for age, sex, study area, education, smoking, alcohol consumption, physical activity, systolic blood pressure (except where the variable of interest). Shaded squares represent the odds ratio (OR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall OR. The open diamond represents the overall OR and its 95% CI.

BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

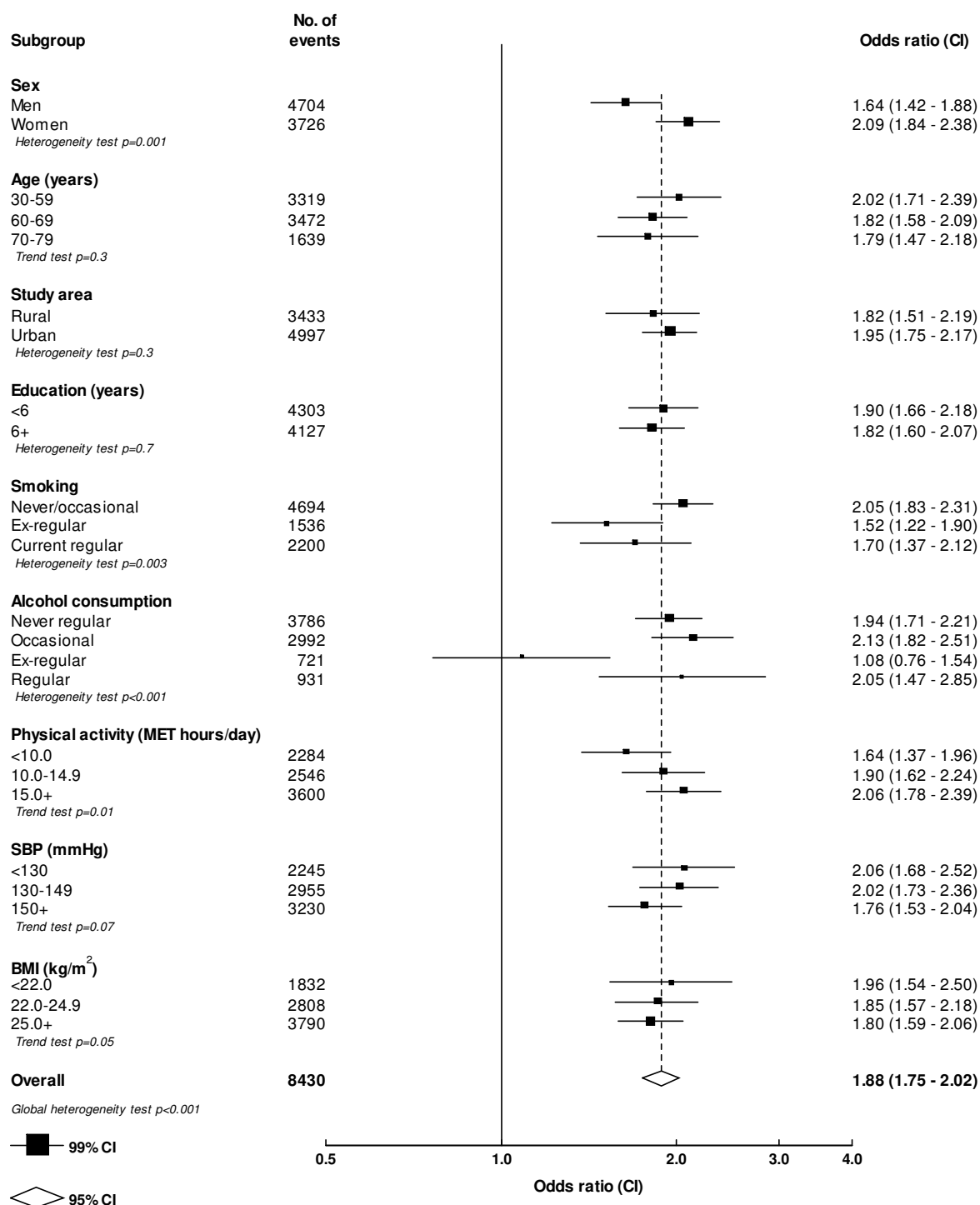


Figure 5.8. Adjusted odds ratios for prevalent cardiovascular diseases by duration of diabetes diagnosis

Adjusted for age, sex, study area, education, smoking, alcohol consumption, physical activity, systolic blood pressure. Squares represent the odds ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs).

Reference group: individuals without diabetes. Heterogeneity tests based on individuals with self-reported diabetes.

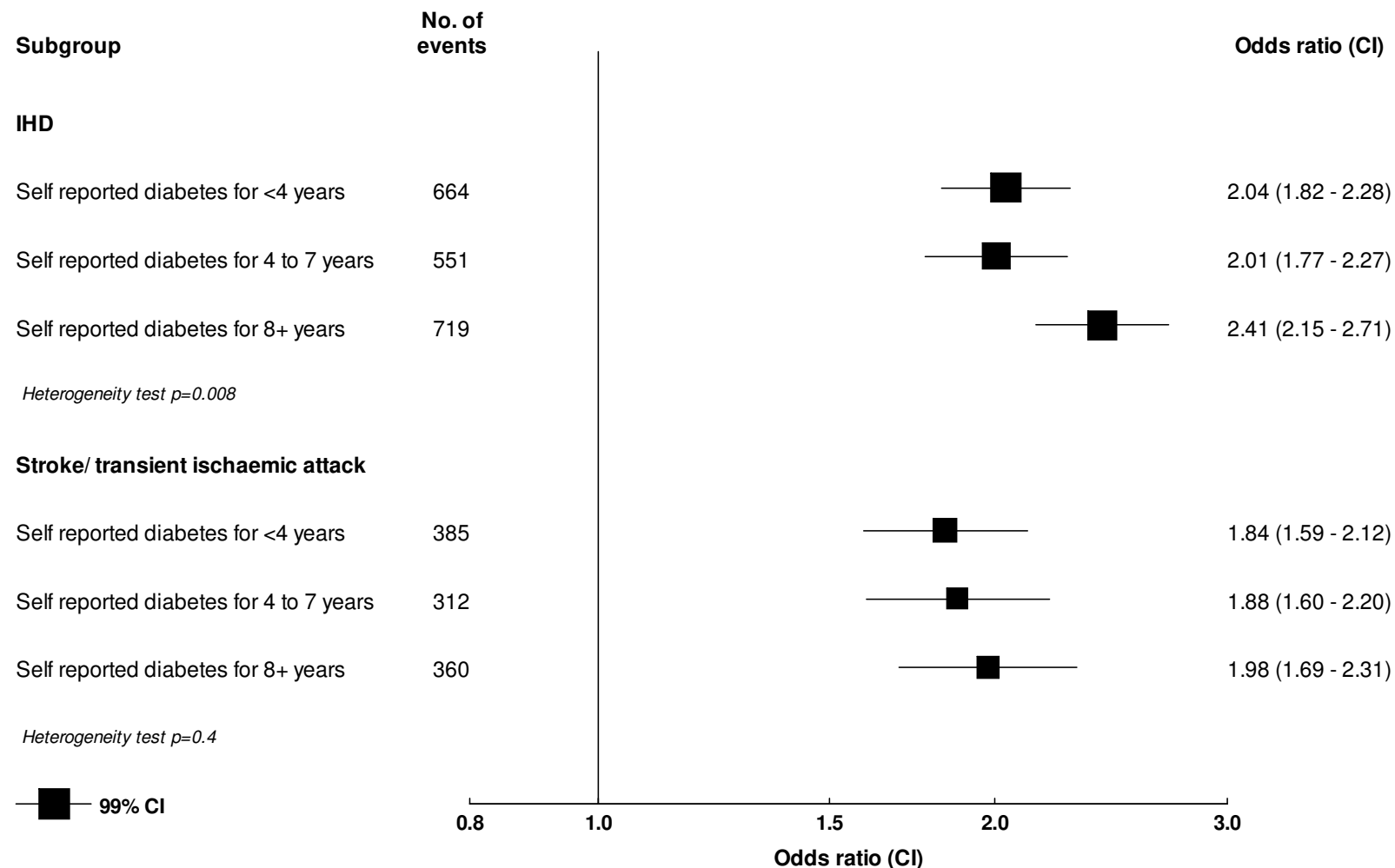


Figure 5.9. Adjusted odds ratios for prevalent cardiovascular diseases by use of hypoglycaemic medication in self-reported diabetes

Adjusted for age, sex, study area, education, smoking, alcohol consumption, physical activity, systolic blood pressure. Squares represent the odds ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs).

Reference group: individuals without diabetes. Heterogeneity tests based on individuals with self-reported diabetes.

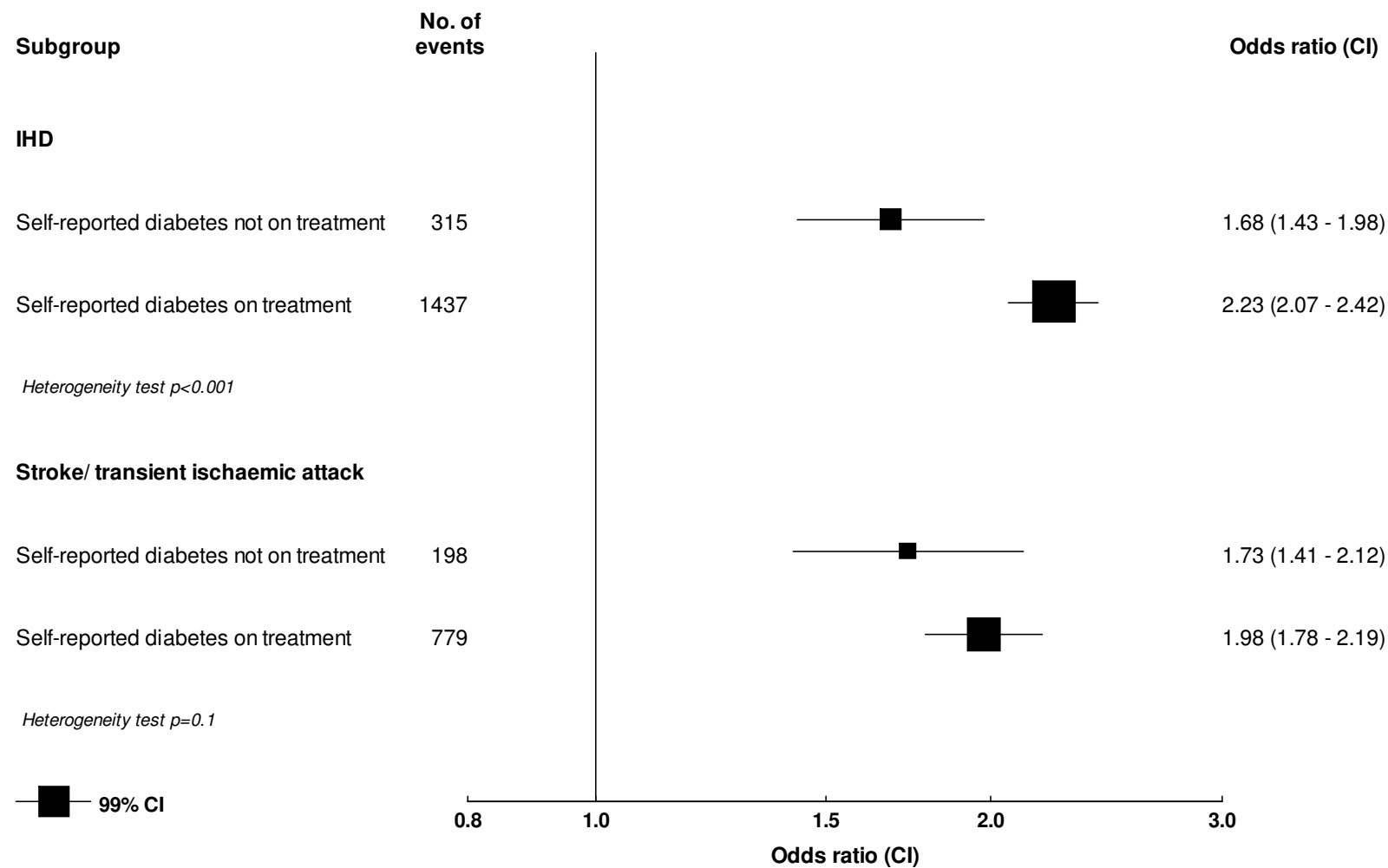


Table 5.4. Adjusted odds ratios for prevalent cardiovascular diseases by screen-detected diabetes status in men and women

	Screen-detected diabetes		No screen-detected diabetes		Model A ^a		Model B ^b		Model C ^c	
	Events	Participants	Events	Participants	OR	95% CI	OR	95% CI	OR	95% CI
Ischaemic heart disease										
<i>Men</i>										
30-59 years	71	3474	1712	145672	0.69	0.54-0.88	0.98	0.76-1.25	0.99	0.76-1.27
60-69 years	83	1400	1972	37895	1.02	0.81-1.29	1.11	0.88-1.40	1.18	0.93-1.49
70-79 years	48	663	1129	14687	1.33	0.98-1.80	1.35	0.99-1.83	1.41	1.04-1.93
Overall	202	5537	4813	198254	1.01	0.87-1.17	1.09	0.94-1.27	1.17	1.01-1.36
<i>Women</i>										
30-59 years	148	5359	3494	224886	0.59	0.50-0.70	0.97	0.81-1.15	1.00	0.84-1.19
60-69 years	201	2374	3160	44071	1.02	0.88-1.19	1.07	0.91-1.25	1.13	0.97-1.32
70-79 years	90	812	1258	14392	0.99	0.78-1.25	1.00	0.79-1.27	1.04	0.82-1.32
Overall	439	8545	7912	283349	0.94	0.85-1.05	1.00	0.90-1.11	1.07	0.97-1.19
<i>Men and women</i>										
30-59 years	219	8833	5206	370558	0.63	0.55-0.72	0.88	0.76-1.01	1.00	0.87-1.15
60-69 years	284	3774	5132	81966	1.02	0.90-1.16	1.08	0.95-1.22	1.14	1.00-1.29
70-79 years	138	1475	2387	29079	1.11	0.92-1.33	1.13	0.93-1.36	1.17	0.97-1.41
Overall	641	14082	12725	481603	0.97	0.89-1.05	1.02	0.94-1.12	1.10	1.01-1.20
Stroke/ transient ischaemic attack										
<i>Men</i>										
30-59 years	70	3474	1614	145672	0.64	0.50-0.82	1.09	0.84-1.41	1.17	0.90-1.51
60-69 years	68	1400	1726	37895	1.02	0.80-1.32	1.20	0.92-1.55	1.25	0.97-1.62
70-79 years	60	663	859	14687	0.70	0.53-0.93	0.78	0.59-1.04	0.83	0.62-1.10
Overall	198	5537	4199	198254	0.83	0.72-0.97	1.01	0.86-1.17	1.07	0.92-1.25
<i>Women</i>										
30-59 years	91	5359	1377	224886	0.37	0.30-0.46	0.70	0.57-0.88	0.76	0.61-0.95
60-69 years	111	2374	1225	44071	0.65	0.53-0.79	0.77	0.62-0.94	0.81	0.66-1.00
70-79 years	30	812	524	14392	1.14	0.78-1.67	1.29	0.88-1.89	1.37	0.93-2.01
Overall	232	8545	3126	283349	0.65	0.57-0.75	0.79	0.69-0.91	0.85	0.74-0.98
<i>Men and women</i>										
30-59 years	161	8833	2991	370558	0.48	0.41-0.57	0.80	0.68-0.95	0.88	0.74-1.04
60-69 years	179	3774	2951	81966	0.79	0.67-0.92	0.92	0.78-1.08	0.97	0.83-1.14
70-79 years	90	1475	1383	29079	0.85	0.68-1.06	0.95	0.76-1.19	1.00	0.80-1.26
Overall	430	14082	7325	481603	0.73	0.66-0.81	0.88	0.80-0.98	0.95	0.85-1.05

^aModel A: adjusted for age, sex (where appropriate) and study area; ^bModel B: additionally adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure; ^cModel C: additionally adjusted for waist-to-hip ratio.

CI, confidence interval; OR, odds ratio.

Figure 5.10. Adjusted odds ratios for prevalent cardiovascular diseases by baseline random plasma glucose levels in participants without self-reported diabetes

Adjusted for age, sex, study area, education, smoking, alcohol consumption, physical activity, systolic blood pressure.

Squares represent the odds ratios (OR) with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs). Numbers above the CIs are the ORs and numbers below the CIs are of ischaemic heart disease or stroke/transient ischaemic attack events.

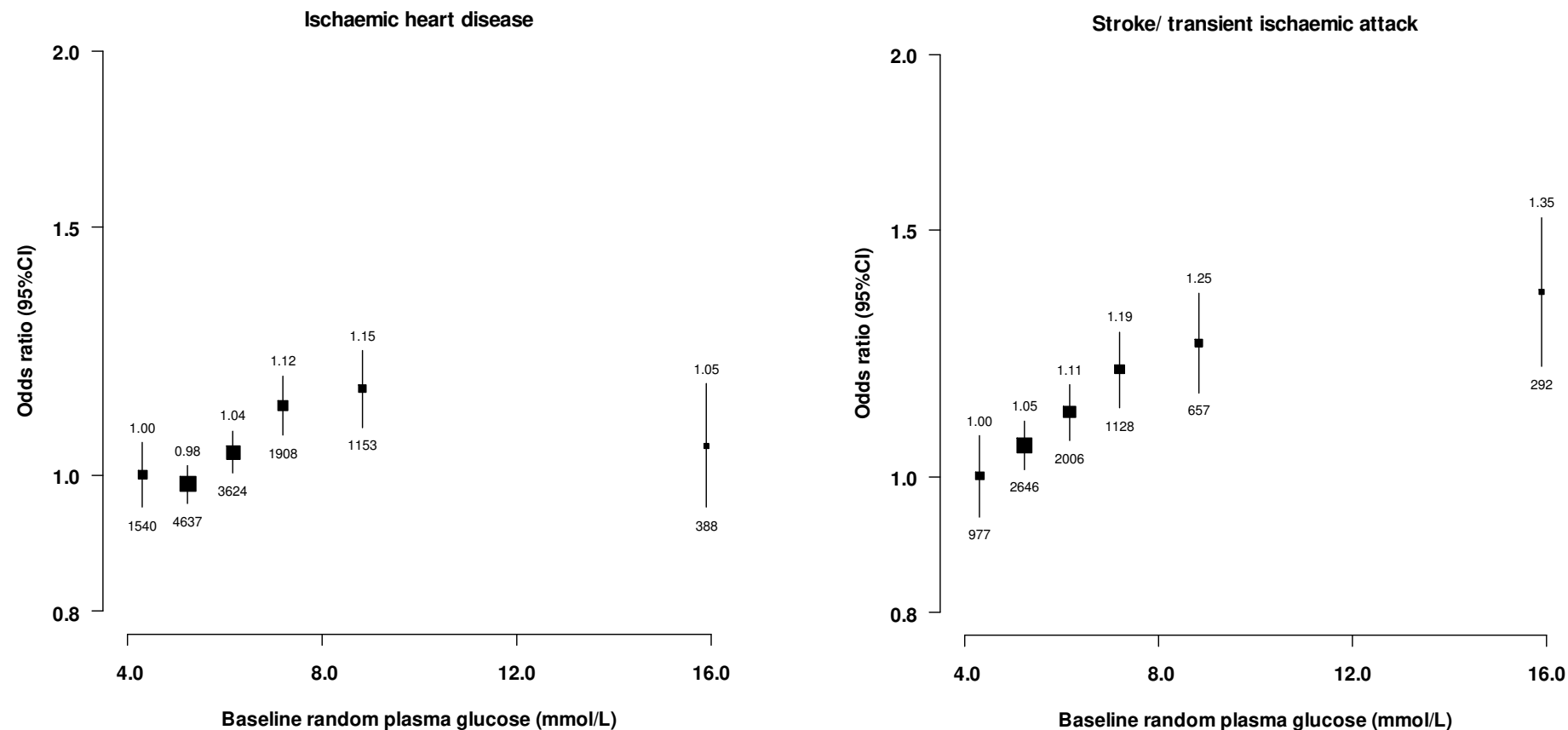


Table 5.5. Adjusted odds ratios for prevalent cardiovascular diseases per 1 mmol/L higher baseline random plasma glucose at random plasma glucose levels <11.1 mmol/L in individuals without self-reported diabetes

	Events	Participants	Model A ^a		Model B ^b		Model C ^c	
			OR	95% CI	OR	95% CI	OR	95% CI
Ischaemic heart disease								
Men								
30-59 years	1724	144231	1.14	1.10-1.18	1.07	1.03-1.12	1.05	1.01-1.10
60-69 years	1980	37725	1.05	1.01-1.09	1.03	1.00-1.07	1.02	0.98-1.06
70-79 years	1144	14762	1.03	0.99-1.08	1.03	0.98-1.08	1.02	0.97-1.07
Overall	4848	196718	1.05	1.02-1.07	1.03	1.01-1.06	1.02	0.99-1.04
Women								
30-59 years	3522	223164	1.16	1.13-1.19	1.08	1.05-1.12	1.05	1.02-1.08
60-69 years	3202	44305	1.04	1.01-1.07	1.03	1.00-1.06	1.01	0.98-1.04
70-79 years	1290	14565	1.08	1.04-1.13	1.08	1.03-1.13	1.07	1.02-1.12
Overall	8014	282034	1.05	1.03-1.07	1.04	1.02-1.06	1.02	1.00-1.04
Men and women								
30-59 years	5246	367395	1.15	1.12-1.18	1.08	1.06-1.10	1.05	1.03-1.07
60-69 years	5182	82030	1.04	1.02-1.06	1.03	1.01-1.05	1.02	0.99-1.04
70-79 years	2434	29327	1.06	1.03-1.09	1.05	1.02-1.09	1.05	1.01-1.08
Overall	12862	478752	1.05	1.03-1.06	1.04	1.02-1.05	1.02	1.00-1.03
Stroke/ transient ischaemic attack								
Men								
30-59 years	1618	144231	1.17	1.13-1.22	1.02	0.98-1.06	1.05	1.01-1.09
60-69 years	1734	37725	1.06	1.03-1.11	1.04	1.00-1.08	1.03	0.99-1.07
70-79 years	872	14762	1.07	1.02-1.13	1.05	1.00-1.11	1.04	0.99-1.10
Overall	4224	196718	1.07	1.05-1.10	1.04	1.02-1.07	1.03	1.00-1.05
Women								
30-59 years	1402	223164	1.21	1.16-1.27	1.10	1.05-1.15	1.07	1.02-1.12
60-69 years	1257	44305	1.09	1.04-1.14	1.07	1.02-1.12	1.06	1.01-1.10
70-79 years	531	14565	1.05	0.99-1.13	1.05	0.98-1.12	1.03	0.97-1.11
Overall	3190	282034	1.09	1.06-1.12	1.06	1.03-1.09	1.05	1.02-1.08
Men and women								
30-59 years	3020	367395	1.19	1.16-1.22	1.04	1.01-1.07	1.06	1.03-1.09
60-69 years	2991	82030	1.07	1.04-1.10	1.05	1.02-1.08	1.04	1.01-1.07
70-79 years	1403	29327	1.07	1.02-1.11	1.05	1.01-1.09	1.04	0.99-1.08
Overall	7414	478752	1.08	1.06-1.10	1.05	1.03-1.07	1.04	1.02-1.05

^aModel A: adjusted for age, sex (where appropriate) and study area; ^bModel B: additionally adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure; ^cModel C: additionally adjusted for waist-to-hip ratio.

CI, confidence interval; OR, odds ratio.

Figure 5.11. Adjusted odds ratios for prevalent cardiovascular diseases per 1 mmol/L higher baseline random plasma glucose at random plasma glucose levels <11.1 mmol/L in individuals without self-reported diabetes, progressively adjusted for potential confounding factors

Squares represent the odds ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs).

BMI, body mass index; DBP, diastolic blood pressure; SBP, systolic blood pressure.

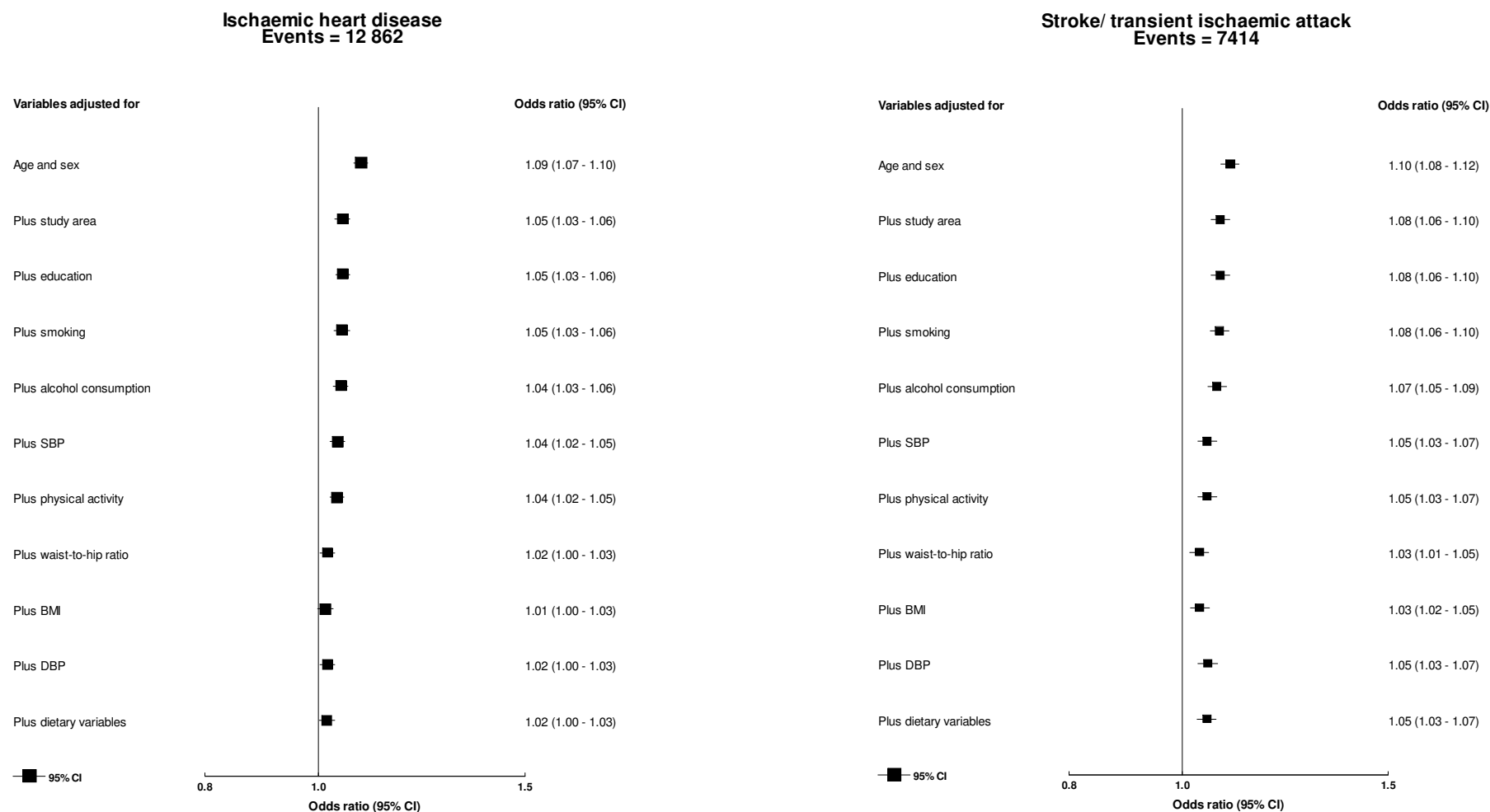


Figure 5.12. Adjusted odds ratios for prevalent cardiovascular diseases per 1 mmol/L higher baseline random plasma glucose at random plasma glucose levels <11.1 mmol/L in individuals without self-reported diabetes, by study area

Adjusted for age, sex, education, smoking, alcohol consumption, physical activity, systolic blood pressure. Shaded squares represent the odds ratio (OR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall OR. The open diamond represents the overall OR and its 95% CI.

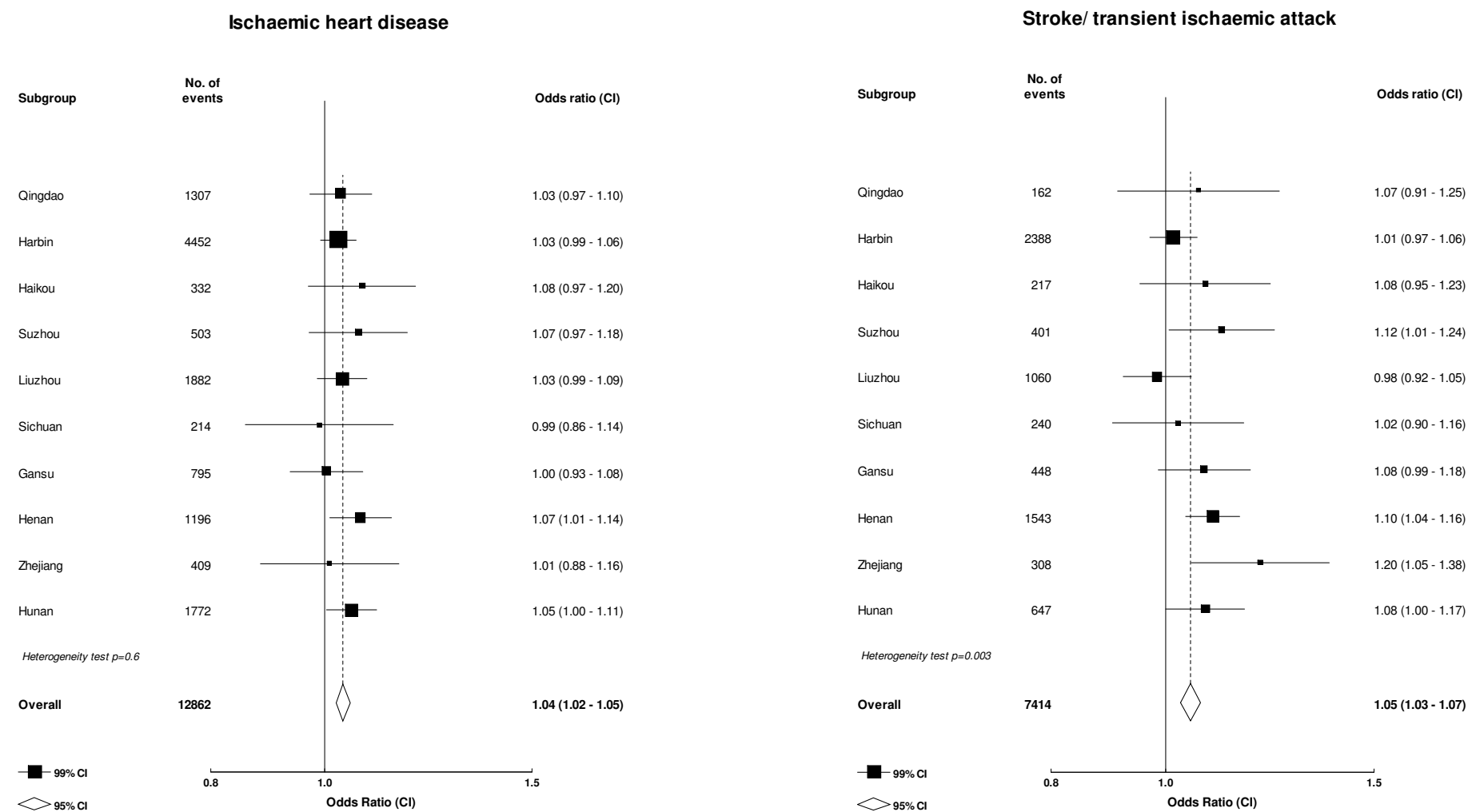


Figure 5.13. Adjusted odds ratios for prevalent ischaemic heart disease per 1mmol/L higher baseline random plasma glucose at random plasma glucose levels <11.1 mmol/L in population subgroups among participants without self-reported diabetes

Adjusted for age, sex, study area, education, smoking, alcohol consumption, physical activity, systolic blood pressure (except where the variable of interest). Shaded squares represent the odds ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall OR. The open diamond represents the overall OR and its 95% CI.

BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

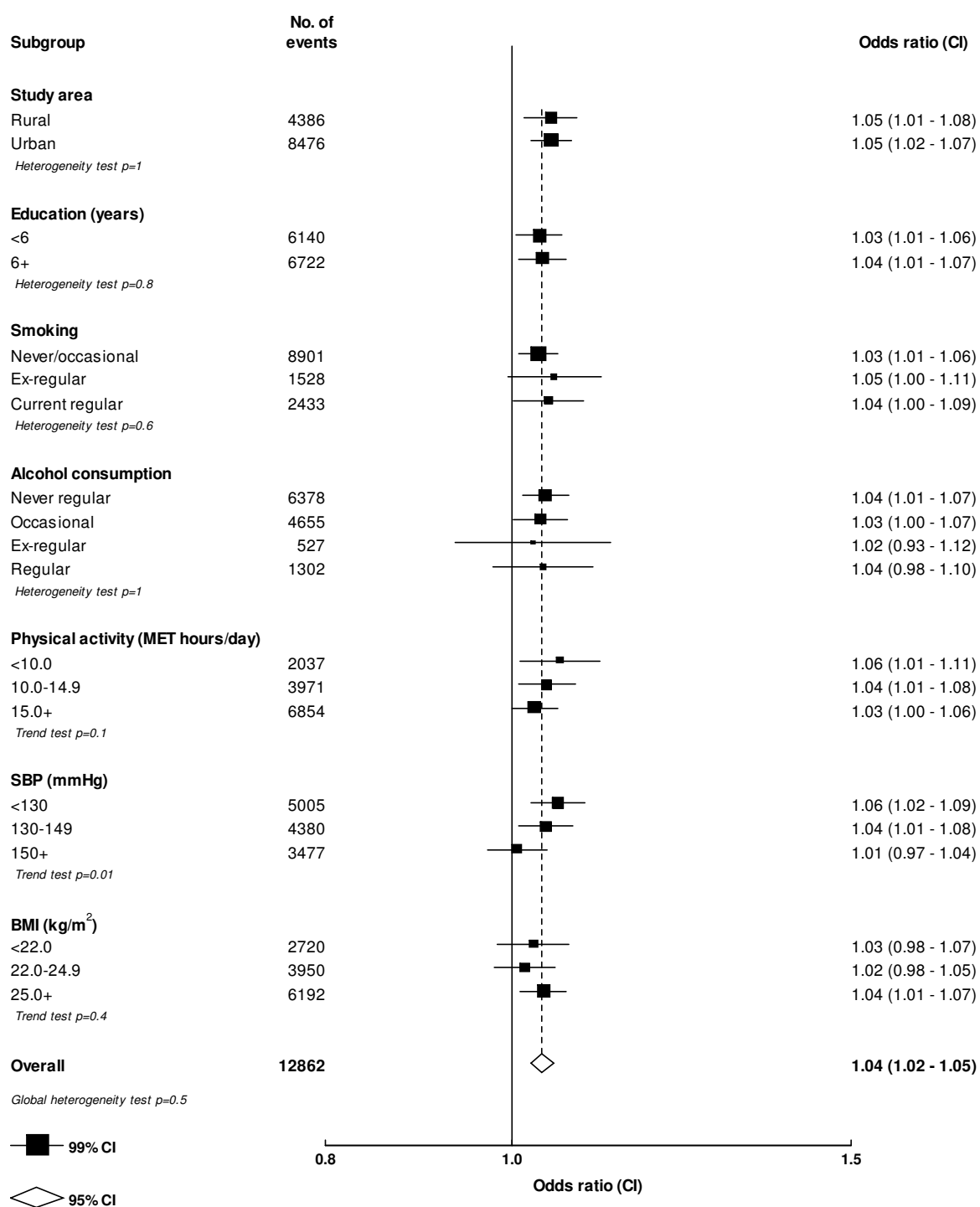


Figure 5.14. Adjusted odds ratios for prevalent stroke/ transient ischaemic attack per 1mmol/L higher baseline random plasma glucose at random plasma glucose levels <11.1 mmol/L in population subgroups among participants without self-reported diabetes

Adjusted for age, sex, study area, education, smoking, alcohol consumption, physical activity, systolic blood pressure (except where the variable of interest). Shaded squares represent the odds ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall OR. The open diamond represents the overall OR and its 95% CI.

BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

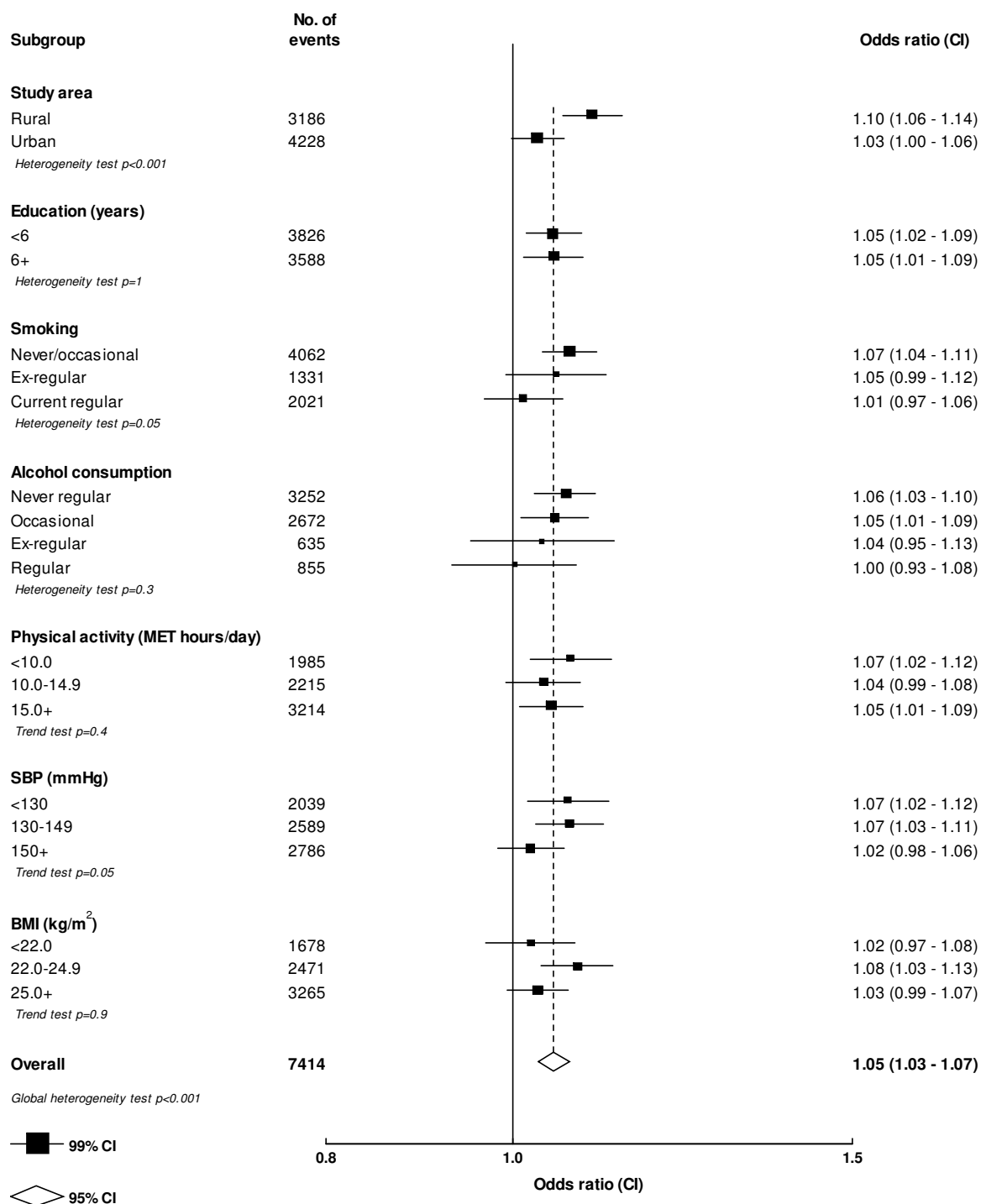
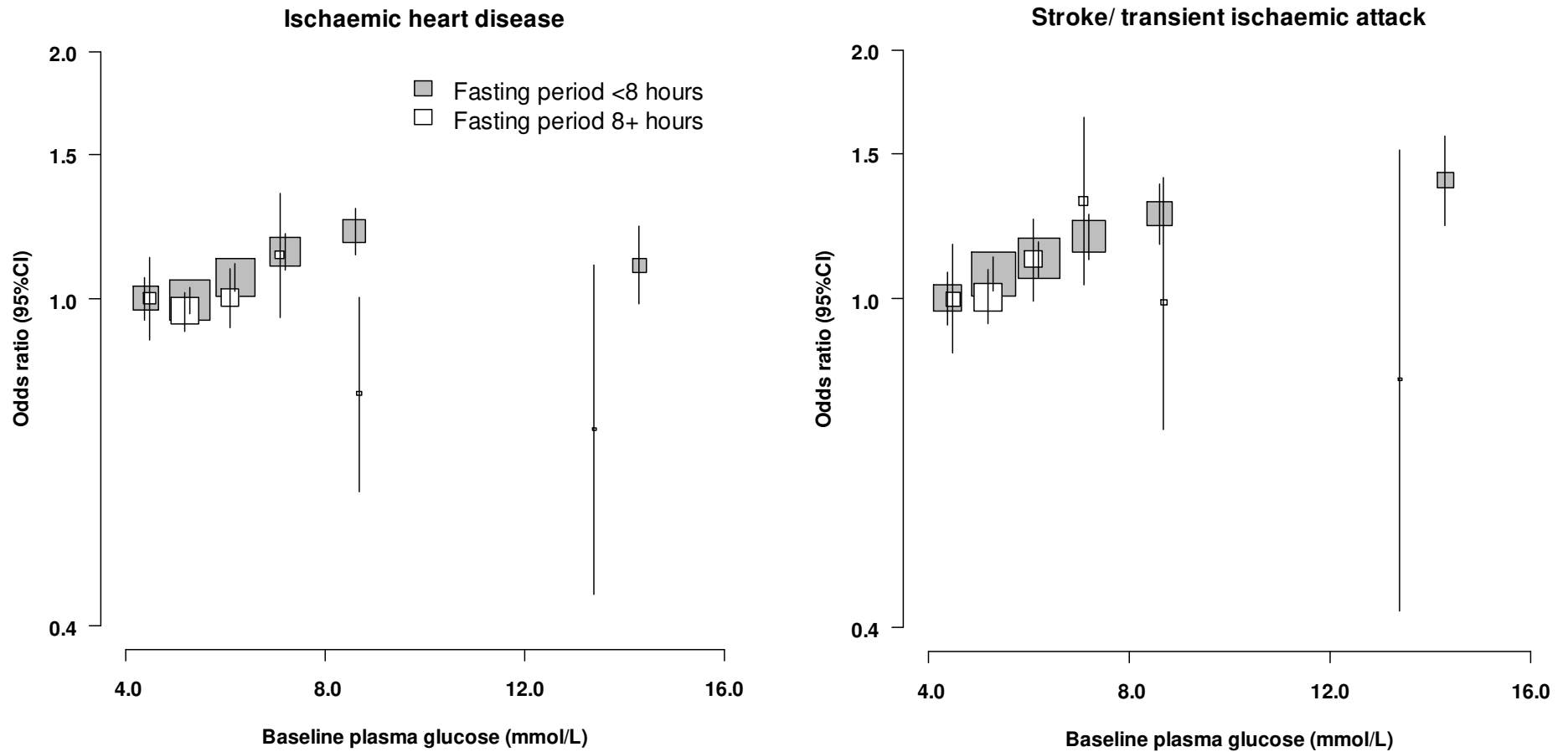


Figure 5.15. Adjusted odds ratios for prevalent cardiovascular diseases by baseline random plasma glucose levels stratified by fasting time in participants without self-reported diabetes

Adjusted for age, sex, study area, education, smoking, alcohol consumption, physical activity and systolic blood pressure.

Squares represent the odds ratios with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).



Chapter 6. Diabetes and the risk of incident cardiovascular disease

6.1 Background

Worldwide, approximately 400 million people have diabetes, and the prevalence is increasing rapidly, including in China.^{65, 86} As described in Chapter 2, diabetes has been shown to be associated with an approximately two-fold increased risk of IHD in Western populations.^{6, 90, 127-129, 133} Compared with Western populations, the classic picture of type 2 diabetes in China has been one of leaner individuals with low pancreatic beta-cell function and, compared with Western populations, greater susceptibility to microvascular diseases and cancer than macrovascular complications.⁸⁸ Despite this, similar diabetes-associated IHD risks have been reported in the limited number of Chinese studies.¹³⁰⁻¹³² However, the relatively small size,¹³⁰⁻¹³² potentially out-dated risk estimates¹³² or highly selected study populations¹³⁰ limits the value of these studies. The association of diabetes with stroke, and particularly stroke subtypes, is less well established across all populations.^{6, 90, 127, 129-131, 134, 136, 138} This is particularly true for HS; few studies have included adequate numbers of accurately phenotyped HS events to enable robust characterisation of the association with diabetes, and findings to-date have been conflicting.^{6, 90, 129} Associations of diabetes with stroke subtypes are, however, of considerable relevance to China, where stroke, particularly HS, rates are high.⁵⁵

Since the 1980s, when diabetes was rare in China,⁸⁵ there has been a rapid and substantial increase in prevalence, with ~10% of Chinese adults now estimated to have diabetes.^{65, 86, 87} This growing epidemic further highlights the importance of a robust understanding of the association of diabetes with CVDs. Furthermore, other factors, such as adiposity, blood pressure and type and level of physical activity, frequently differ between China and the West, but the relevance of these for the association of diabetes with CVDs is currently unclear in Chinese populations. Whilst the cross-sectional analyses presented in Chapter

5 provide some insight into the association of diabetes with CVDs in Chinese populations, their cross-sectional design, investigation only of self-reported, non-fatal events and lack of information on IHD and stroke subtypes limits their value.

In an attempt to address some of these gaps in understanding, this chapter aimed to examine, using prospective data from the CKB, a) the incidence and correlates of major CVDs; b) the associations of diabetes, both self-reported and screen-detected, with risks of major CVDs including IHD and stroke subtypes; and c) the relevance of other cardiovascular risk factors, to these associations.

6.2 Methods

6.2.1 Exposure definitions

As described in section 4.2.1, participants were classified as having self-reported diabetes if they answered yes to the question “Has a doctor ever told you that you had had diabetes?” in the baseline survey. Participants were classified as having screen-detected diabetes if they did not have self-reported diabetes at baseline, but had a measured plasma glucose level ≥ 7.0 mmol/L if time since last food ≥ 8 hours, or ≥ 11.1 mmol/L if time since last food < 8 hours, or a FPG level ≥ 7.0 mmol/L on subsequent testing.

6.2.2 Outcome definitions

Data on CVD outcomes were obtained from death certificates, CDC disease reporting systems and from the national HI system (Section 3.7). The primary outcomes investigated were CVD mortality (ICD-10 I00-25, I27-88, I95-99), MI (I21-23), major coronary event (MCE: non-fatal MI or fatal IHD [I20-25]), IS (I63), ICH (I61), TS (I60, I61, I63, I64) and major occlusive vascular disease (MOVD: IS or MCE). Fatal and non-fatal events were investigated separately and combined, where relevant, and events were classified as fatal if death from any cause was recorded within 28 days of disease onset. First-ever-in-a-lifetime IHD and stroke events were examined to avoid bias resulting from their influence

on subsequent CVD events, and to limit the effects of reverse causality. Detailed outcome definitions are presented in Table E.1. Participants were followed-up until January 1 2014, when 25 488 (5.0%) participants had died and 2411 (0.5%) were lost to follow-up.

6.2.3 Selection of confounders

Confounding variables were selected for inclusion in the statistical models using the approach described for the cross-sectional analyses (Section 5.2.3). Variables were identified as possible confounders based on a priori knowledge of underlying biological mechanisms and associations. The ultimate selection of variables for inclusion in the main statistical models was based on demonstrated statistical associations with diabetes (Section 4.3.2) and with CVD outcomes, on the findings of causal diagrams^{239, 240} (Figure D.1), and on examination of the effects on risk estimates of progressive addition of potential confounders to the Cox proportional hazards regression models (Section 6.3.2).

6.2.4 Participant exclusions

In addition to the exclusion criteria applied for analyses of the associations of diabetes with prevalent IHD and stroke/TIA (Section 4.2.3), participants reporting a history of doctor-diagnosed IHD or stroke/TIA (n=23 050) at baseline were excluded. After applying these criteria, 488 760 participants (199 896 men, 288 864 women) remained for inclusion in the analyses (Figure 4.1).

6.2.5 Statistical analyses

Incidence and correlates of incident cardiovascular diseases

Direct standardisation was used to compare cardiovascular mortality rates and MOVD, MCE, IS and ICH incidence rates across sexes, age groups and study areas, standardised to the age, sex, study area and education structure of the whole study population, as appropriate. Associations of incident MOVD, MCE, IS, ICH and of CVD mortality with potential confounding factors were examined using Cox proportional hazards models. All analyses were stratified by age-at-risk (5 year age groups), study area and sex, and

adjusted for education and, separately, additionally adjusted for alcohol intake, smoking, SBP, physical activity, BMI, waist circumference, hip circumference and consumption of fresh fruit, meat, dairy products, egg, wheat and tea. Intermediate levels of adjustment were investigated for some associations; in addition to basic adjustment for age, sex, study area and education, associations with smoking were adjusted for alcohol consumption and vice versa. Similarly, associations with adiposity measures were adjusted for each other as previously described (Section 4.2.4) and associations with DBP were adjusted for SBP.

Associations of diabetes with incident cardiovascular diseases

All analyses were done separately for self-reported and screen-detected diabetes, with a comparator group of individuals with no diabetes. Direct standardisation was used to calculate age-, sex-, and study area-standardised disease incidence rates, using the total study population as the standard. Cox proportional hazards models were used to obtain HRs and associated CIs for the associations between baseline diabetes status and the described endpoints. Age-at-risk <35 and >79 years was excluded since there were very small event numbers at these ages. Participants were followed-up until 1st January 2014 and were censored at the date of their first event (as described in Table E.1), loss to follow-up or death. Models were stratified by age-at-risk (5-year age groups), study area and, where appropriate, sex. As such, the proportional hazards assumption was not assumed between the strata of age groups, study areas or sex. Potential confounding factors were then progressively added to the models to examine the effect of additional adjustment for these factors.

Based on these analyses and the additional considerations described in Section 6.2.3, the statistical models for the main and subsidiary analyses were stratified by age-at-risk, study area and, where appropriate, sex and were adjusted simultaneously (except where the variable of interest in subsidiary analyses to investigate effect modification) for education (no formal education, primary school, middle school, high school, college/university), smoking (never, occasional, ex-regular, current regular), alcohol (never, occasional intake,

ex-regular, reduced intake, weekly intake), SBP (<100, 100-109, 110-119, 120-129, 130-139, 140-149, 150-159, 160-169, ≥ 170 mmHg), and physical activity (<10, 10-19.9, 20-29.9, 30-39.9 ≥ 40 MET hours/day). SBP and physical activity were included in models as categorical variables to avoid any inappropriate assumptions being made about their distribution. Additional analyses were undertaken with SBP and physical activity instead included as continuous variables, which did not alter the calculated risk estimates.

Comparison of HRs for the first four and subsequent years of follow-up revealed no evidence of departure from the proportional hazards assumption. Adjusted HRs were calculated across strata of other cardiovascular risk factors (age, sex, region, education, smoking, alcohol, SBP, physical activity, BMI) and across the 10 study areas. In addition, for self-reported diabetes, adjusted HRs were calculated across strata of baseline duration of diabetes diagnosis and diabetes treatment status. Chi-square tests for trend and heterogeneity were applied to the log HRs and their SEs.²⁴² Chi-square statistics and degrees of freedom were summed for heterogeneity tests across all subgroup analyses to obtain a global test for heterogeneity.²⁴² In further analyses, effect estimates were also adjusted for waist-to-hip ratio (<0.75, 0.75 to <0.80, 0.80 to <0.85, 0.85 to <0.90, 0.90 to <0.95, ≥ 0.95). Based broadly on the ADA targets for glycaemic control,²⁶⁰ participants with self-reported diabetes were categorised according to their baseline blood glucose levels: good control was defined as a blood glucose level <7.2 mmol/L if time since last food was ≥ 8 hours or a blood glucose level <10.0 mmol/L if time since last food was <8 hours, and poor control was defined as a blood glucose level ≥ 7.2 mmol/L if time since last food was ≥ 8 hours or a blood glucose level ≥ 10.0 mmol/L if time since last food was <8 hours. Finally, the HR for MOVD associated with numbers of other presenting cardiovascular risk factors (hypertension: self-reported hypertension, mean SBP ≥ 140 mmHg or DBP ≥ 90 mmHg; overweight or obesity: BMI ≥ 25 kg/m²; ever regular smoking; physical activity: <10 MET hours/day) among individuals with diabetes was assessed. Analyses stratifying by glycaemic control and by diabetes diagnosis duration and treatment, and those

investigating the effects of the number of other CVD risk factors used the floating absolute risk method.²⁴¹

In additional analyses, the strength of the associations between self-reported diabetes and incident MOVD, MCE, IS and ICH were calculated separately for disease events identified through linkage to the HI system only, through the CDC disease reporting system only and through both (defined as reporting of the same CVD event type in both data sources within a seven day period). For stroke outcomes, the strength of associations with self-reported diabetes were additionally calculated separately for events confirmed, and not confirmed, by CT/MR scan. These latter analyses were restricted to stroke events identified through the CDC disease reporting system, since procedural data in the HI system are not currently available.

Macros developed within CTSU were used to undertake the Cox proportional hazards analyses and to calculate directly standardised disease incidence and mortality rates. SAS version 9.3 (SAS Institute Inc., Cary, North Carolina, USA) was used to conduct the statistical analyses. Line graphs and forest plots were produced using R version 2.13.1 (R Foundation for Statistical Computing, Vienna, Austria); plots presenting the results of the Cox proportional hazards analyses were produced using an R graphics programme developed within CTSU.

6.3 Results

Among 488 760 participants without prior CVD, the mean (SD) baseline age was 51 (11) years and 2.7% (n=13 284) reported a history of doctor-diagnosed diabetes, with a further 2.7% (n=13 051) having screen-detected diabetes. The prevalence of both self-reported and screen-detected diabetes was slightly higher in women than in men (2.9% versus 2.5% and 2.8% versus 2.6%, respectively). Overall, 16.6% of these participants with self-reported diabetes reported using any CVD medication at baseline; 15.1%, 2.1% and <1% reported

taking anti-hypertensives, aspirin and statins, respectively (76.0% reported taking diabetes medications) (Table E.2).

6.3.1 Incidence and correlates of cardiovascular diseases

During ~3.4 million person-years of follow-up, 6916 (1.4%) participants died from CVD, with standardised mortality rates of 357.5 per 100 000 person years in men and 136.9 per 100 000 person years in women. There were 22 814 MOVD events recorded during follow-up, representing an incidence of 941.5 and 592.5 per 100 000 person years in men and women, respectively. IS (n= 19 778) accounted for the greatest proportion of MOVD events, with incidence rates of 784.7 and 530.0 per 100 000 person years in men and women, respectively. There were 4423 ICH events (205.4 and 96.6 per 100 000 person years in men and women, respectively) and 3465 MCE, again with a higher incidence in men than in women (171.7 versus 69.2 per 100 000 person years).

The incidence of all major CVDs examined increased with age. Incidence was higher in men than in women at all ages >40 years, after accounting for study area and education, with more marked sex-differences at older ages (Figure 6.1). There was considerable regional variation in CVD incidence (Figures 6.2 to 6.6). Overall, the age- and education-standardised CVD mortality rate was similar in urban and rural areas (163.2 and 154.9 per 100 000 person years, respectively). The incidence of major ischaemic CVDs was, however, higher in urban, than in rural, areas (MOVD: urban 914.1 versus rural 451.3 per 100 000 person years; MCE: urban 107.9 versus rural 74.0 per 100 000 person years; IS: urban 820.3 versus rural 383.3 per 100 000 person years). The ICH incidence rate, however, was higher in rural, than in urban, areas (126.0 versus 77.3 per 100 000 person years).

The associations of the main potential confounding factors with incident MOVD and ICH, encompassing both ischaemic and haemorrhagic CVD subtypes, are presented in Tables E3 to E.6 and Figures E.1 to E.4 (associations with CVD mortality, IS and MCE were similar

to those with MOVD, and are not therefore presented). These confounder analyses are not discussed in detail here, as this is beyond the scope of this thesis. In summary, there were independent associations with several lifestyle factors, most notably elevated risks of ischaemic CVDs in current regular smokers and of all CVDs in less physically active participants, and with some dietary variables (most consistently lower risks of CVDs in participants more frequently consuming fresh fruit). There were also positive associations, particularly for ischaemic CVDs, with several adiposity measures and with blood pressure.

6.3.2 Diabetes and risk of cardiovascular disease

Individuals with self-reported diabetes had a two-fold increased risk of CVD mortality (adjusted HR 2.07, 95% CI 1.90-2.26, Table 6.1). The risk was higher in women than in men (2.37 [2.11-2.65] vs 1.73 [1.50-2.00], $p < 0.001$) (Figure 6.7), although the absolute excess mortality associated with diabetes was greater in men due to higher CVD mortality rates in men. With the exception of the youngest participants, the sex-difference in relative risk was evident at all ages. Self-reported diabetes was also independently associated with increased risks of MOVD (1.77, 1.69-1.86), MCE (2.44, 2.18-2.73) and IS (1.68, 1.60-1.77), with somewhat greater risks for fatal, than non-fatal, MI and IS (Table 6.1). For MOVD, MCE and IS, there was no clear evidence of sex-differences in adjusted HRs (p for heterogeneity=0.7, 0.2 and 0.9, respectively) (Figures 6.7 and 6.8). When fatal and non-fatal events were examined separately, there was some suggestion of greater diabetes-associated risks in women than in men for fatal, but not non-fatal, MI and IS, but these sex-differences were not significant (Figure 6.9). For MOVD and IS there were stronger self-reported diabetes-associated risks in younger, than older, age groups (p for trend < 0.001) (Figures 6.7 and 6.8). Additional adjustment for waist-to-hip ratio modestly, but non-significantly, attenuated the HRs for MOVD, MCE and IS, but not for CVD mortality (Table 6.1). Further adjustment for other potential confounding variables, including DBP and dietary variables, did not markedly attenuate these associations (Figures 6.10 and 6.11). For CVD mortality and MCE the diabetes-associated risks varied little across study areas

(p for heterogeneity=0.3 and 0.2, respectively), while for MOVD and IS there was some evidence of variation between study areas (p for heterogeneity <0.001) (Figures 6.12 and 6.13).

For CVD mortality the self-reported diabetes associated-risk was somewhat stronger in never-regular smokers and never-regular alcohol drinkers (Figure 6.14). For MOVD (Figure 6.15) and IS (Figure 6.16) there were stronger associations in rural areas (p for heterogeneity <0.001 and 0.003, respectively) and possibly also among less highly educated individuals (p for heterogeneity=0.06 and 0.03, respectively). The risk of MCE varied little across population subgroups (Figure 6.17).

For ICH, the diabetes-associated HR was more modest (1.24, 95% CI 1.07-1.44), but was more extreme for fatal 1.59 (1.30-1.96), than for non-fatal 1.00 (0.81-1.23), events (Table 6.1). There was no clear difference in adjusted HRs for ICH between men and women (p for heterogeneity=0.8) or across age groups (p for trend=0.4) (Figure 6.18) or study areas (p for heterogeneity=0.7) (Figure 6.19). Additional adjustment for potential confounding factors did not materially alter the association (Figure 6.20). The diabetes-associated risk of ICH varied little across strata of other cardiovascular risk factors (Figure 6.21). For TS, self-reported diabetes was associated with a 61% excess risk (1.61, 1.54-1.69) (Table 6.1).

Among individuals with self-reported diabetes, the risk of major ischaemic CVDs increased with longer duration of diabetes (p for trend=0.003, 0.004 and 0.04 for MOVD, MCE and IS, respectively), but no such trend was seen for CVD mortality (p =0.1) or ICH (p =1.0) (Figure 6.22). The risks of CVD mortality, MOVD, IS and MCE were modestly, but non-significantly, higher in individuals who reported taking oral hypoglycaemic agents or insulin than in those who did not (p for heterogeneity=0.3, 0.2, 0.2 and 0.4, respectively), but risks of ICH were very similar in the two groups (Figure 6.23). Risks of CVD mortality, MOVD, MCE and IS were higher among participants with poorer glycaemic control than among those with better control (p for heterogeneity <0.001) (Figure 6.24). During the 7-year follow-up, 8896

participants were newly diagnosed with diabetes. Exclusion of these individuals from the analyses had little effect on the self-reported diabetes-associated CVD risks (Table E.7).

Among participants with screen-detected diabetes, the excess CVD risks were somewhat less extreme than those associated with self-reported diabetes. For CVD mortality, the adjusted HR was 1.66 (95% CI 1.51-1.83), while for MOVD, MCE and IS they were 1.50 (1.41-1.58), 1.62 (1.40-1.86), and 1.48 (1.40-1.57), respectively (Table 6.2). The excess risk for ICH was more modest (1.17, 1.01-1.36). There was some heterogeneity across study areas in screen-detected diabetes-associated HRs for MOVD (p for heterogeneity=0.02) and IS (p for heterogeneity=0.007) (Figures E.5 and E.6), but not for CVD mortality, MCE or ICH (p for heterogeneity=0.5, 0.8 and 0.4, respectively) (Figures E.5-E.7). For MOVD (Figure E.8) and IS (Figure E.9) the screen-detected diabetes-associated risk was somewhat greater in younger participants (p for trend <0.001), in residents of rural areas (p for heterogeneity=0.003 and <0.001, respectively) and in the more physically active (p for trend=0.006 and 0.008, respectively). There was little evidence of heterogeneity in HRs for CVD mortality (Figure E.10) or MCE (Figure E.11) across strata of other risk factors. For ICH, there was some evidence of a stronger association in men (p for heterogeneity=0.03) and in individuals with a lower SBP (p for trend=0.04), although the relevance of these observations, which may, at least partly, reflect multiple testing ($p=0.4$ for global heterogeneity test), is unclear (Figure E.12).

Among individuals with self-reported diabetes, the risk of MOVD increased progressively with number of other presenting cardiovascular risk factors, with adjusted HRs of 1.68 (1.53-1.85), 2.08 (1.93-2.23) and 2.82 (2.59-3.06) in those with one, two and three or more risk factors, respectively, compared with participants with self-reported diabetes who had none of these risk factors (p for trend <0.001; Figure 6.25). A similar gradient was seen for screen-detected diabetes (p for trend <0.001).

There was no significant difference in the self-reported diabetes-associated risks of MCE or ICH by data source (CDC disease reporting system only, HI system only or both) (p for heterogeneity=0.2). There were, however, modestly lower diabetes-associated HRs for MOVD and IS identified through HI data only (HR 1.68, 99% CI 1.53-1.83 and 1.60, 1.46-1.76, respectively) than through the CDC disease reporting system only (1.98, 1.74-2.26 and 1.94, 1.68-2.22, respectively) or through both data sources (1.85, 1.58-2.16 and IS 1.80, 1.54-2.12, respectively) (MOVD: p for heterogeneity=0.02; IS: p for heterogeneity=0.01). The majority of IS (87.8%) and ICH (92.5%) events identified through the CDC disease reporting system were confirmed by CT or MR scan. There was weak evidence of a stronger association of diabetes with non-scan-confirmed IS than with scan-confirmed IS (2.24, 1.71-2.93 and 1.79, 1.59-2.01, respectively) (p for heterogeneity=0.05). For ICH no significant difference was seen by scan status (1.16, 0.35-3.89 and 1.01, 0.67-1.50, respectively) (p for heterogeneity=0.8).

6.4 Discussion

The presented analyses provide the first large-scale prospective evidence of the cardiovascular consequences of diabetes among adults in China. They showed that self-reported, doctor-diagnosed diabetes was associated with 1.5- to 2.5-fold increased risks of cardiovascular mortality and incident IHD and IS. Moreover, they provide evidence of a significant, though more modest, adverse effect of self-reported diabetes on incident ICH risk. These diabetes-related excess risks tended to increase progressively with longer diabetes duration, with increasing number of other cardiovascular risk factors and possibly also with poorer glycaemic control. Individuals with screen-detected diabetes had significantly, although more modestly, increased CVD risks compared with individuals without diabetes.

6.4.1 Incidence of cardiovascular diseases

Reliable, contemporary estimates of IHD and stroke incidence and mortality rates in China are limited, with no recent incidence data available.⁵⁴ Some informative, but less contemporary, data on IHD incidence come from the Sino-MONICA project, a prospective community-based active disease surveillance study which collected data on all acute coronary events from medical records, patient or proxy interviews and death certificates,²⁶¹ between 1987 and 1993 in a population of approximately 5 million men and women aged 35-64 years in 16 mainland China provinces.⁵⁸ Age-standardised IHD incidence varied widely across Sino-MONICA populations in the included provinces, and although there were no clear trends over the study period, incidence rates were generally higher in northern than in southern areas, and in men, than in women. The same study also provided data on stroke incidence.⁵⁸ There were large variations in stroke incidence rates between provinces, with higher rates in northern, than in southern, study areas. Longer follow-up (1984 to 2004) of men and women aged 25 to 74 years in one Sino-MONICA population in Beijing, including ~14 500 stroke events, showed a 1.7% (95% CI 2.7 to 1.2, $p < 0.05$) annual reduction in age-standardised HS incidence throughout the study period, whilst the incidence of IS increased by an average 8.7% (4.3 to 8.9%, $p < 0.05$) annually.²⁶² IHD and stroke incidence rates in the CKB were generally higher than those previously published.^{58, 262, 263} However, in light of likely temporal trends, differences in study populations and diagnostic criteria, and the non-representative nature of the CKB population, there is limited value in directly comparing rates in the CKB with these other studies. Comparison of patterns in CVD incidence between population subgroups within studies is, however, more relevant. The higher incidence of stroke, than of IHD, in the CKB is consistent with findings from the Sino-MONICA study⁵⁸ and with the classical description of patterns of CVD in China.²⁶⁴ Increasing incidence of CVDs with age⁵⁵ and higher incidence of IHD and stroke in men than in women, seen in the CKB, have also been shown previously.^{55, 58, 263} Previous studies suggest both IHD⁵⁴ and TS⁵⁵ incidence rates are higher in urban than in rural areas,

and similar findings were seen in the CKB. Data from previous studies on stroke subtypes are limited, but the CKB showed that IS rates were higher in urban, than in rural, areas, while the converse was true for ICH incidence, possibly reflecting regional differences in blood pressure, which shows a particularly strong association with ICH risk in the CKB.²⁶⁵

6.4.2 Diabetes and risk of cardiovascular disease

As in the CKB, previous studies of predominantly Western populations have relatively consistently shown approximate doubling of IHD risk in diabetes.^{6, 129-131, 133} The comparability of risk estimates suggests the differences described between Asian and Western populations in the relative importance of insulin resistance and pancreatic beta-cell dysfunction in diabetes aetiology⁸⁸ have little, if any, impact on associated IHD risk. Several studies have also reported on the association of diabetes with TS^{129, 130} or IS.^{6, 129, 131, 136} Compared with IHD, these have shown similar, or slightly stronger, diabetes-associated IS risk^{6, 129, 131} and similar, or moderately weaker, TS risk.^{129, 130} The association of screen-detected diabetes with IS in the CKB is largely consistent with data from studies of mostly Western populations.⁶ The association with self-reported diabetes was moderately weaker than in those studies,^{6, 129} and than in a small prospective Chinese study of 30 000 men and women,¹³¹ and a study, based on routine data, of half a million Chinese adults in Taiwan.¹³⁶ The relatively high proportion of lacunar infarcts among IS cases in China,²⁶⁶ which may be detected through CT/MR scanning without apparent focal neurological deficit, and the relatively low IS case fatality might contribute to more modest effects in the CKB.

Existing evidence relating ICH to diabetes is more limited than for IS,^{6, 91, 129, 131, 136} due partly to lower ICH rates in Western populations,^{6, 129} and partly to lack of widespread imaging, particularly in earlier studies. Where available, previous findings have been inconsistent; some have shown moderate, positive associations with diabetes,^{6, 90, 136} and others no association.^{91, 129, 131} Two of the largest studies, one with ~1000⁶ ICH events and one with

~2300⁹⁰ events, found modest elevated risks of ICH associated with known diabetes (HRs 1.56⁶ and 1.28⁹⁰). The CKB included almost 5000 well-characterised ICH cases, more than in previous studies combined, and has provided reliable evidence of a positive, independent association between self-reported diabetes and ICH. The different diabetes-associated risks of stroke subtypes highlight the importance of their separate investigation, and the large number of well-characterised stroke events (~90% of validated stroke events had been confirmed on CT/MRI) is a strength of the CKB. The apparent difference in excess risk between fatal and non-fatal CVD events could reflect poorer survival following such events in diabetes^{253, 267} or more severe disease in fatal cases.

In the CKB, self-reported diabetes-associated CVD mortality risks were higher in never-regular smokers and never-regular alcohol drinkers. As well as lower absolute cardiovascular risks, the disproportionately higher percentages of women (who had higher diabetes-associated risks of CVD mortality) in these groups may explain the higher risks. The self-reported diabetes-associated cardiovascular risk, especially for IS, was greater in rural than in urban areas. This may reflect the higher proportion of undiagnosed diabetes cases in rural areas, such that self-reported diabetes may represent more severe disease in these areas. More frequent use of diabetes medications and higher average baseline RPG among individuals with self-reported diabetes in rural, than in urban, areas supports this hypothesis.

Self-reported diabetes-associated IS risks were lower in Harbin and Haikou, but the reasons for this are not immediately obvious. Harbin also had notably high IS incidence (Section 6.3.1), consistent with previous observations on stroke rates in northern regions of China more generally.²⁶⁸ The north-south gradient in stroke reported previously in China has been attributed to higher levels of hypertension in northern regions,²⁶⁸ but blood pressure was not found to be higher in CKB participants in Harbin, than in other study areas. Importantly, the lower diabetes-associated risks of IS in these study areas could reflect the play of chance.

The presented findings provide the first large-scale evidence for adult Chinese that although, on average, diabetes is associated with approximately two-fold increased CVD risk, the risk is also dependent on the presence of other cardiovascular risk factors. This is particularly relevant in China, where there are high levels of other risk factors, including smoking in men and hypertension,²⁶⁹ and highlights the need for multi-factorial approaches to managing disease risk.

Whilst tight glycaemic control has been shown to reduce the risk of microvascular complications of diabetes,^{270, 271} controversy remains with regard to its influence on the risk of macrovascular complications. Whilst acknowledging the limitations of single RPG measurements as an indicator of glycaemic control in diabetes, findings from the CKB are consistent with previous observational data, showing greater CVD risk in association with higher glucose or glycated haemoglobin levels among individuals with diabetes.^{272, 273} Despite these observational findings, trials of intensive glycaemic control in diabetes have been inconclusive. Initial follow-up in some individual studies has shown no convincing benefits on CVD risk,^{270, 274, 275} and others have even reported increased mortality risks.²⁷⁶ In a recent published data meta-analysis of 12 trials, more intensive glycaemic control in type 2 diabetes had no significant effect on cardiovascular mortality (HR 1.11, 95% CI 0.95-1.35) but a possible protective effect on non-fatal MI (0.85, 0.76-0.95).²⁷⁷ The relevance of this finding in the CKB, and in previous observational analyses,^{272, 273} is thus unclear. Interestingly, however, in three randomised controlled trials (RCTs) of intensive glycaemic control in diabetes, findings from post-trial follow-up were suggestive of lower CVD risk associated with intensive control, despite little post-intervention difference in glycaemic control between trial arms.²⁷⁸⁻²⁸⁰ The UK Prospective Diabetes Study of almost 4000 newly diagnosed type 2 diabetes patients found significant reductions in the risk of MI (n=1204) with intensive control after up to 10 years post-intervention follow-up (up to 30 years total follow-up).²⁷⁸ The study showed HRs of 0.85 (95% CI 0.74-0.97) in those allocated to intensive control with a sulfonylurea or insulin compared with their corresponding standard

therapy group, and 0.67 (0.51-0.89) in those allocated to intensive control with metformin treatment, compared with their corresponding standard therapy group. There was no significant effect on stroke risk (0.91 [0.73-1.13] and 0.80 [0.50-1.27], respectively) (n=452), but this may reflect inadequate power. The Diabetes Control and Complications Trial examined the effects of intensive versus standard therapy in ~1500 type 1 diabetes patients; during a mean 17-year follow-up (11.5 years post-intervention) intensive treatment was associated with a 42% (9-63%) reduction in the risk of CVD.²⁷⁹ Most recently, the Veterans Affairs Diabetes Trial, in almost 2000 type 2 diabetes patients, showed a just significant reduction in the risk of major cardiovascular events (0.83, 0.70-0.99) associated with intensive (versus standard) treatment after approximately 10 years' follow-up (~4 years post-intervention).²⁸⁰ There was, however, no reduction in CVD mortality (0.88, 0.64-1.20). The mechanisms underlying this phenomenon, termed the "legacy effect", are poorly understood, but good control, particularly in the early stages of diabetes, may result in long-term benefits for CVD risk.²⁷⁸ In contrast to the uncertainty surrounding the benefits of intensive glycaemic control for CVD risk, there is clear evidence for beneficial effects of appropriate management of other CVD risk factors, including blood pressure²⁸¹ and lipids²⁸² in diabetes. Despite this, few CKB participants with self-reported diabetes at baseline reported use of conventional cardio-protective therapy. 51.2% had elevated blood pressure at baseline (SBP $\geq$ 140 mmHg and/or DBP $\geq$ 90 mmHg),²⁸³ including 69.6% of those treated with anti-hypertensives. Data on lipids in the CKB are not currently available.

A notable finding in the CKB is the absence of consistent sex-differences in diabetes-associated risks of CVDs; a higher diabetes-associated risk in women was evident only for CVD mortality. Several studies have reported greater diabetes-associated relative risk for IHD in women, typically 30-50% higher than in men,^{6, 133} and, although less consistently,^{6, 90} for stroke.¹³⁸ In the small USA-PRC of ~1000 Chinese adults with ~100 IHD events, the diabetes-associated HRs were more extreme in women than in men for IHD (women: HR 2.99; men: 1.51) and for IS (women: 2.89; men: 1.08).¹³² Similar sex-differences in

diabetes-associated risk of IS were seen in a study of ~1 million adults in Taiwan, based on routine healthcare data (2.32 vs. 1.89).¹³⁶ Both studies were inadequately adjusted for potential confounding factors, with factors such as smoking, blood pressure, physical activity and lipids not considered. By contrast, in a study of over 65 000 elderly residents of Hong Kong, with more extensive adjustment for potential confounding factors including socio-economic status, alcohol use, smoking, physical activity and adiposity, there was no sex-difference in diabetes-associated risk of IHD (men: 2.20, 1.89-2.57; women: 2.25, 1.98-2.55).¹³⁰ Other more recent studies, in a variety of settings, have failed to show convincing sex-differences in diabetes-associated risks of a range of CVDs, including a study of 1.9 million individuals in the UK (using routine healthcare data from mainly the 2000s)⁹⁰ and a cohort of 150 000 middle-aged adults in Mexico City, recruited in 1999-2001 and undergoing continued follow-up.²⁸⁴

The previously-described sex-difference in diabetes-associated disease risks has been attributed to a number of factors. It has been suggested that diabetes is associated with a more adverse CVD risk factor profile in women, such that the extent of differences in CVD risk factor abnormalities, including abnormalities of lipids and blood pressure, between women with and without diabetes is greater than between men with and without diabetes.^{245, 285-288} This is supported by the observed attenuation of sex-differences with adjustment for other known CVD risk factors.^{245, 286, 289} The sex-differences may also reflect differences in the appropriateness and success of diabetes and CVD treatment favouring men.^{288, 290, 291} More novel hypotheses for the sex-difference in diabetes-associated risks include a greater tendency for abnormalities of coagulation, endothelial function and oxidative stress among women, than men, with diabetes,^{136, 288, 292, 293} and possible, but currently poorly understood, roles for gestational diabetes or polycystic ovarian syndrome.²⁸⁸

In the CKB there were no clear sex-differences in the frequency of use of diabetes or CVD medications in self-reported diabetes (Section 4.3.1). There was a modestly greater difference in mean SBP levels between women with and without diabetes than between

men with and without diabetes ($p < 0.001$), but such an effect was not seen for other measured risk factors (Table E.8). There was also no clinically significant difference in RPG levels between men and women with diabetes (Section 4.3.1). These factors, along with robust adjustment for a wide range of potential confounding factors, may help to explain the lack of any clear sex-differences in diabetes-associated risks for incident IHD and stroke in the CKB. The observed greater diabetes-associated relative risk of CVD mortality in women could reflect sex-differences in survival following CVD events.²⁹⁴

Separate examination of screen-detected diabetes in our study should help to limit bias from lifestyle changes following diagnosis. The more modest effects of screen-detected, than self-reported, diabetes in the CKB may reflect shorter duration of, or less severe glycaemic aberration in, screen-detected diabetes, but also greater misclassification of screen-detected diabetes.²¹⁴⁻²¹⁷

6.4.3 Accuracy of ischaemic heart disease and stroke diagnoses

CVD outcome data included in the presented analyses were obtained from three sources: the CDC disease reporting system, the national HI system and death registries. Some cohort studies, particularly smaller or older studies, have relied on self-report or on clinical review to identify disease outcomes.^{131, 243, 295} Whilst the latter approach, in particular, may assist accurate phenotyping, it is resource intensive and may limit sample size or event numbers.²⁹⁶ Obtaining outcome data from multiple, pre-existing data sources, as in the CKB, can enable relatively cost-efficient identification of large numbers of disease events over long follow-up periods, and can reduce susceptibility to selection and recall biases,²⁹⁷ as well as limiting biases inherent in the use of a single data source.²⁹⁸ Extraction of data may, however, be resource intensive, particularly from administrative data sources such as HI data, which, since they are collected for cost-reimbursement, may be susceptible to their own unique biases.^{296, 297} Appropriate use of these data is thus clearly reliant upon an understanding of their quality.^{297, 298}

An outcome adjudication process, including reported cases of IHD and stroke prior to 1st January 2014, is currently underway in the CKB, as described (Section 3.7). An attempt was made to retrieve medical notes for all study participants with an IHD or stroke event recorded in any of the three data sources. Medical records were retrieved for over 12 500 participants (~80% of those selected) with a recorded diagnosis of IHD. Based on medical record review, the overall positive predictive value (PPV) for IHD events was 83.8%. PPVs for IHD subtypes, however, were lower (MI 51.3% and angina 37.8%), and more detailed investigations are underway to better understand these preliminary findings. For stroke the retrieval rate of medical notes was 82.9% (n=17 478). The TS PPV was excellent at 91.3%, with good, but lower, PPVs for IS (70.8%) and ICH (70.2%). There was a moderately higher PPV for TS events recorded in HI data (72.5%) than in the CDC disease reporting system (63.8%). Importantly, the current inability to retrieve medical records for a proportion of CKB participants with IHD or stroke events recorded in one of the three data sources used may have distorted these findings.

Although data have not, to my knowledge, been published on the validity of CVD diagnoses recorded in the Chinese HI system or in the disease surveillance system, evaluations of the validity of cause of death data have been reported.^{299, 300} These demonstrated reasonable validity of cerebrovascular disease deaths in both urban (sensitivity 81.6%, PPV 88.4%) and rural (73.6%, PPV 75.9%) areas of China, but, again, lower validity for IHD (urban: sensitivity 69.4%, PPV 84.4%; rural: sensitivity 53.1%, PPV 40.2%). In keeping with that study, the CKB found comparable PPVs for stroke in rural and urban areas (90.9% and 91.7%, respectively), but a modestly lower PPV for IHD in rural, than in urban, areas (81.6% and 85.1%, respectively). Outside China, diverse estimates of the validity of CVD diagnoses in administrative data have been reported. In a systematic review of studies investigating the validity of MI diagnoses in administrative data, including 29 studies, the PPVs ranged from 48.7% to 98.8%,²⁹⁷ whilst for stroke PPVs have ranged between approximately 60% and 95%.³⁰¹⁻³⁰⁴ Less accurate recording of IHD diagnoses in the CKB may have biased

estimates of diabetes-associated IHD risks, likely towards the null, although the use of MCE (fatal IHD and non-fatal MI) as the main IHD outcome would be expected to have reduced the impact of such inaccuracies.

6.4.4 Limitations

The relative strengths and weaknesses of the diabetes definitions used in these analyses have been described in Section 4.4.2. Screen-detected diabetes in the CKB may be particularly susceptible to misclassification resulting from the use of RPG²¹⁶ measured on a glucometer²³³ for diagnosis. Whilst use of fasting time data would be expected to reduce the impact of the limitations of this measurement technique, it is likely that a proportion of diabetes cases remained undiagnosed,²¹⁴⁻²¹⁷ resulting in underestimated diabetes-associated CVD risks. Diagnoses of diabetes made during the follow-up period were not routinely accounted for in the presented analyses, which could contribute to underestimation of diabetes-associated risks. However, sensitivity analyses excluding individuals newly diagnosed with diabetes during follow-up did not materially alter risk estimates.

In addition to the possible impacts of inaccuracies in recording of CVD outcomes, as described above, the data sources used in the CKB may not capture milder, non-fatal CVD events. Although clearly superior to cross-sectional analyses, observational analyses such as these remain susceptible to the effects of bias and residual confounding, limiting the conclusions that can be drawn with respect to causality of the association of diabetes with CVD. In particular, the nature of recruitment to the CKB introduces the possibility of a “healthy volunteer effect”, with potential for underestimation of diabetes-associated risks, particularly given the relatively short follow-up period.³⁰⁵ Furthermore, diabetes, particularly type 2 diabetes, frequently co-exists with other cardiovascular risk factors;⁹⁴ although the risk estimates presented were extensively adjusted for potential confounders, the current inability to adjust for the effects of lipids and likely inaccuracies in measurement of other variables may have led to some residual confounding. In addition, although diabetes-

associated CVD HRs were similar for the first and second halves of follow-up, the follow-up period was relatively short, and it will be of value to repeat the analyses after longer follow-up, excluding the first few years to reduce potential for reverse causality.

6.4.5 Conclusion

In China, over 10% of the adult population is estimated to have diabetes.⁸⁷ The present nationwide study provides the first large-scale, prospective evidence from mainland China that individuals with diabetes are at significantly elevated risk of CVDs, similar in magnitude to that observed in Western populations.^{6, 129} With further adverse changes in lifestyle, including diet and physical activity, the prevalence of diabetes will likely increase further in China, especially among young adults,⁶⁵ foreshadowing a substantial diabetes-attributable burden of CVDs.

Figure 6.1. Incidence of cardiovascular diseases by age and sex

Standardised to study area and education structure of study population.

Squares represent CVD incidence and vertical lines represent corresponding 95% confidence intervals.

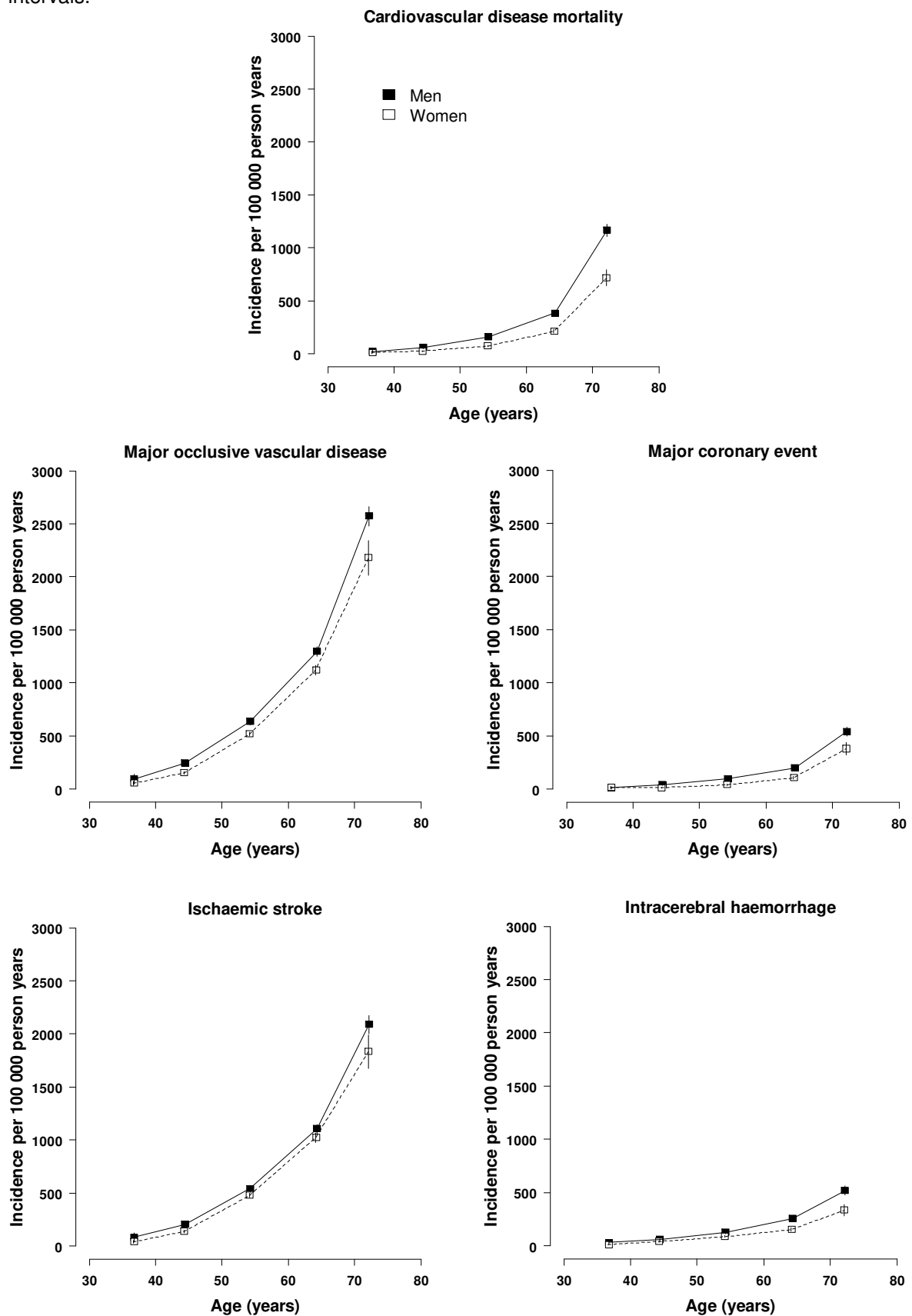


Figure 6.2. Incidence of cardiovascular disease mortality by study area

Standardised to age and education structure of study population.

Error bars represent 95% confidence intervals.

R, rural; U, urban.

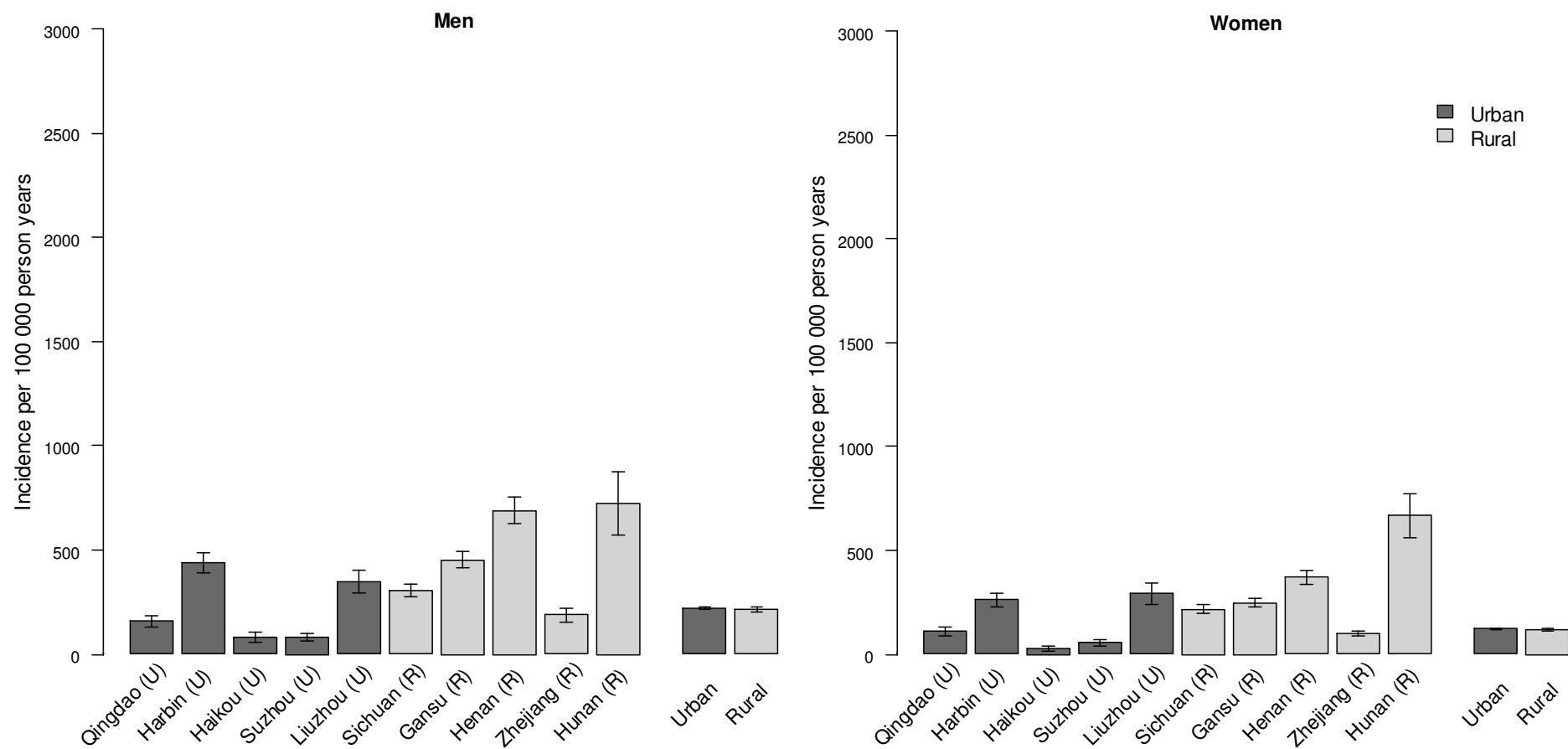


Figure 6.3. Incidence of major occlusive vascular disease by study area

Standardised to age and education structure of study population.

Error bars represent 95% confidence intervals.

R, rural; U, urban.

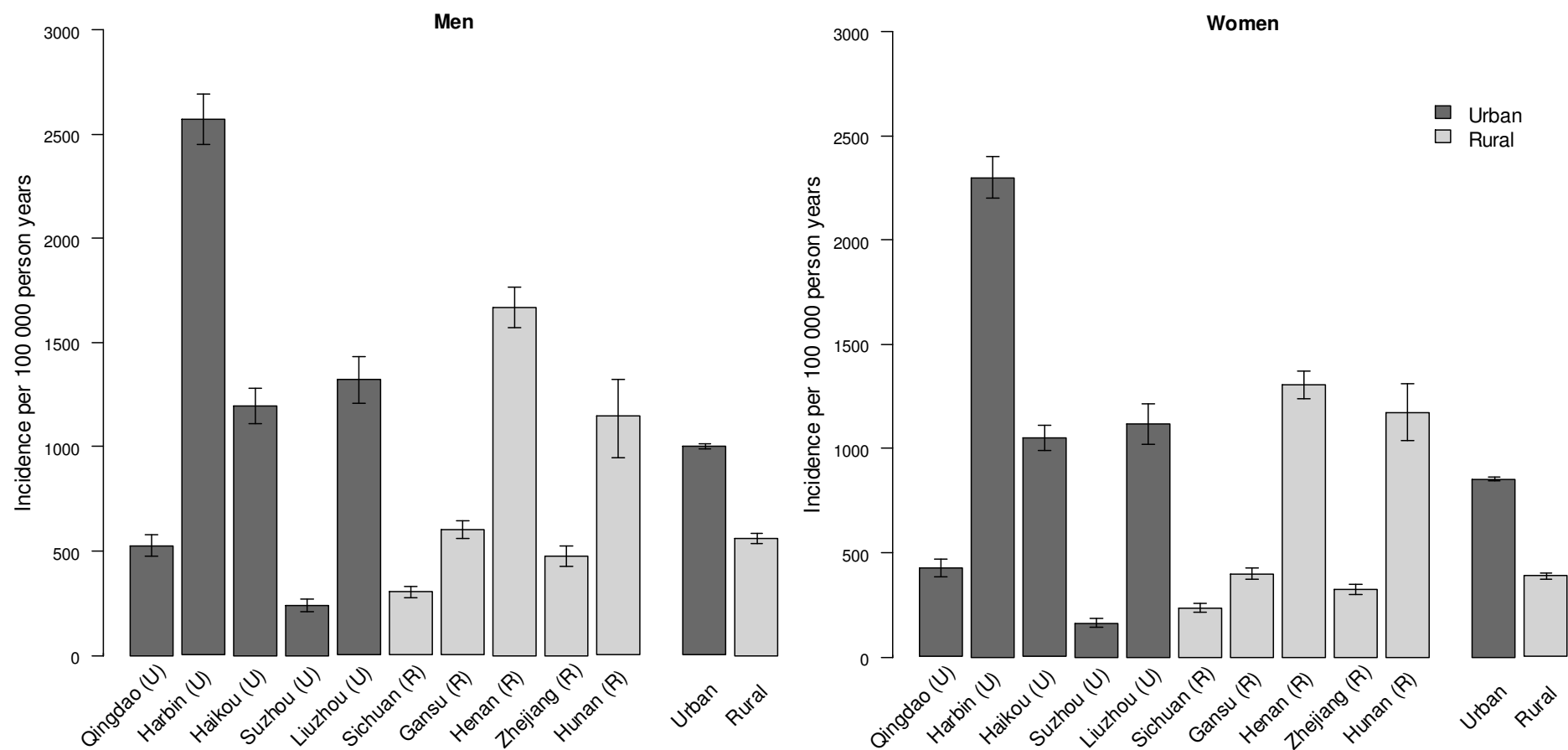


Figure 6.4. Incidence of major coronary event by study area

Standardised to age and education structure of study population.

Error bars represent 95% confidence intervals.

R, rural; U, urban.

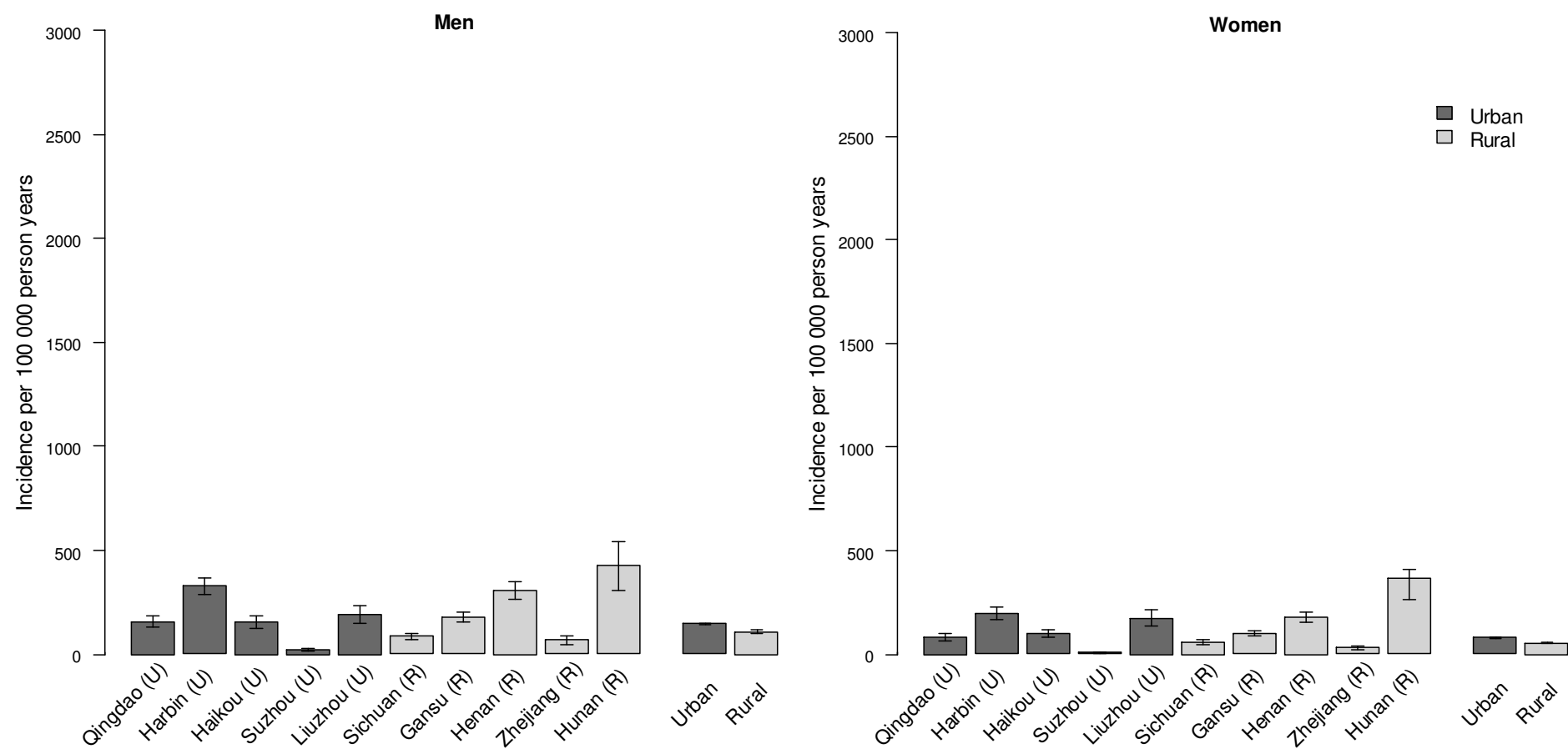


Figure 6.5. Incidence of ischaemic stroke by study area

Standardised to age and education structure of study population.
Error bars represent 95% confidence intervals.
R, rural; U, urban.

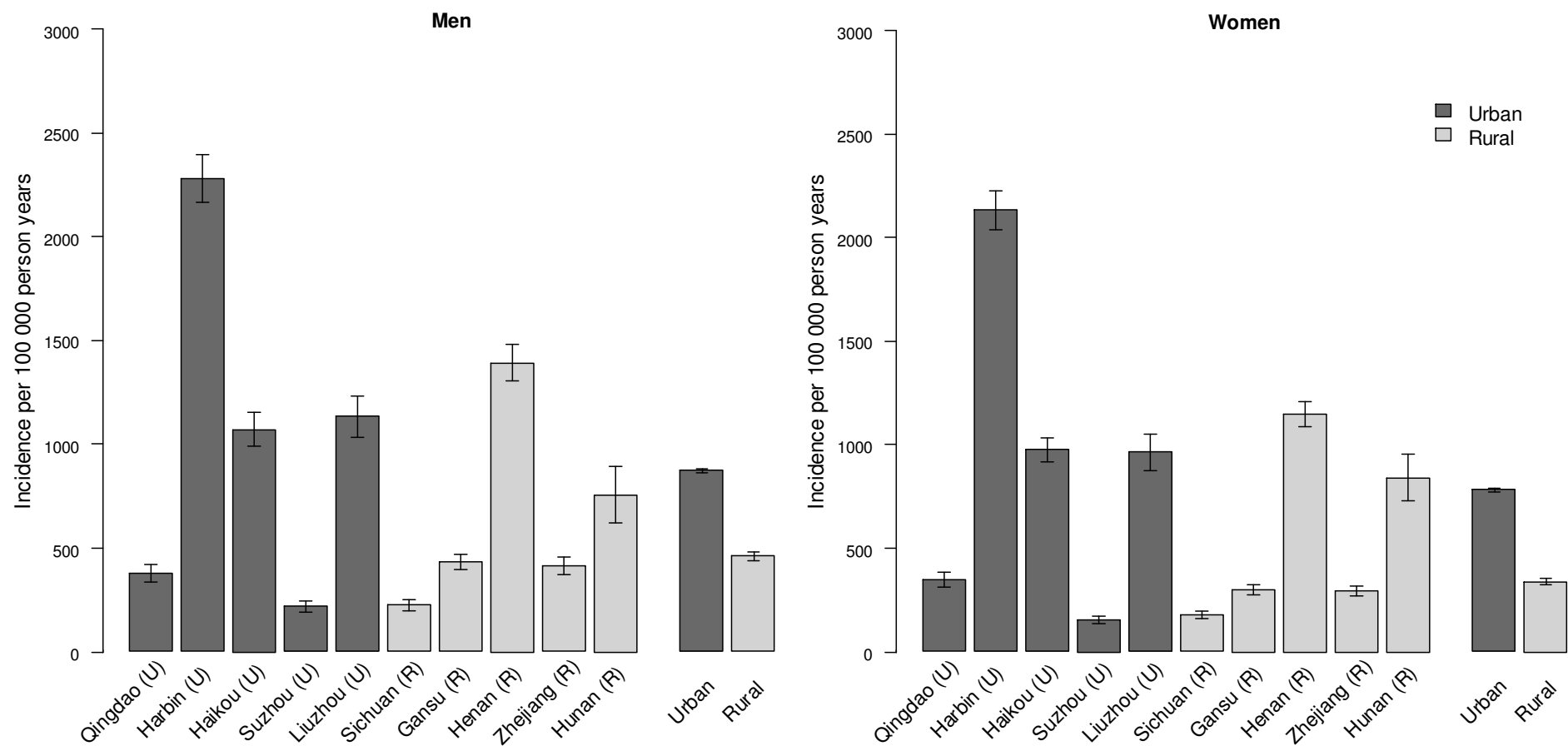


Figure 6.6. Incidence of intracerebral haemorrhage by study area

Standardised to age and education structure of study population.

Error bars represent 95% confidence intervals.

R, rural; U, urban.

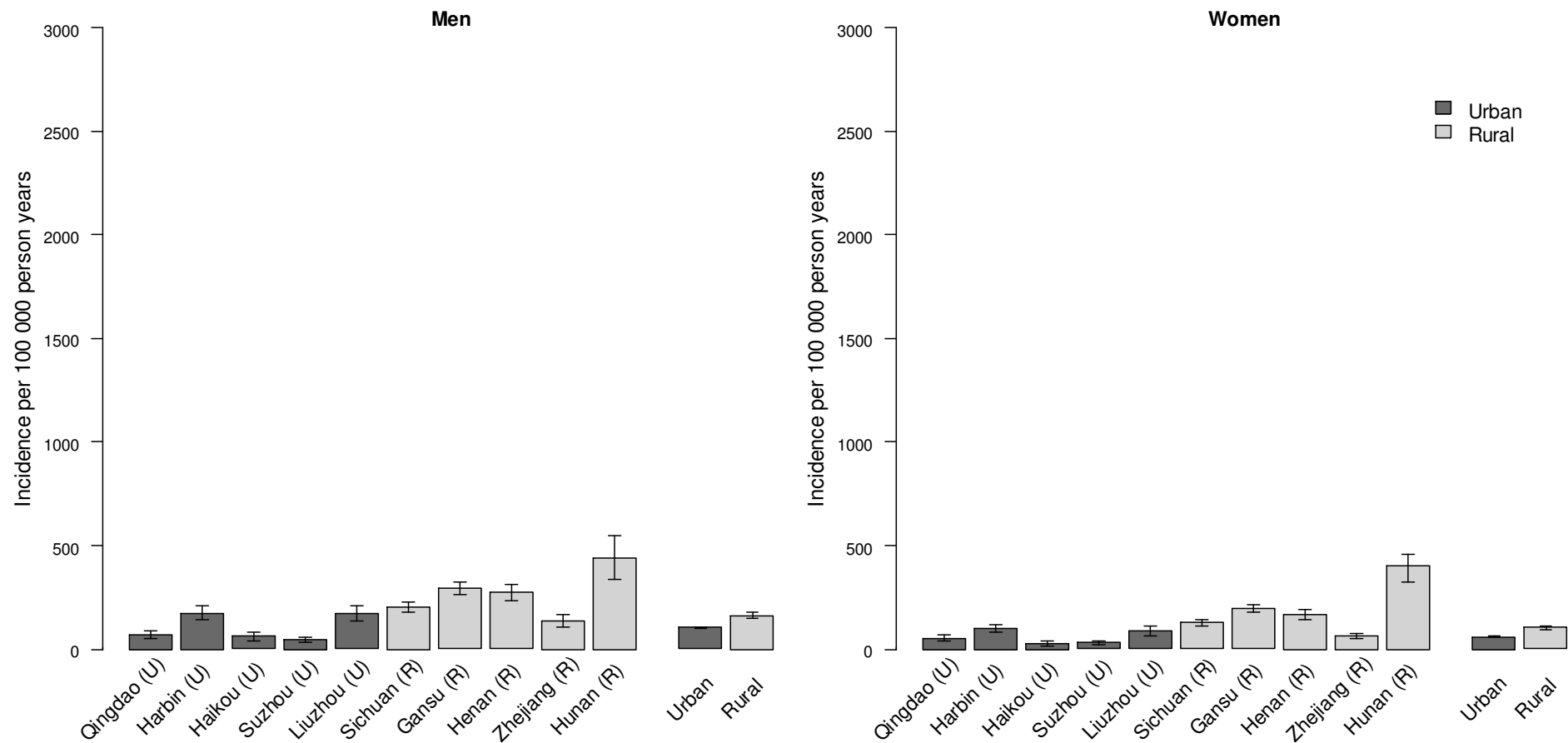


Table 6.1. Adjusted hazard ratios for incident cardiovascular diseases by self-reported diabetes status

Outcomes	Diabetes		No diabetes		Model A ^b		Model B ^c		Model C ^d	
	Events	Rate per 100 000 person years ^a	Events	Rate per 100 000 person years ^a	HR	95% CI	HR	95% CI	HR	95% CI
Cardiovascular disease mortality	571	351.1	6345	133.5	2.49	2.28-2.72	2.07	1.90-2.26	2.07	1.89-2.26
Major occlusive vascular disease	1938	1177.8	20876	536.8	2.03	1.94-2.13	1.77	1.69-1.86	1.71	1.63-1.79
Ischaemic heart disease										
Fatal myocardial infarction	127	88.1	1237	25.7	3.03	2.51-3.65	2.67	2.21-3.23	2.58	2.13-3.12
Non-fatal myocardial infarction	122	78.2	1087	27.3	2.52	2.08-3.06	2.18	1.80-2.65	2.10	1.73-2.56
Major coronary event	368	211.9	3097	69.2	2.80	2.50-3.13	2.44	2.18-2.73	2.37	2.12-2.65
Stroke										
<i>Ischaemic stroke</i>										
Fatal	43	24.5	416	8.6	2.65	1.92-3.66	2.20	1.59-3.04	2.21	1.59-3.07
Non-fatal	1589	952.6	17730	463.6	1.91	1.81-2.01	1.67	1.59-1.76	1.60	1.52-1.69
Any	1632	977.2	18146	472.1	1.92	1.82-2.02	1.68	1.60-1.77	1.62	1.53-1.70
<i>Intracerebral haemorrhage</i>										
Fatal	100	109.6	1891	42.6	2.01	1.64-2.47	1.59	1.30-1.96	1.62	1.32-2.00
Non-fatal	92	66.3	2340	61.8	1.22	0.99-1.51	1.00	0.81-1.23	1.00	0.81-1.23
Any	192	175.9	4231	104.4	1.54	1.33-1.78	1.24	1.07-1.44	1.25	1.08-1.45
<i>Total stroke</i>										
Fatal	168	154.8	2575	56.9	2.20	1.88-2.58	1.78	1.52-2.09	1.81	1.54-2.13
Non-fatal	1749	1068.6	21157	553.5	1.84	1.75-1.94	1.60	1.52-1.68	1.54	1.47-1.62
Any	1917	1223.3	23732	610.4	1.87	1.78-1.96	1.61	1.54-1.69	1.56	1.49-1.64

^aAge-, sex- and study area- standardised rates; ^bModel A: stratified by age, sex and study area; ^cModel B: additionally adjusted for education, smoking, alcohol, physical activity and systolic blood pressure; ^dModel C: additionally adjusted for waist-to-hip ratio.

Events classified as fatal if death from any cause recorded within 28 days of disease onset.

CI, confidence interval; HR, hazard ratio.

Figure 6.7. Age- and sex-specific adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease by self-reported diabetes status

Stratified by age, sex (where appropriate) and study area and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs). The open diamond represents the overall HR and its 95% CI.

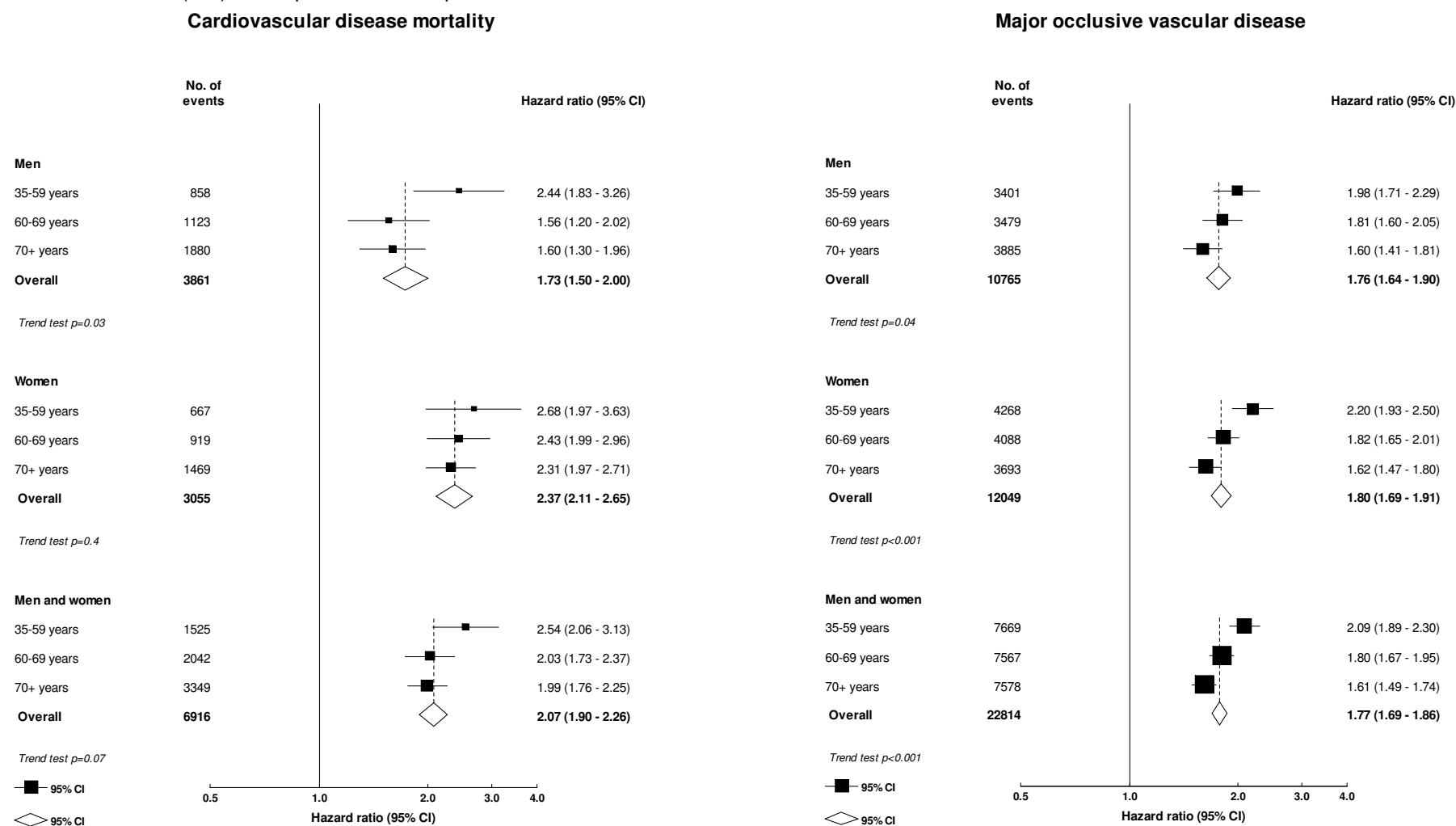


Figure 6.8. Age- and sex-specific adjusted hazard ratios for major coronary event and ischaemic stroke by self-reported diabetes status

Stratified by age, sex (where appropriate) and study area and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs). The open diamond represents the overall HR and its 95% CI.

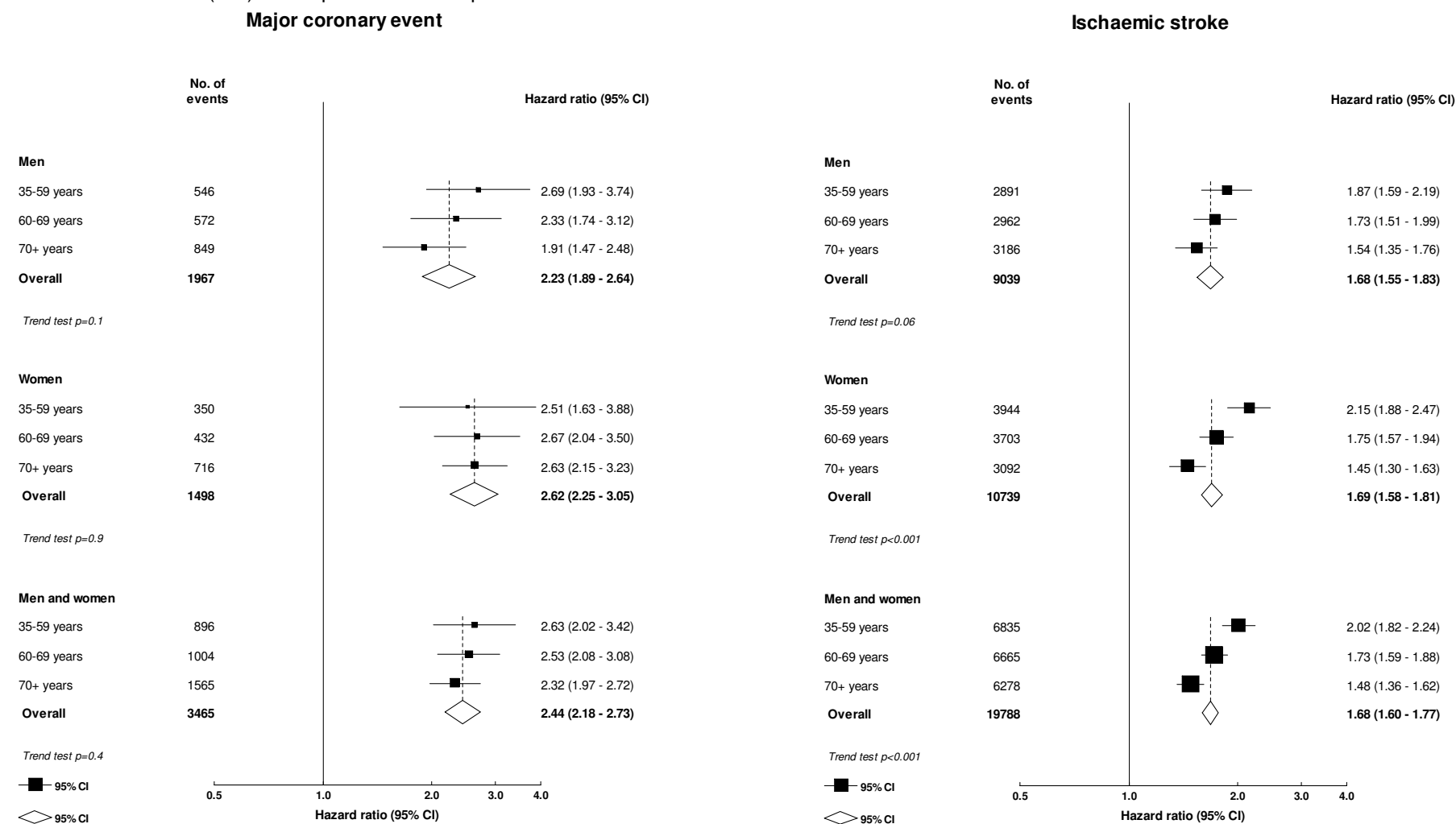


Figure 6.9. Sex-specific adjusted hazard ratios for incident cardiovascular diseases by self-reported diabetes status

Stratified by age and study area and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs).

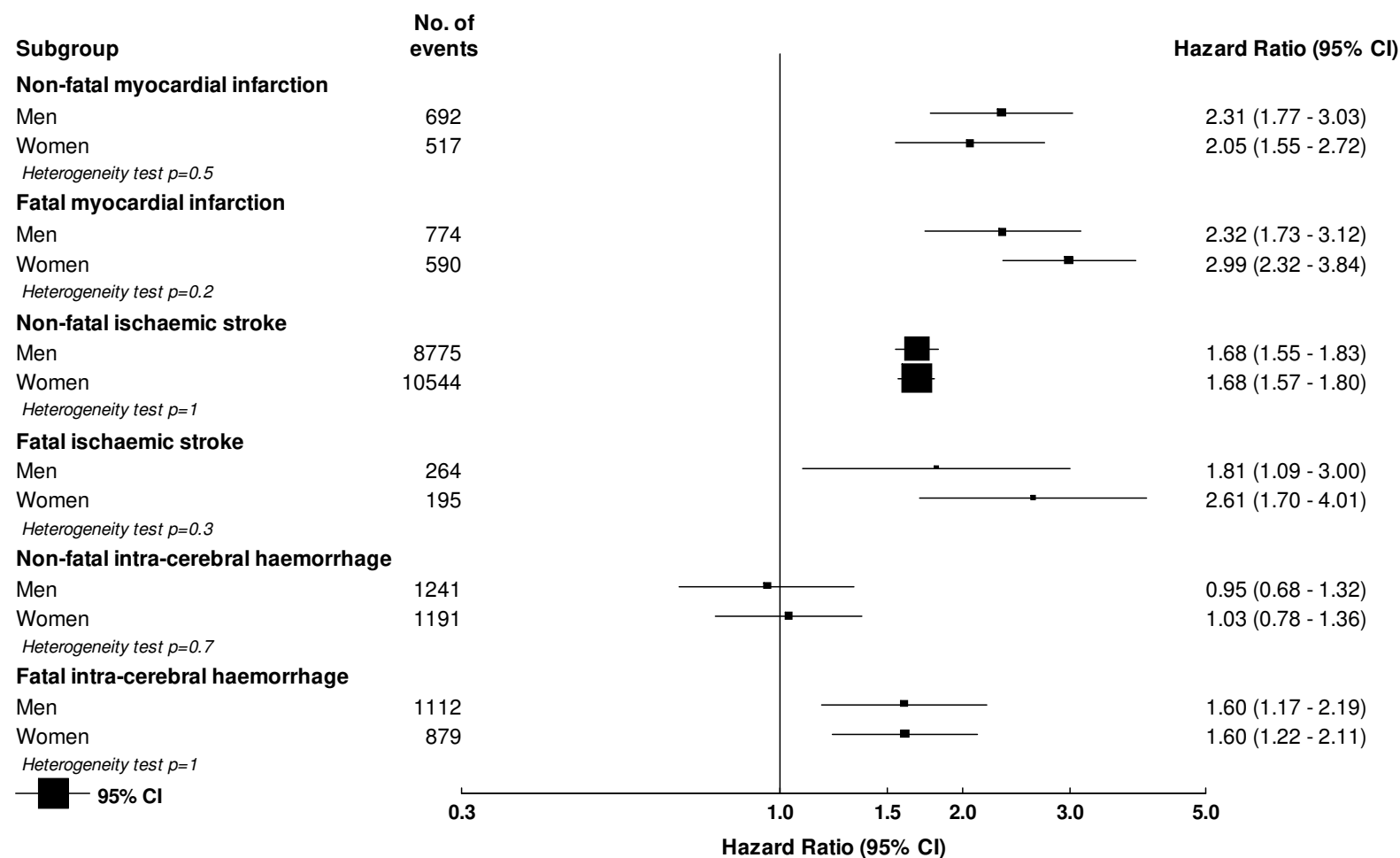


Figure 6.10. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease by self-reported diabetes status progressively adjusted for potential confounding factors

Squares represent the hazard ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs).

BMI, body mass index; DBP, diastolic blood pressure; SBP, systolic blood pressure.

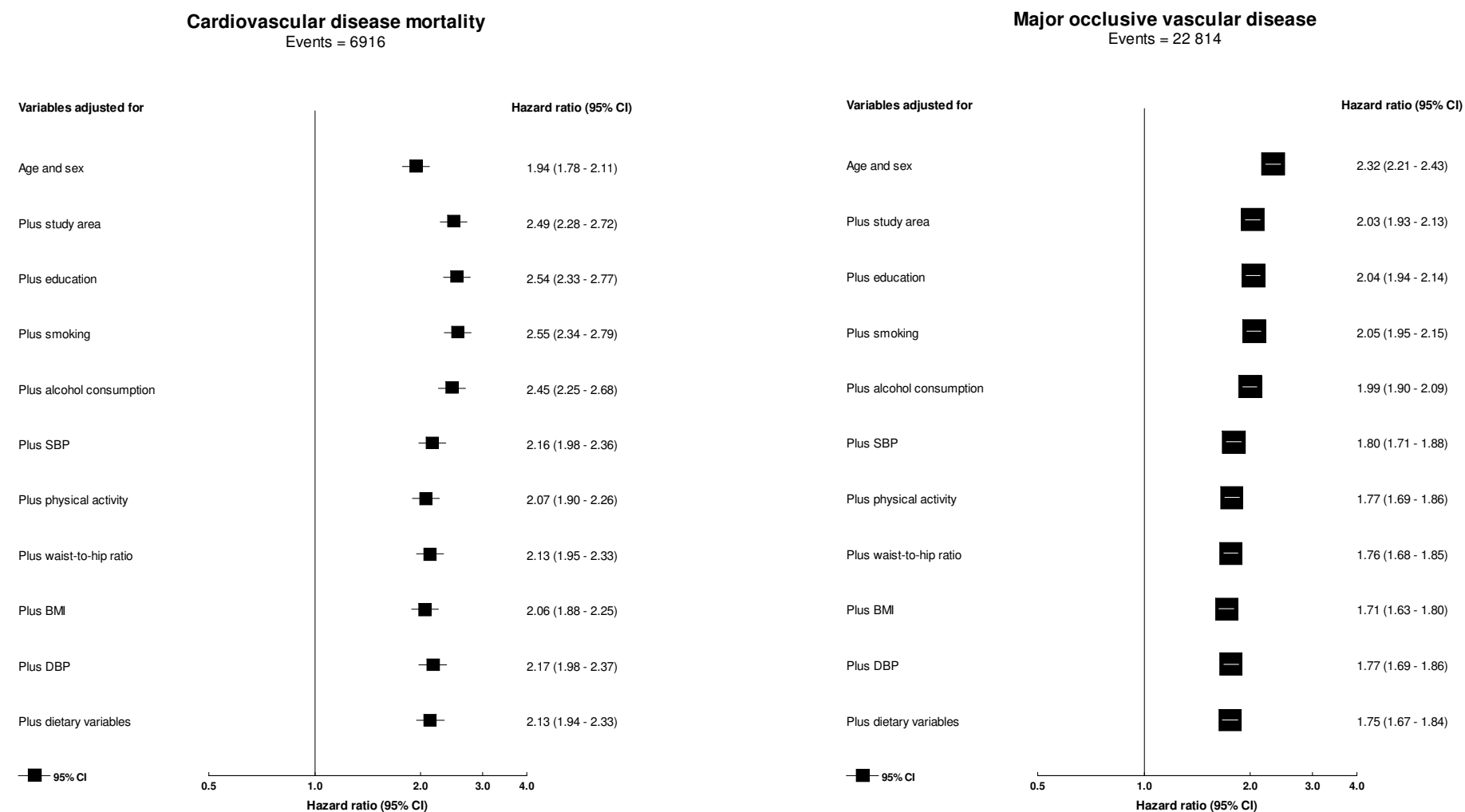


Figure 6.11. Adjusted hazard ratios for major coronary event and ischaemic stroke by self-reported diabetes status progressively adjusted for potential confounding factors

Squares represent the hazard ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs).

BMI, body mass index; DBP, diastolic blood pressure; SBP, systolic blood pressure.

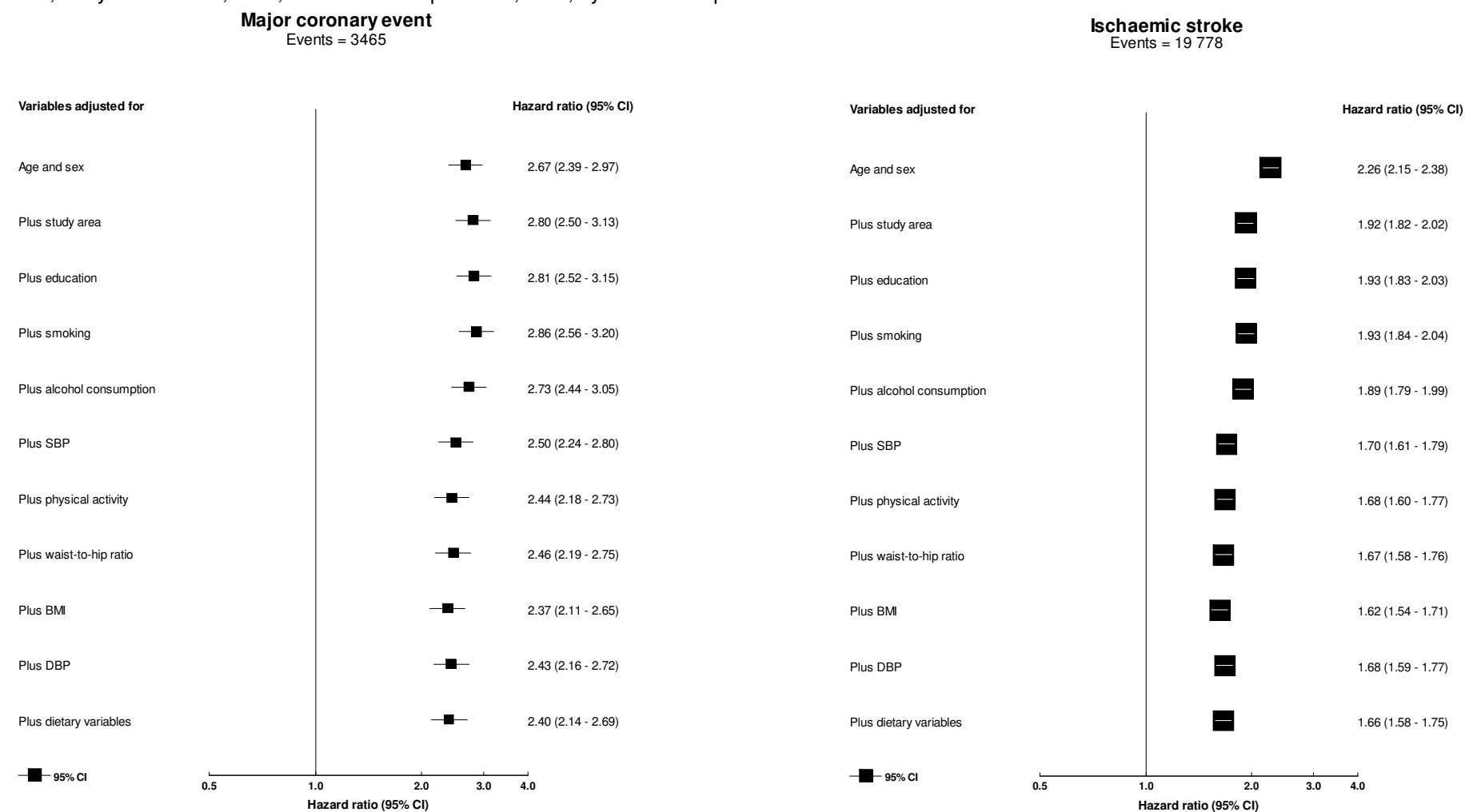


Figure 6.12. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease by self-reported diabetes status, by study area

Stratified by age and sex and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall HR. The open diamond represents the overall HR and its 95% CI.

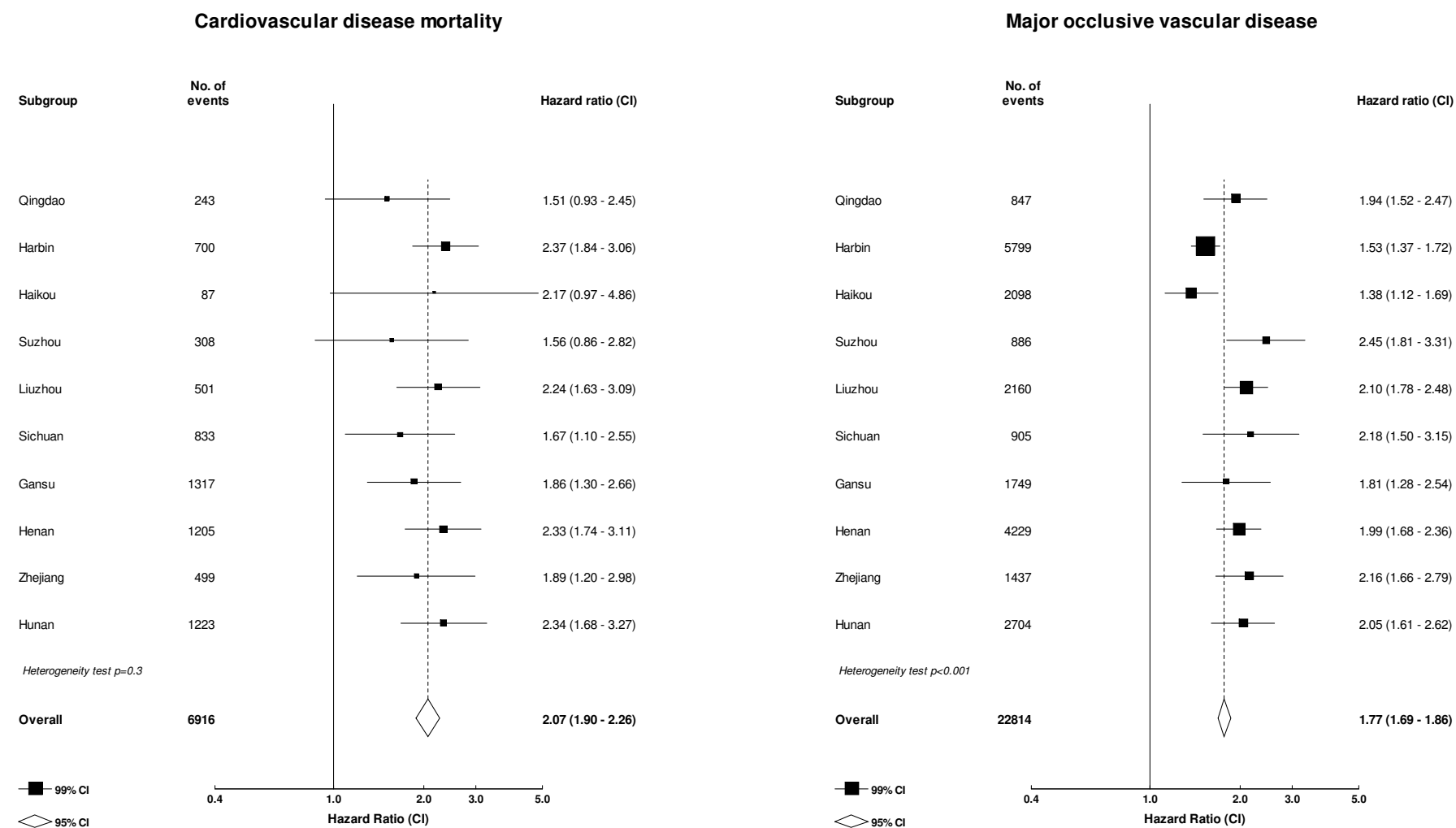


Figure 6.13. Adjusted hazard ratios for major coronary event and ischaemic stroke by self-reported diabetes status, by study area

Stratified by age and sex and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall HR. The open diamond represents the overall HR and its 95% CI.

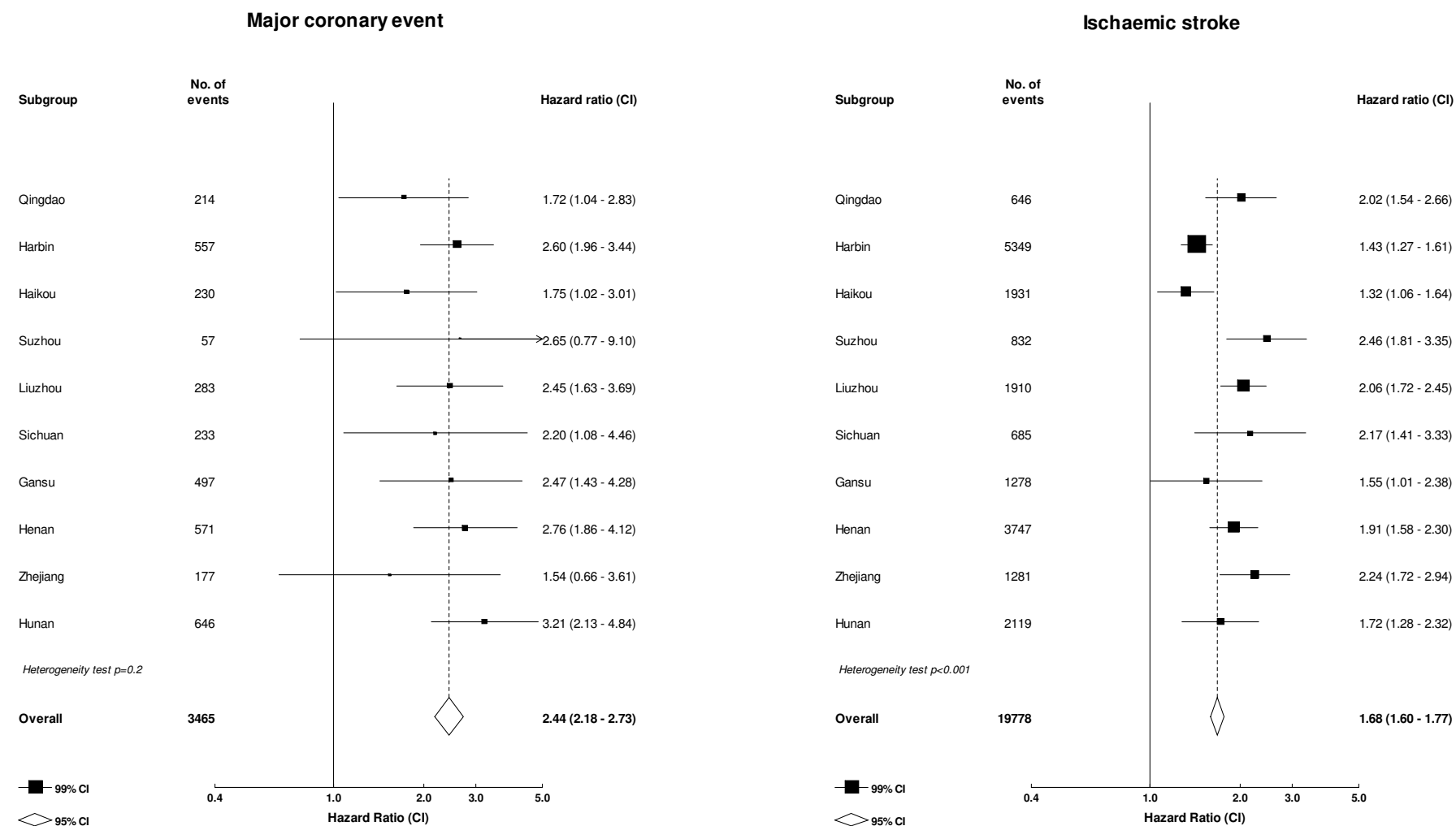


Figure 6.14. Adjusted hazard ratios for cardiovascular disease mortality by self-reported diabetes status in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratios (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

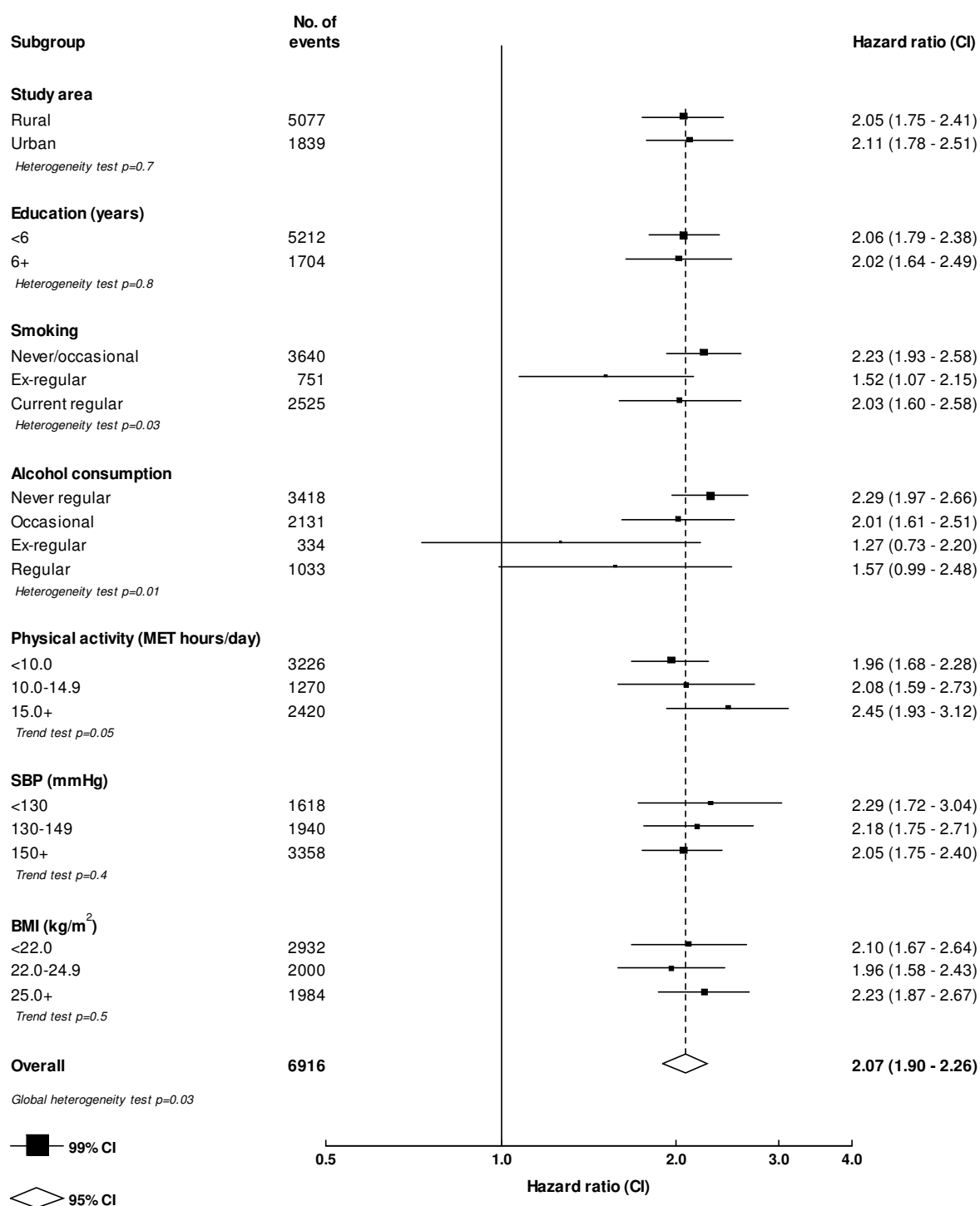


Figure 6.15. Adjusted hazard ratios for major occlusive vascular disease by self-reported diabetes status in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratios (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

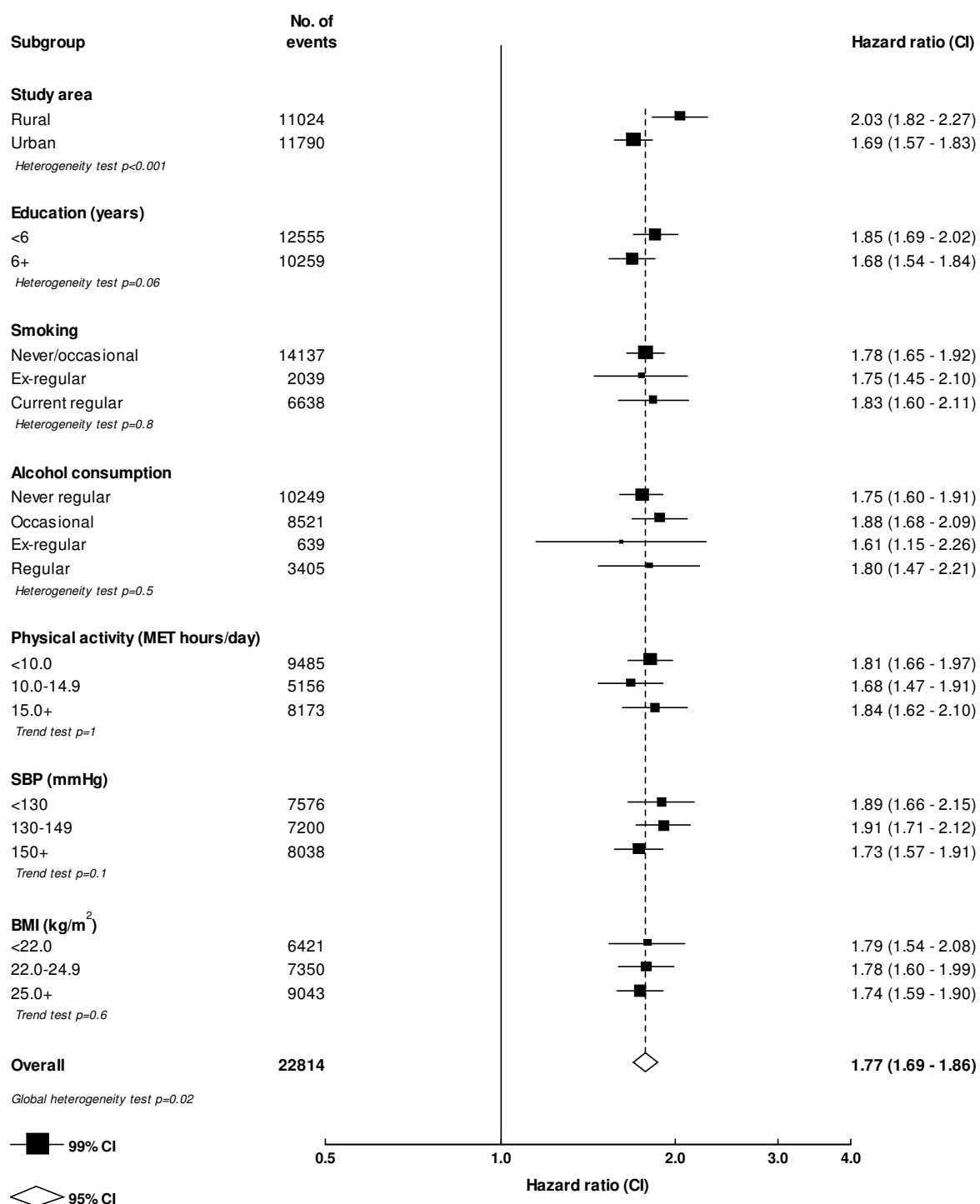


Figure 6.16. Adjusted hazard ratios for ischaemic stroke by self-reported diabetes status in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratios (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

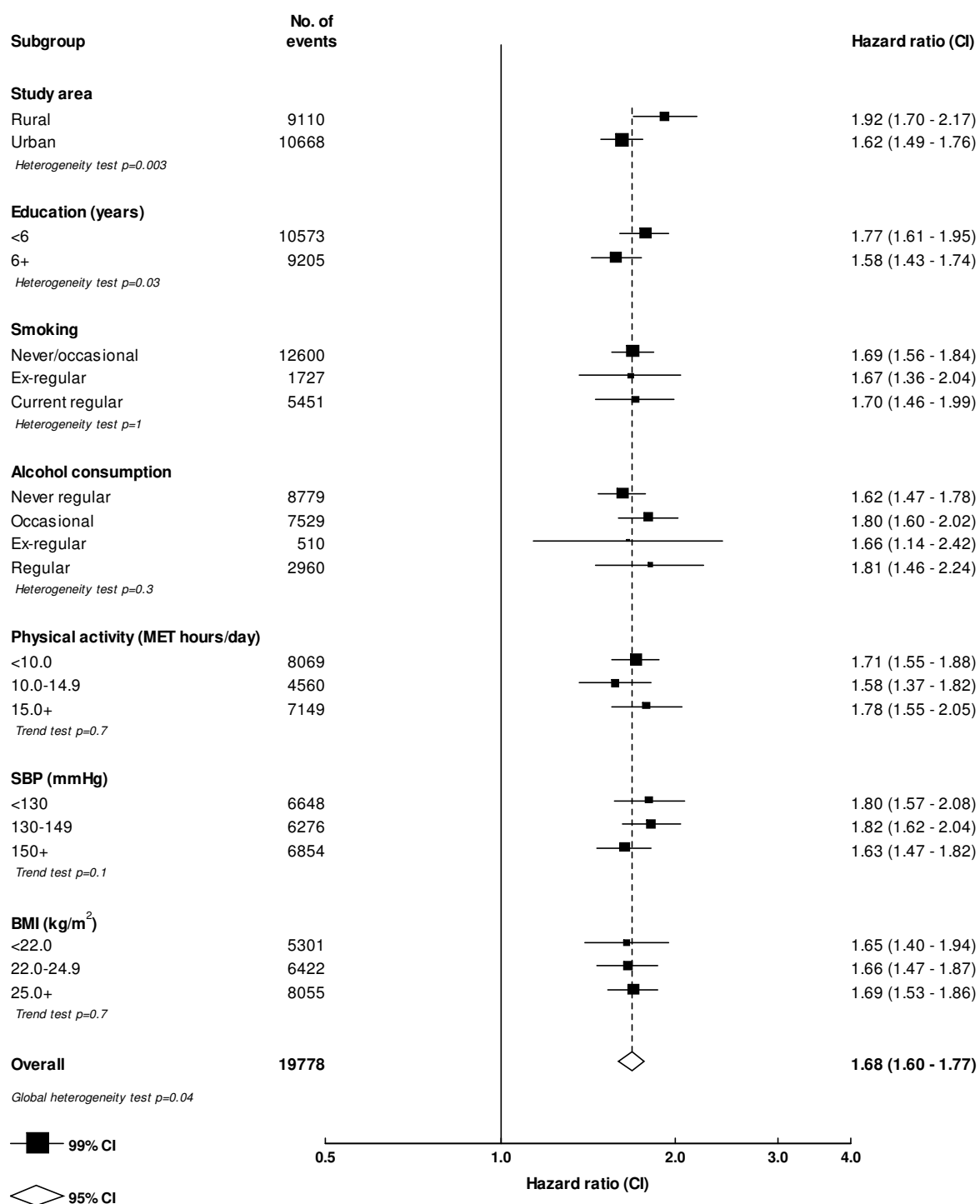


Figure 6.17. Adjusted hazard ratios for major coronary event by self-reported diabetes status in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratios (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

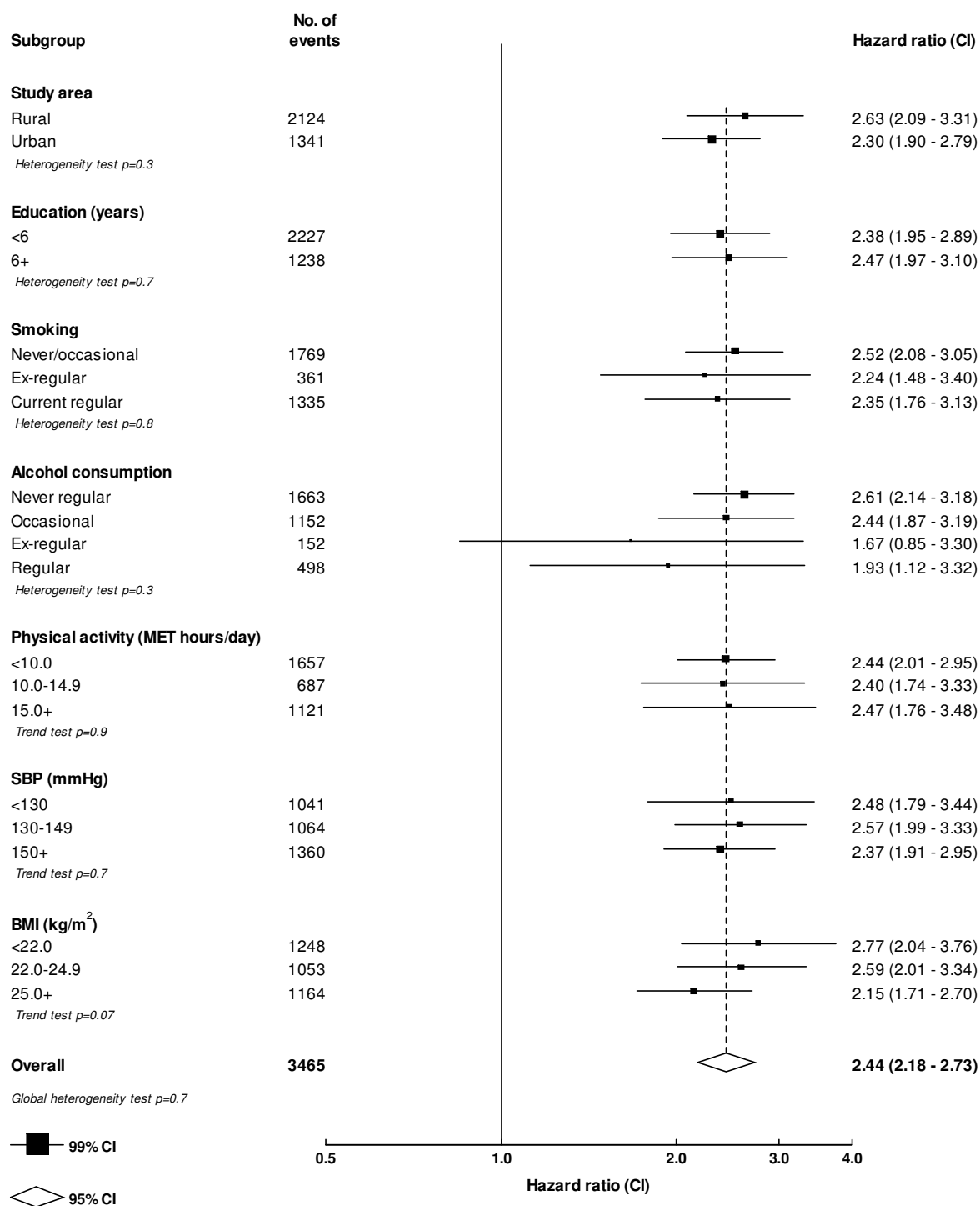


Figure 6.18. Age- and sex-specific adjusted hazard ratios for intracerebral haemorrhage by self-reported diabetes status

Stratified by age, sex (where appropriate) and study area and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs). The open diamond represents the overall HR and its 95% CI.

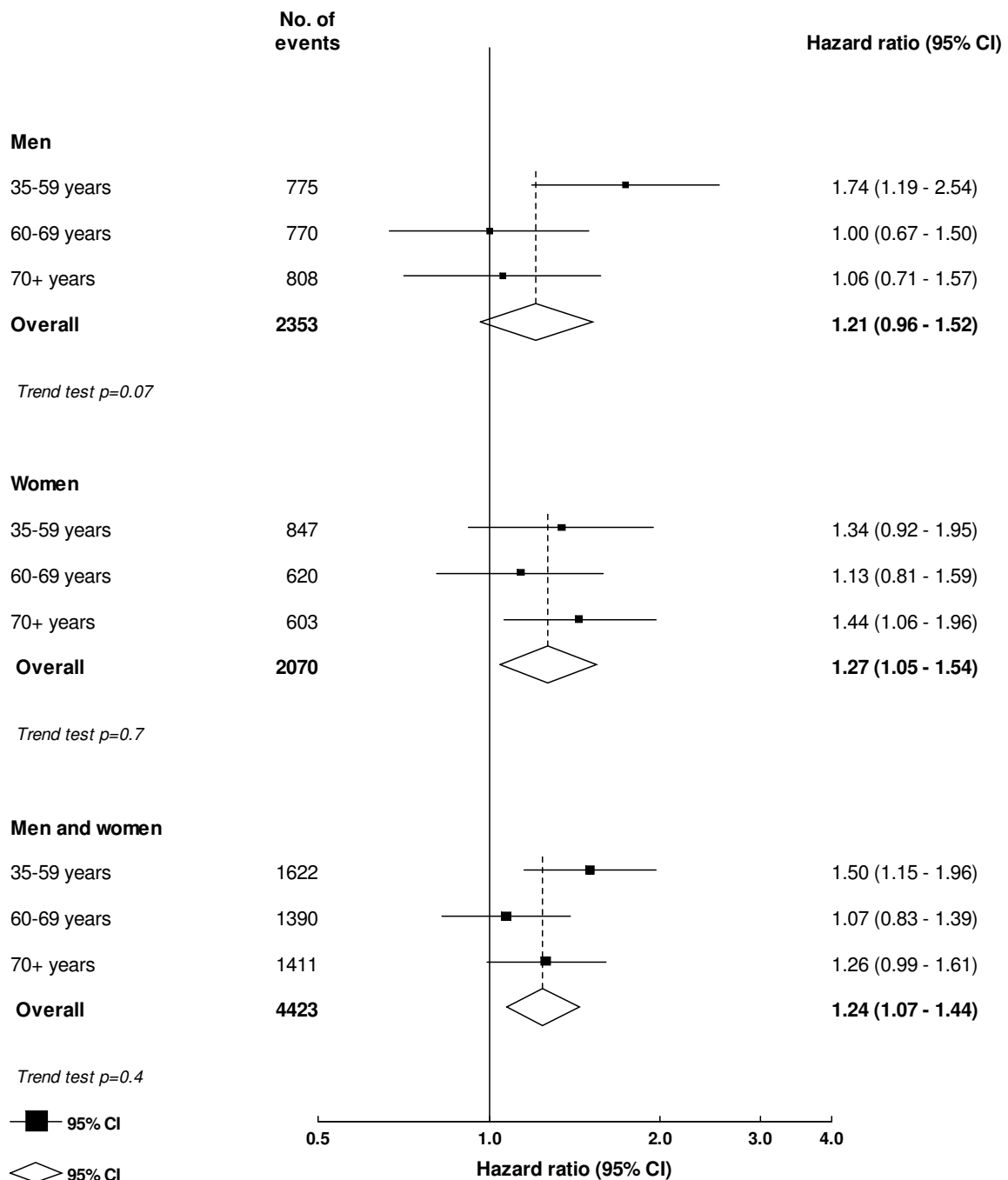


Figure 6.19. Adjusted hazard ratios for intracerebral haemorrhage by self-reported diabetes status, by study area

Stratified by age and sex and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall HR. The open diamond represents the overall HR and its 95% CI.

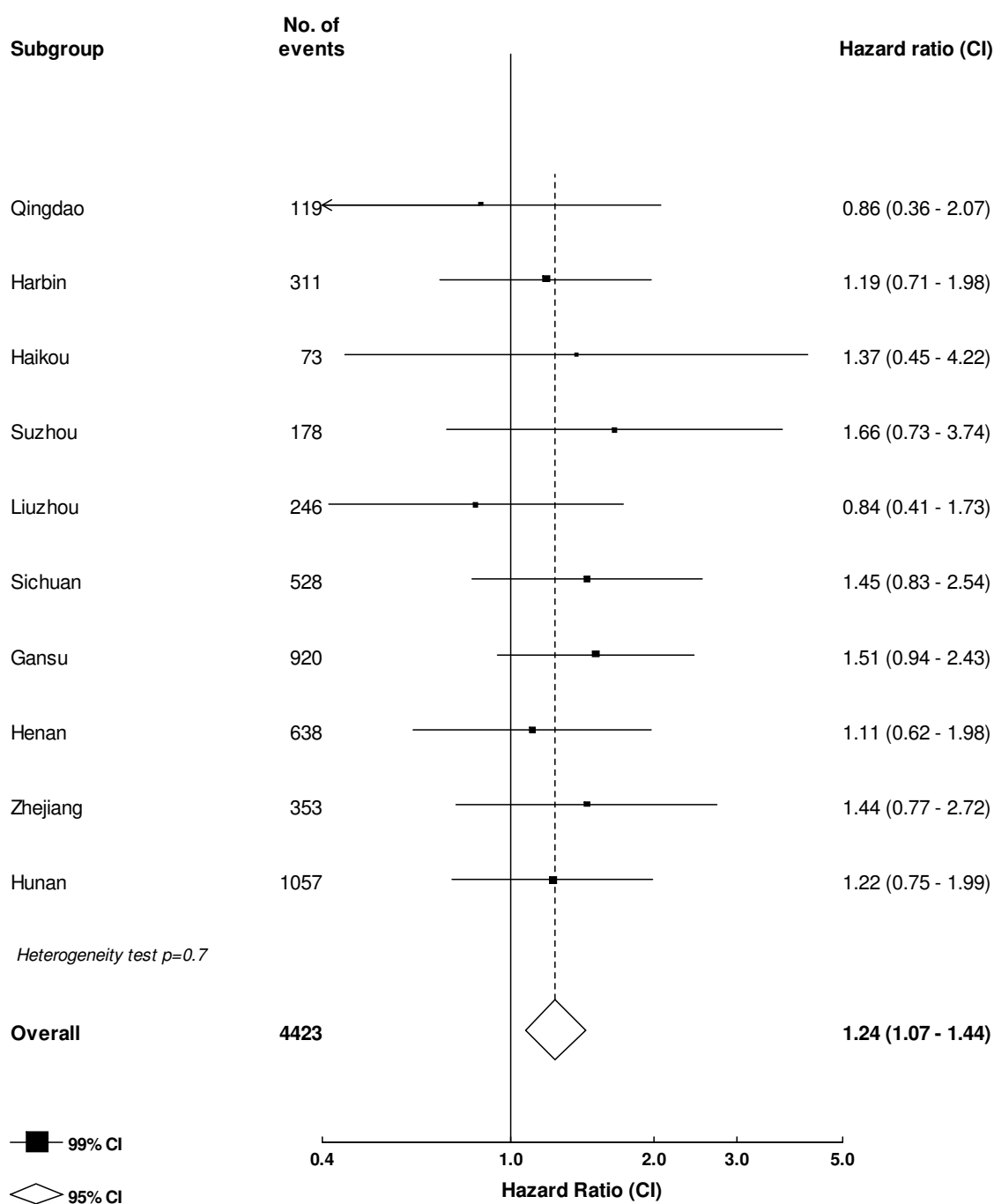


Figure 6.20. Adjusted hazard ratios for intracerebral haemorrhage by self-reported diabetes status progressively adjusted for potential confounding factors

Squares represent the hazard ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs).

BMI, body mass index; DBP, diastolic blood pressure; SBP, systolic blood pressure.

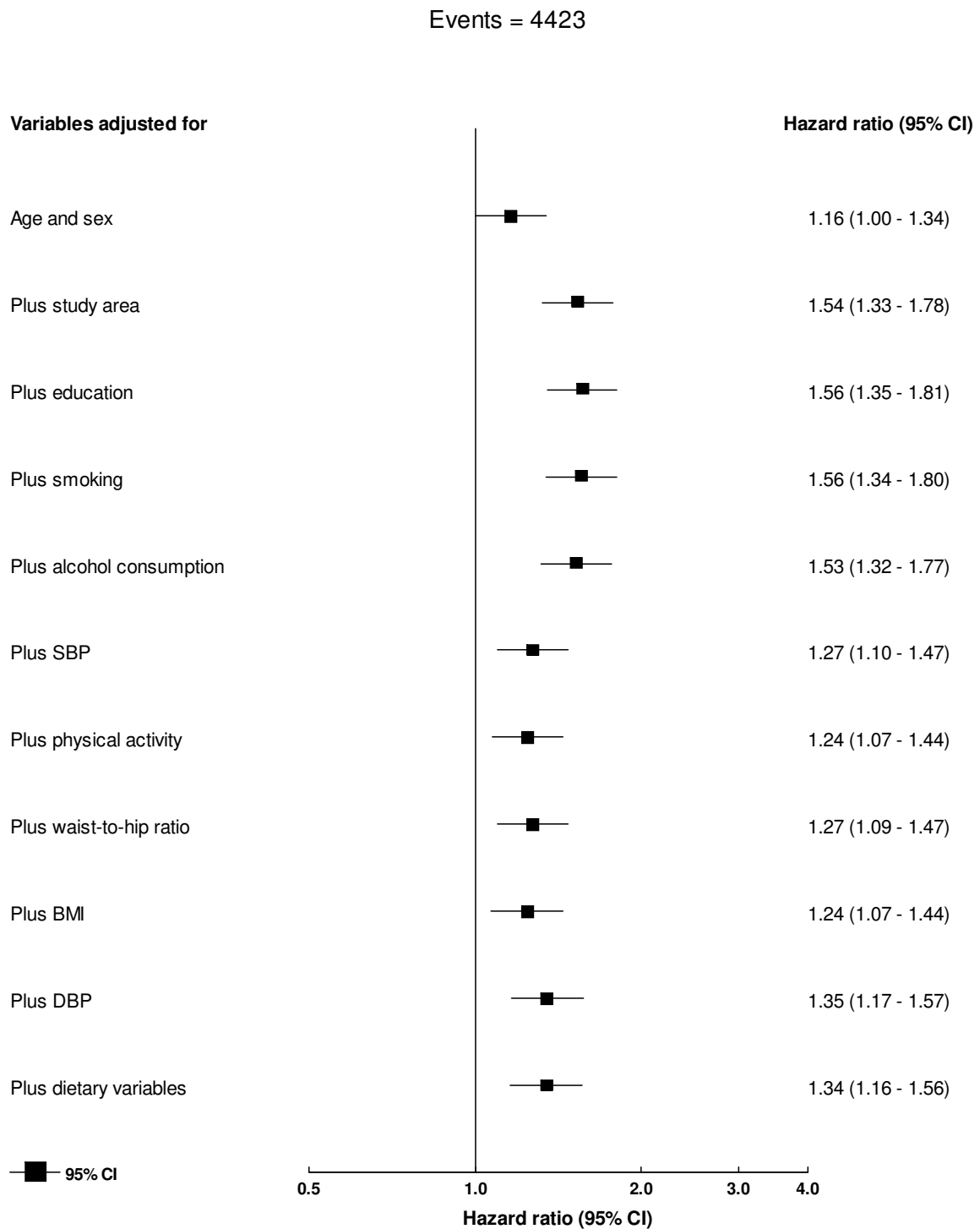


Figure 6.21. Adjusted hazard ratios for intracerebral haemorrhage by self-reported diabetes status in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

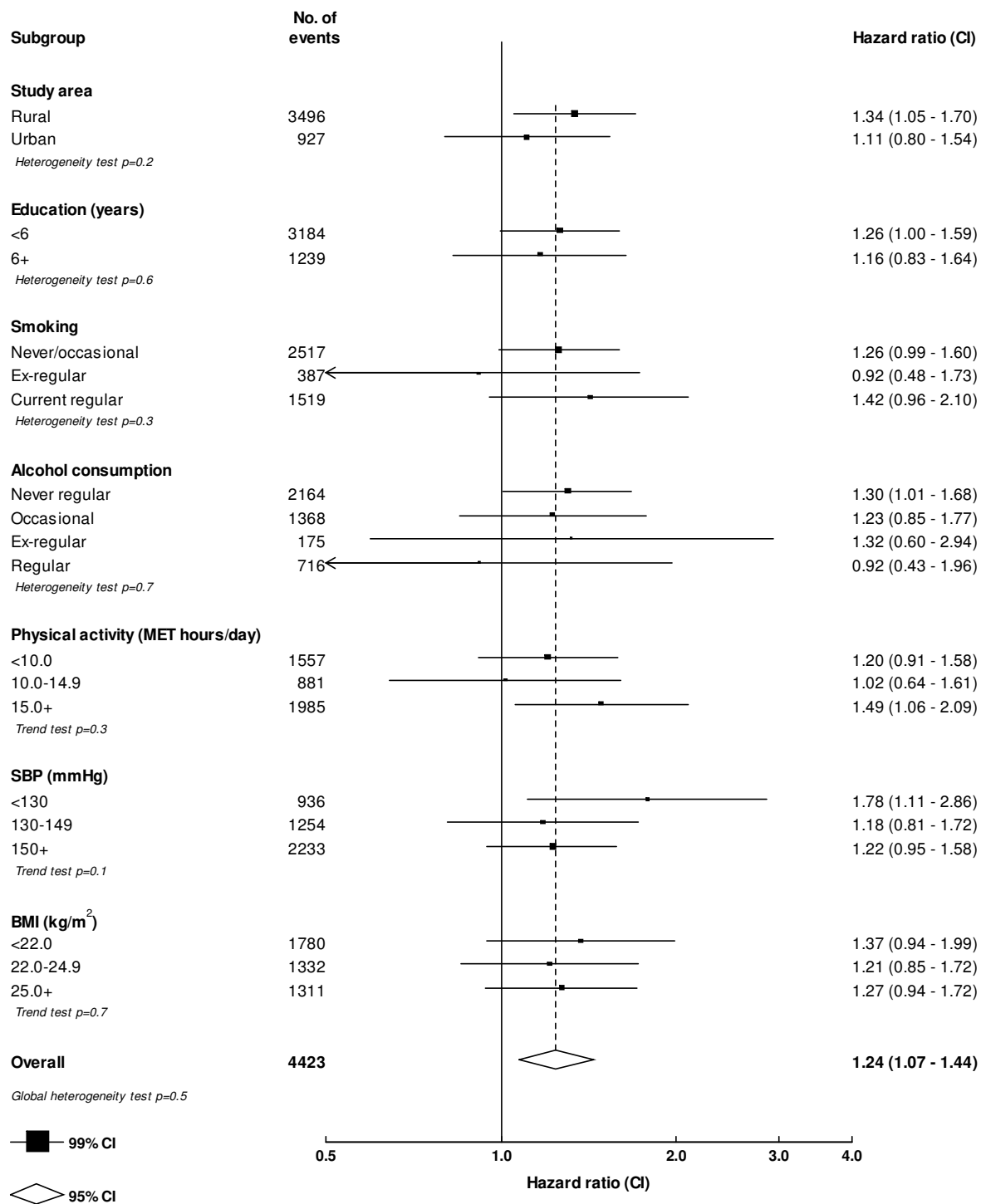


Figure 6.22. Adjusted hazard ratios for cardiovascular diseases by duration of diabetes diagnosis

Stratified by age, sex and study area and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Squares represent the hazard ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs).

Reference group: individuals without diabetes. Trend tests based on individuals with self-reported diabetes.

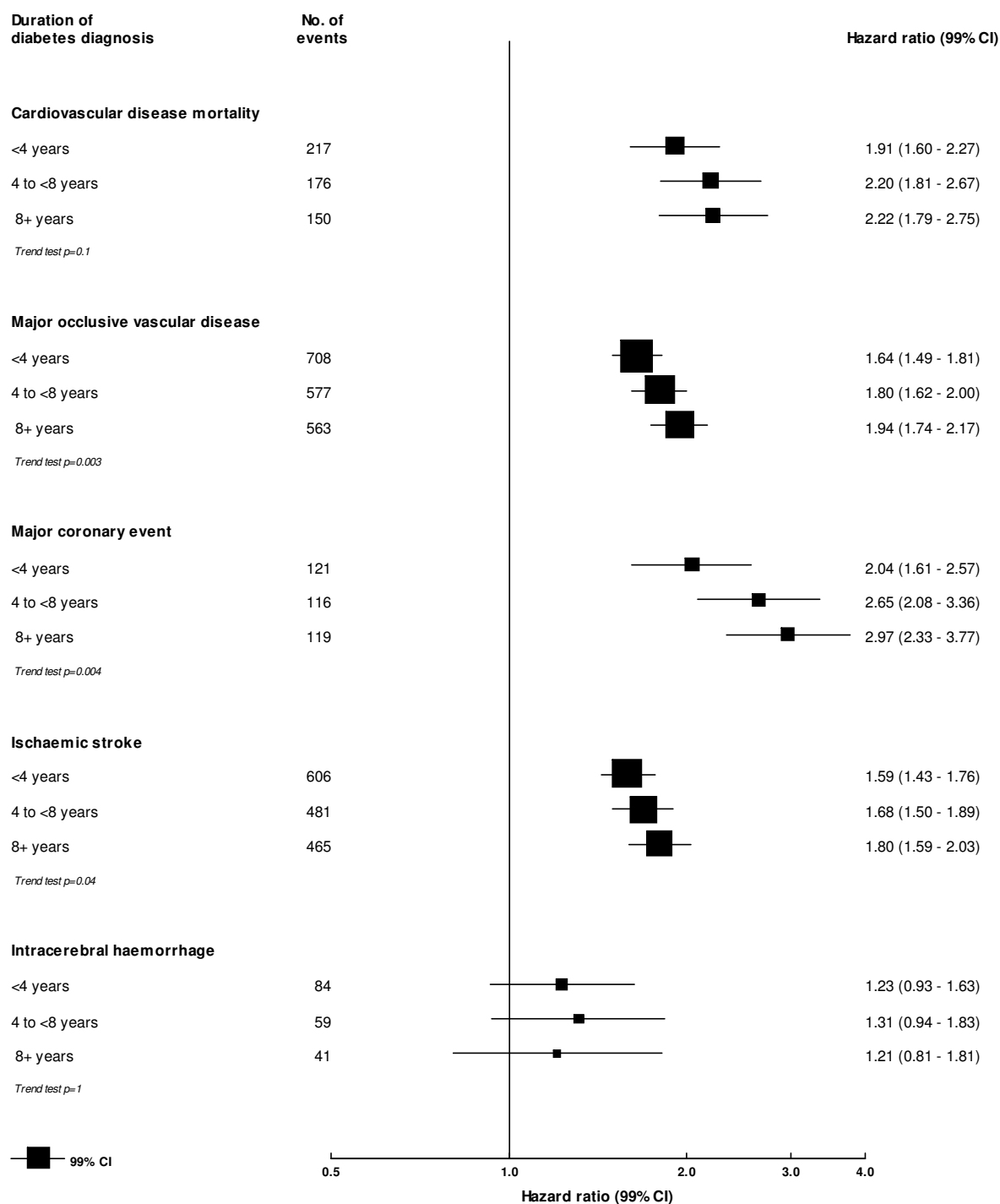


Figure 6.23. Adjusted hazard ratios for cardiovascular diseases by hypoglycaemic medication treatment status in self-reported diabetes

Stratified by age, sex and study area and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Squares represent the hazard ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). Reference group: individuals without diabetes. Heterogeneity tests based on individuals with self-reported diabetes. Treatment: oral hypoglycaemic agents or insulin.

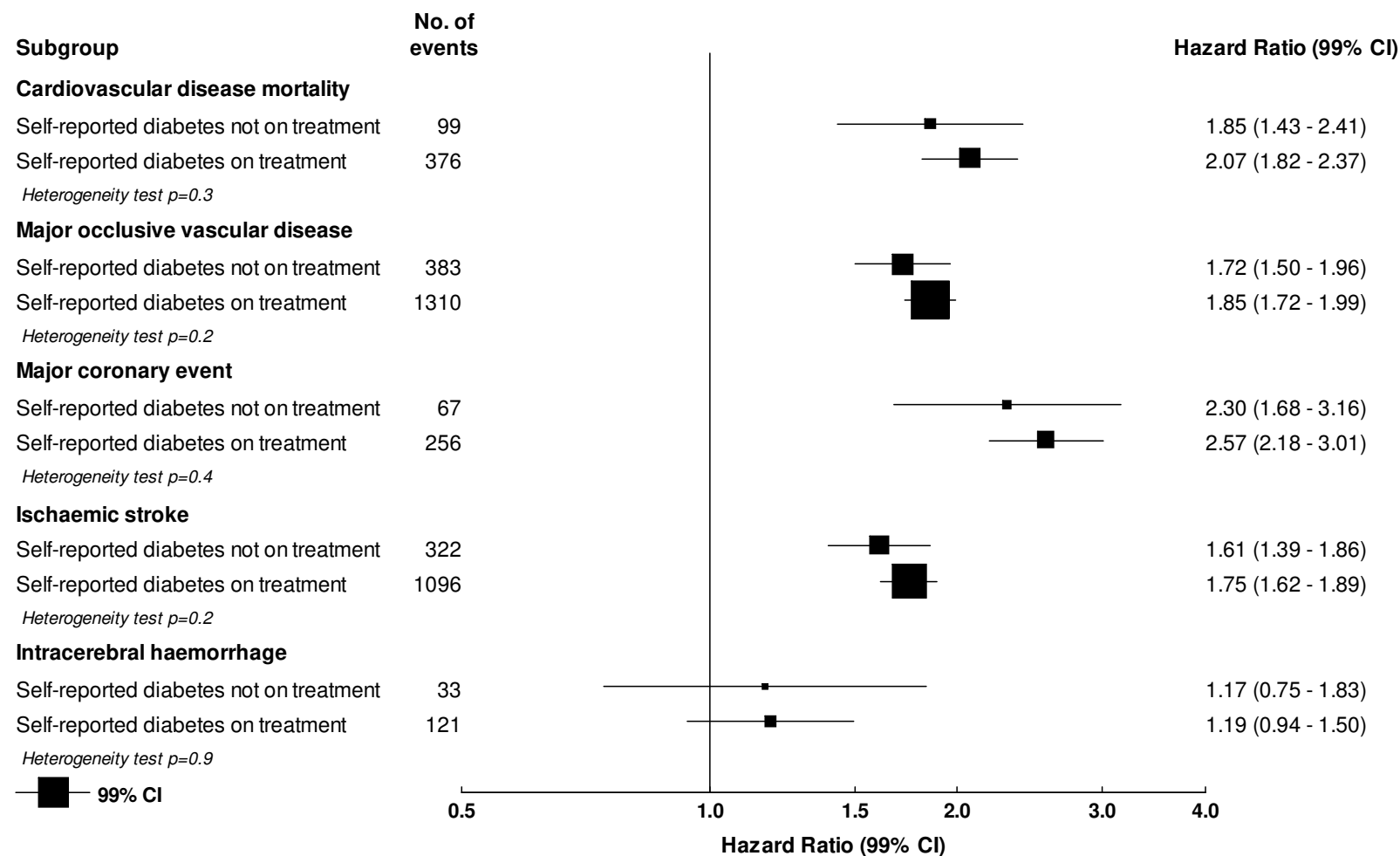


Figure 6.24. Adjusted hazard ratios for cardiovascular diseases by baseline plasma glucose level in self-reported diabetes

Stratified by age, sex and study area and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Squares represent the hazard ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs).

Reference group: individuals without diabetes. Heterogeneity tests based on individuals with self-reported diabetes.

“Good control”: blood glucose level <7.2 mmol/L if time since last food was ≥8 hours or a blood glucose level <10.0 mmol/L if time since last food was <8 hours; “Poor control”: blood glucose level ≥7.2 mmol/L if time since last food was ≥8 hours or a blood glucose level ≥10.0 mmol/L if time since last food was <8 hours.

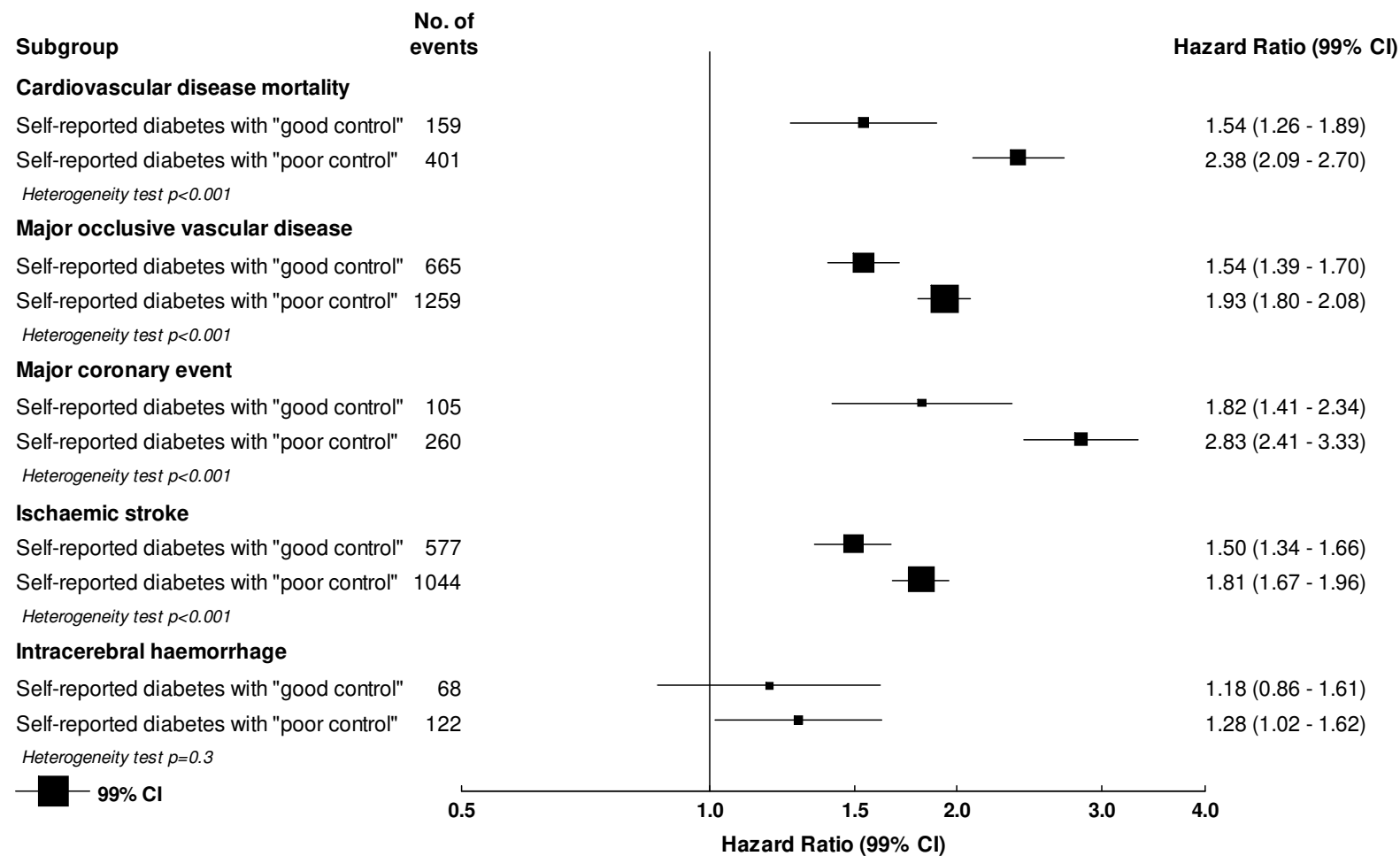


Table 6.2. Adjusted hazard ratios for incident cardiovascular diseases by screen-detected diabetes status

Outcomes	Diabetes		No diabetes		Model A ^b		Model B ^c		Model C ^d	
	Events	Rate per 100 000 person years ^a	Events	Rate per 100 000 person years ^a	HR	95%CI	HR	95%CI	HR	95%CI
Cardiovascular disease mortality	437	294.0	6345	133.5	1.98	1.80-2.18	1.66	1.51-1.83	1.66	1.51-1.83
Major occlusive vascular disease	1356	1021.1	20876	536.8	1.73	1.64-1.83	1.50	1.41-1.58	1.44	1.36-1.52
Ischaemic heart disease										
Fatal myocardial infarction	81	60.6	1237	25.7	1.90	1.52-2.39	1.68	1.34-2.12	1.61	1.28-2.03
Non-fatal myocardial infarction	68	52.3	1087	27.3	1.70	1.33-2.18	1.51	1.18-1.93	1.44	1.13-1.85
Major coronary event	213	140.0	3097	69.2	1.84	1.60-2.11	1.62	1.40-1.86	1.57	1.36-1.81
Stroke										
<i>Ischaemic stroke</i>										
Fatal	33	19.3	416	8.6	2.19	1.53-3.13	1.85	1.29-2.64	1.84	1.28-2.64
Non-fatal	1144	869.5	17730	463.6	1.70	1.60-1.81	1.47	1.39-1.57	1.41	1.33-1.50
Any	1177	888.8	18146	472.1	1.71	1.62-1.82	1.48	1.40-1.57	1.42	1.34-1.51
<i>Intracerebral haemorrhage</i>										
Fatal	116	96.8	1891	42.6	2.02	1.67-2.44	1.62	1.34-1.96	1.65	1.36-2.00
Non-fatal	73	69.1	2340	61.8	1.05	0.83-1.32	0.81	0.64-1.03	0.81	0.64-1.03
Any	189	165.9	4231	104.4	1.49	1.28-1.72	1.17	1.01-1.36	1.18	1.02-1.37
<i>Total stroke</i>										
Fatal	168	129.1	2575	56.9	2.05	1.75-2.40	1.66	1.42-1.95	1.69	1.44-1.98
Non-fatal	1268	981.7	21157	553.5	1.63	1.54-1.73	1.39	1.32-1.48	1.34	1.27-1.42
Any	1436	1110.8	23732	610.4	1.67	1.59-1.77	1.42	1.34-1.49	1.37	1.30-1.45

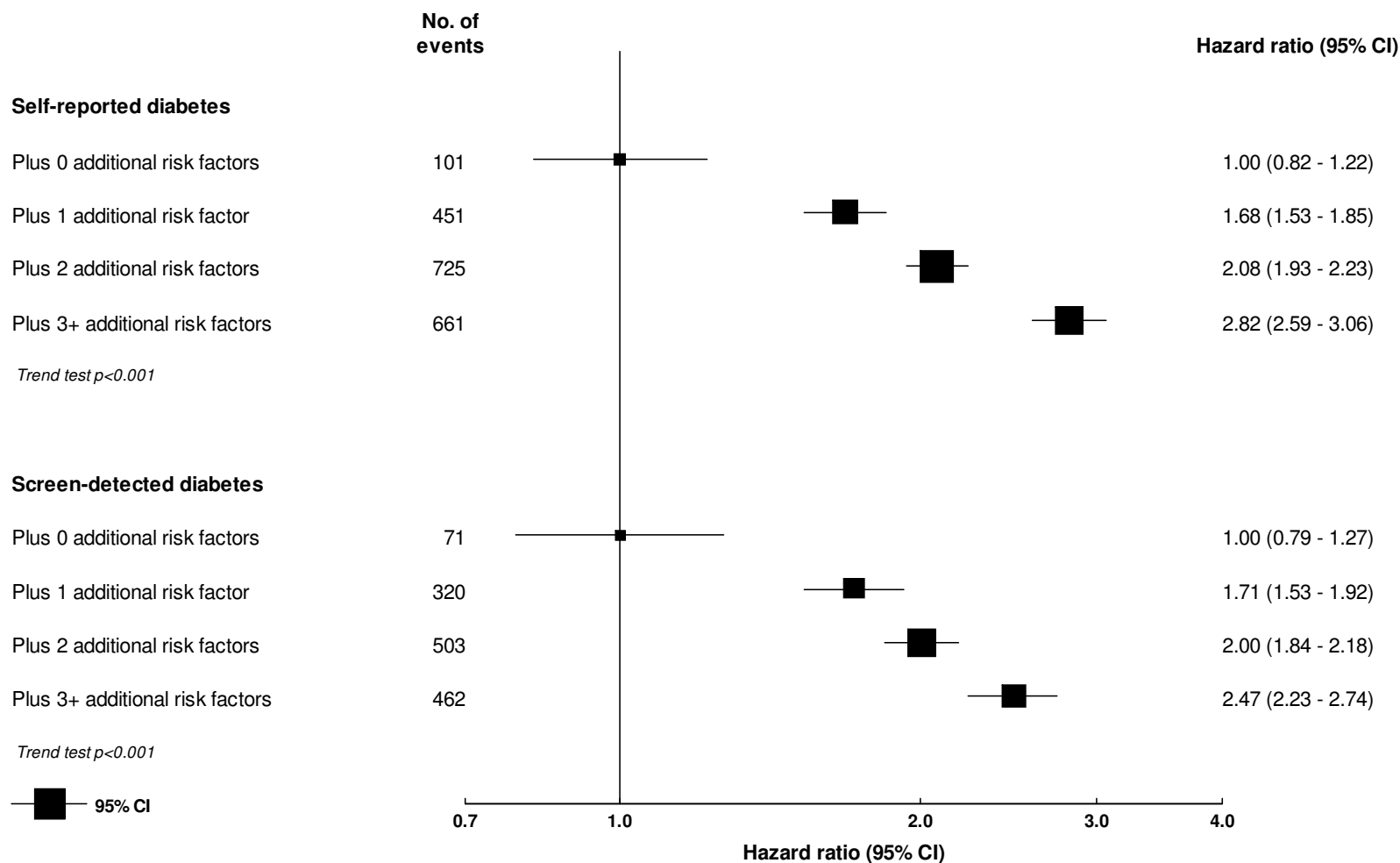
^aAge-, sex- and study area- standardised rates; ^bModel A: stratified by age, sex and study area; ^cModel B: additionally adjusted for education, smoking, alcohol, physical activity and systolic blood pressure; ^dModel C: additionally adjusted for waist-to-hip ratio.

Events classified as fatal if death from any cause recorded within 28 days of disease onset.

CI, confidence interval; HR, hazard ratio.

Figure 6.25. Adjusted hazard ratios for major occlusive vascular disease by number of cardiovascular risk factors at baseline among individuals with diabetes

Stratified by age, sex and study area and adjusted for education and alcohol. Cardiovascular risk factors: hypertension (self-reported hypertension, mean systolic blood pressure ≥ 140 mmHg or diastolic blood pressure ≥ 90 mmHg), overweight or obese (body mass index ≥ 25 kg/m²), ever regular smoking, physical inactivity (<10 metabolic equivalent of task hours/day). Squares represent the hazard ratio (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 95% confidence intervals (CI).



Chapter 7. Plasma glucose and the risk of incident cardiovascular disease

7.1 Background

As described in Chapter 2, and further confirmed in analyses of data from the CKB (Chapters 5 and 6), diabetes is a major, independent risk factor for CVD in both Western and Chinese populations.^{6, 90} It is unclear, however, whether the glycaemic threshold for diabetes, defined on the basis of an elevated risk of microvascular complications,⁷⁷ is also appropriate for characterising macrovascular disease risk.²⁴⁸ There is reasonably good evidence, particularly from Western populations, that individuals with blood glucose levels in the pre-diabetes range (ie, FPG 6.1-6.9 mmol/L or post-load plasma glucose 7.8-11.0 mmol/L⁷⁸) also have elevated CVD risks when compared with individuals with lower blood glucose levels.³⁰⁶ Even within the pre-diabetes range, however, uncertainty remains about the magnitude of risk and the relevance of other factors, such as age, blood pressure and adiposity to this risk. Furthermore, as discussed in Chapter 2, substantial uncertainty remains as to whether lower levels of blood glucose are associated with lower risk of CVDs below the glycaemic threshold for pre-diabetes,²⁴⁸ and, if so, whether the association is continuous¹⁰³ or whether a threshold exists at a lower blood glucose level.^{6, 102} The documented global increase in mean blood glucose levels⁷⁹ highlights the importance of a better understanding of these. Moreover, most previous studies have focussed on FBG, with limited evidence about the role of RBG in CVD risk. Given we spend most of our time in a non-fasting state, RBG may be a useful and more practical measure for assessing the relevance of glucose levels to disease risk.¹⁴⁸

As described, there has been a rapid and marked increase in the prevalence of diabetes in China over recent decades^{86, 87} and, although trends in pre-diabetes and in blood glucose levels in the population are less well documented, high rates of pre-diabetes have been reported.⁸⁷ Despite this, there is little reliable prospective evidence about the relevance to

CVD risk of blood glucose levels in the pre-diabetes or “normoglycaemic” ranges in China. Cross-sectional analyses in the CKB found evidence supporting a positive association of RPG levels below the diabetic threshold with prevalent CVDs (Chapter 4), but the evidence was limited by the potential for reverse causality and by investigation of self-reported, composite IHD and cerebrovascular disease outcomes. This chapter addresses these limitations, presenting relevant prospective data from the CKB. The objectives of this chapter were to examine a) the association of baseline and usual RPG levels with the risks of incident CVDs, including IHD and stroke subtypes; b) the relevance of fasting time to these associations; and c) whether other CVD risk factors, such as age, sex, adiposity and blood pressure modify these associations.

7.2 Methods

7.2.1 Exposure definitions

Plasma glucose levels were handled in the same way as described for the cross-sectional analyses (Sections 4.2.2 and 5.2.1). The small number (n=231) of baseline plasma glucose levels recorded as being above the maximum reading on the SureStep Plus glucometer (>27.8 mmol/L)¹⁸⁵ were categorised as ≥ 11.1 mmol/L in analyses examining plasma glucose as a categorical variable and were imputed as 27.8 mmol/L in analyses examining plasma glucose as a continuous variable.

7.2.2 Outcome definitions

The primary CVD outcomes investigated were as described in Chapter 6, including CVD mortality (ICD-10 I00-25, I27-88, I95-99), MI (I21-23), MCE (non-fatal MI or fatal IHD [I20-25]), IS (I63), ICH (I61), TS (I60, I61, I63, I64) and MOVD (IS, non-fatal MI or fatal IHD). Detailed descriptions of these outcomes are presented in Table E.1.

7.2.3 Selection of confounders

The process of selecting confounders for inclusion in statistical models examining the associations of plasma glucose with incident CVDs was as described previously (Sections 5.2.3 and 6.2.3). An initial list of potential confounders was identified based on a priori knowledge of the associations and mechanisms involved, and included age, sex, study area, education, smoking, alcohol, physical activity, consumption of various foodstuffs, measures of adiposity and blood pressure. The ultimate selection of confounders for inclusion in the main statistical models was based on demonstrated statistical associations of the potential confounding variables with RPG (Section 4.3.4) and with CVD outcomes, on findings of causal diagrams^{239, 240} (Figure D.1) and on examination of the effects of adjustment for the potential confounders in regression analyses examining the associations of plasma glucose with incident CVDs (Section 7.3.2).

7.2.4 Participant exclusions

In addition to the exclusion criteria applied for the cross-sectional analyses of the associations of plasma glucose with prevalent CVDs (Section 5.2.4), participants reporting a history of doctor-diagnosed IHD or stroke/TIA at baseline (n=20 044) were excluded. 467 508 participants (191 555 men, 275 953 women) remained after applying these exclusions (Figure 4.1).

7.2.5 Statistical analyses

Incidence and correlates of cardiovascular diseases

Direct standardisation, using the total study population as the standard, was used to calculate age-, sex-, study area- and education-standardised (as appropriate) CVD incidence rates across sexes, age groups and study areas. The associations of other potential confounding variables (education, smoking, alcohol consumption, physical activity, consumption of fresh fruit, dairy products, eggs, meat, wheat and tea, BMI, waist circumference, hip circumference, waist-to-hip ratio, body fat percent, SBP and DBP) with

incident MCE, IS, ICH, MOVD and with CVD mortality were examined using Cox proportional hazards models. All associations were examined stratifying by age-at-risk, sex and study area and adjusted for education. In order to investigate the independence of the associations, analyses were repeated additionally adjusting for alcohol consumption, smoking, SBP, physical activity, BMI, waist circumference, hip circumference and consumption of fresh fruit, dairy products, eggs, meat, wheat and tea. Intermediate levels of adjustment were also investigated for a subset of analyses as described in previous chapters: in addition to basic adjustment, associations with smoking were adjusted for alcohol consumption, and vice versa, associations with adiposity measures were adjusted for each other and associations with DBP were adjusted for SBP.

Associations of plasma glucose levels with incident cardiovascular diseases

Cox proportional hazards models were used to estimate HRs and associated CIs for the associations of baseline plasma glucose levels with the incident CVD outcomes described (Section 7.2.2). Age-at-risk <35 and >79 years was excluded. Participants were followed-up until 1st January 2014 and were censored at the date of their first event, loss to follow-up or death. Baseline RPG was categorised into seven groups (<4.3, 4.3-5.2, 5.3-5.7, 5.8-6.7, 6.8-7.7, 7.8-11.0, ≥ 11.1 mmol/L), selected to include the OGTT 2-hour post glucose-load thresholds for diabetes and IGT,⁷⁸ and to ensure reasonable numbers of participants and CVD events in all groups. The floating absolute risk method, which provides variance estimates across all exposure categories,²⁴¹ was used to calculate HRs, with the lowest category (<4.3 mmol/L) as the referent group. As described in Chapter 4, a single RPG measurement may not accurately reflect an individual's usual RPG level due to random measurement error and more systematic changes in glucose levels within individuals over time.¹⁷³ Usual RPG levels, that is RPG levels adjusted for regression dilution bias, were calculated as the mean resurvey RPG levels for individuals in each of the seven baseline RPG categories. Usual RPG levels in the lowest three categories (<4.3 mmol/L: 5.7 mmol/L; 4.3-5.2 mmol/L: 5.9 mmol/L; 5.3-5.7 mmol/L: 6.1 mmol/L) were similar; these categories

were therefore combined when investigating the association of usual RPG with CVD risks. Estimated HRs were examined for departure from log-linearity by testing the type 3 chi-square for regression dilution bias-adjusted RPG as a categorical variable in a model also containing regression dilution bias-adjusted RPG as a continuous variable. If the shape was log-linear, then baseline RPG was also investigated as a continuous variable. Examination of HRs separately for the first four and subsequent years of follow-up showed no strong evidence of departure from the proportional hazards assumption. Log HR estimates for baseline RPG examined as a continuous variable were multiplied by the reciprocal of the regression dilution ratio to obtain regression dilution bias-corrected estimates¹⁷⁴ and presented as HR per 1 mmol/L higher usual RPG.

The Cox proportional hazards models were stratified by age-at-risk, sex (where appropriate) and study area. Potential confounding factors were then progressively added into the models. The main and subsidiary analyses investigating the associations of baseline plasma glucose with risks of CVDs were stratified by age-at-risk, sex (where appropriate) and study area, and were adjusted simultaneously (except where the variable of interest in subsidiary analyses to investigate effect modification) for education (no formal education, primary school, middle school, high school, college/university), smoking (never, occasional, ex-regular, current regular), alcohol (never, occasional, ex-regular, reduced intake, weekly intake), SBP (<100, 100-109, 110-119, 120-129, 130-139, 140-149, 150-159, 160-169, ≥ 170 mmHg) and physical activity (<10, 10-19.9, 20-29.9, 30-39.9, ≥ 40 MET hours/day). Further analyses additionally adjusted effect estimates for waist-to-hip ratio (<0.75, 0.75 to <0.80, 0.80 to <0.85, 0.85 to <0.90, 0.90 to <0.95, ≥ 0.95). Adjusting for SBP, physical activity and waist-to-hip ratio as continuous, rather than categorical variables, did not alter the calculated risk estimates.

Adjusted HRs were compared across strata of other CVD risk factors (age, sex, education, smoking, alcohol consumption, physical activity, SBP and adiposity) using stratum-specific regression dilution ratios to adjust for intra-individual variation in RPG, across fasting period

strata and across the 10 study areas. Chi-square tests for trend and heterogeneity were applied to the log HRs and their SEs and global tests of heterogeneity were calculated by summing separate heterogeneity chi-square statistics for subgroup analyses, and their degrees of freedom, to account for multiple testing and to examine overall consistency of HRs across population subgroups.²⁴²

Separate analyses investigating the association of RPG with CVD risks excluded individuals with a plasma glucose level suggestive of diabetes (ie, ≥ 7.0 mmol/L with time since last food ≥ 8 hours, or ≥ 11.1 mmol/L with time since last food < 8 hours, or FPG level ≥ 7.0 mmol/L on subsequent testing),¹⁸³ individuals from Zhejiang (where participants were asked to fast prior to blood sampling) or individuals diagnosed with diabetes during the follow-up period (based on data from the CDC disease reporting system, the HI system or death certificates). In additional analyses, the strength of the association of baseline RPG levels with incident MOVD, MCE, IS and ICH were compared for events identified through the HI system only, the CDC disease reporting system only and through both data sources (defined as the same CVD event type recorded in both data sources within a seven day period). The strength of association of baseline RPG levels with IS and ICH were also calculated separately for events confirmed by CT/MR scan and events not confirmed by CT/MR scan; these analyses were restricted to events ascertained from the CDC disease reporting system, since data on CT/MR scans are not currently available from the national HI system.

Macros developed within CTSU were used to undertake the Cox proportional hazards analyses and to calculate directly standardised incidence and mortality rates. SAS version 9.3 (SAS Institute Inc., Cary, North Carolina, USA) was used to conduct all statistical analyses. Line graphs and forest plots were produced using R version 2.13.1 (R Foundation for Statistical Computing, Vienna, Austria); plots presenting the results of the Cox proportional hazards analyses were produced using an R graphics programme developed within CTSU.

7.3 Results

Among the 467 508 CKB participants who had no prior history of diabetes or CVD at baseline, the mean (SD) age was 51 (11) years, and 59% were women. The mean (SD) baseline RPG was 5.9 (1.9) mmol/L, slightly higher in women than in men (6.0 vs 5.8 mmol/L) after adjustment for age and study area. 1.7% (n=8144) of participants had a baseline RPG level ≥ 11.1 mmol/L, ie, suggestive of diabetes.

7.3.1 Incidence and correlates of cardiovascular diseases

During ~3.3 million person years of follow-up, there were 6645 CVD deaths, with age-, study area- and education-standardised CVD mortality rates of 347.3 per 100 000 person years in men and 124.8 per 100 000 person years in women. 22 023 MOVD events (men: 905.3 per 100 000 person years; women: 558.6 per 100 000 person years), 3270 MCE (men: 163.3 per 100 000 person years; women: 62.2 per 100 000 person years), 19 153 IS (men: 755.7 per 100 000 person years; women: 502.3 per 100 000 person years) and 4326 ICH (men: 203.5 per 100 000 person years; women: 93.8 per 100 000 person years) were recorded. Although the absolute rates of these major CVDs differed modestly between this study population and that included in analyses of the association of diabetes with incident CVDs (Section 6.3.1), the trends across age groups, sexes and study areas were comparable (Figures F.1 to F.6) and are therefore not described again here. Similarly, the correlates of incident CVDs were comparable between the two populations, and are therefore not presented again for the population included in these blood glucose analyses.

7.3.2 Plasma glucose and risk of cardiovascular disease

Table 7.1 shows adjusted HRs for CVDs associated with baseline RPG levels. For major ischaemic CVDs and CVD mortality the risks increased progressively with higher levels of baseline RPG across the seven categories examined, with no evidence of a threshold below which lower RPG levels were no longer associated with lower risks. For CVD mortality the adjusted HRs were 1.00, 1.01, 1.05, 1.10, 1.16, 1.29 and 2.03 at baseline RPG

levels of <4.3, 4.3-5.2, 5.3-5.7, 5.8-6.7, 6.8-7.7, 7.8-11.0 and ≥ 11.1 mmol/L (p for trend <0.001) (Table 7.1 and Figure 7.1). Similar positive associations with baseline RPG were observed for MCE, IS and MOVD (p for trend <0.001 for all). When compared with the reference category (<4.3 mmol/L), the risk of MOVD and IS was significantly elevated in all baseline RPG categories ≥ 5.3 mmol/L, but the risk of CVD mortality and MCE was significantly elevated only in baseline RPG categories ≥ 7.8 mmol/L. For ICH, the strength of the association with baseline RPG was more modest (p for trend=0.10) (Table 7.1 and Figure 7.2). The shape of the associations of baseline RPG with these major CVDs were similar in men and women (Figures F.7 to F.9).

Figure 7.3 shows the relationship between usual RPG levels (ie, RPG adjusted for regression dilution bias) and the risk of CVD mortality and ischaemic CVDs. There was a positive, log-linear relationship between usual RPG and these CVDs ie, a similar proportional increase in risk for a given increment in RPG, at least at usual RPG levels above 5.9 mmol/L. For CVD mortality, each 1 mmol/L increment in usual RPG was associated with an adjusted HR of 1.11 (95% CI 1.10-1.13), with no clear difference in the strength of the association across age groups (Figure 7.4). However, the association was somewhat stronger in men than in women (p for heterogeneity=0.005), in urban than in rural areas (p for heterogeneity=0.03) and in individuals with higher levels of education (p for heterogeneity=0.009) or lower baseline SBP (p for trend=0.002) (Figure 7.5).

A positive, log-linear association was also found between usual RPG levels and the risk of ischaemic CVDs, with no evidence of a threshold level in the usual RPG range examined (Figure 7.3). For MOVD, 1 mmol/L higher usual RPG was associated with an 8% (HR 1.08, 95% CI 1.07-1.09) greater risk. The association of RPG with MCE was modestly stronger, with an adjusted HR of 1.10 (1.08-1.13) per 1 mmol/L higher usual RPG, somewhat more extreme for fatal, than for non-fatal, IHD events (1.13 [1.09-1.16] vs. 1.05 [1.01-1.10]) (Table 7.2). For IS, each 1 mmol/L higher usual RPG was associated with an adjusted HR of 1.08 (1.07-1.09) (Figure 7.3), and the association was again stronger for fatal events than for

non-fatal events (1.15 [1.10-1.21] vs 1.08 [1.07-1.09]) (Table 7.2). There was a stronger association between RPG and MOVD (p for trend=0.003) (Figure 7.4) and IS (p for trend <0.001) (Figure 7.6) at younger ages, in both men and women, and a modestly stronger association with MOVD in men than in women (p for heterogeneity=0.04). For MCE, no such differences were seen between sexes (p =0.2) or across age groups (p for trend=0.9) (Figure 7.6). For IS, the association was somewhat stronger in rural, than in urban, areas (p for heterogeneity=0.02) (Figure 7.7), but otherwise there was no clear evidence of substantial effect modification of the associations of RPG with major ischaemic CVDs (Figures 7.8 and 7.9).

After adjusting for regression dilution bias, there was an apparent, positive log-linear association between usual RPG and the risk of ICH; each 1 mmol/L higher usual RPG was associated with an adjusted HR of 1.05 (95% CI 1.02-1.07) (Figure 7.10), driven by effects on fatal (1.10, 1.07-1.13), rather than non-fatal (0.98, 0.95-1.02) ICH (Table 7.2). The association of usual RPG with ICH risk appeared to be stronger in men (p for heterogeneity=0.009) (Figure 7.11), in current regular smokers (p for heterogeneity<0.001) and in individuals with lower BMI (p for trend=0.03) (Figure 7.12).

To examine the shape of the association at the upper end of the RPG distribution in more detail, the 7.8-11.0 mmol/L category was further sub-divided into three groups (7.8-8.8 mmol/L, 8.9-9.9 mmol/L, 10.0-11.0 mmol/L). Although this resulted in relatively small numbers of participants and CVD events in these categories, and consequently wide CIs around the risk estimates, the positive, log-linear associations remained (Figures F.10 and F.11).

The shape of the associations described above did not appear to differ across fasting periods, although the relatively small number of participants reporting a fasting period of ≥ 6 hours produced wide CIs (Figures 7.13, 7.14 and 7.15). When plasma glucose levels were dichotomised into non-fasting (fasting period <8 hours) and fasting (fasting period ≥ 8 hours)

there was no clear difference in the strength of association per 1 SD increase in non-fasting (2.0 mmol/L) and fasting (1.1 mmol/L) plasma glucose with CVD mortality (p for heterogeneity=0.1), MOVD (p for heterogeneity=0.3), MCE (p for heterogeneity=0.9) or IS (p for heterogeneity=0.2). For ICH, however, there was an apparently stronger association with non-fasting, than with fasting, plasma glucose levels (1.06 vs 0.92 per 1 SD increment, p for heterogeneity=0.004).

Additional adjustment for waist-to-hip ratio (Table 7.2) and for other potential confounding factors (eg, dietary factors and DBP) did not materially alter the strength of the associations between RPG and CVD risks (Figures 7.16-7.18). There was some heterogeneity in the strength of the association of RPG with MOVD (p<0.001) and IS (p<0.001) across study areas (Figures 7.19 and 7.20). In particular, the association of RPG with both MOVD (1.02, 95% CI 1.01-1.04) and IS (1.02, 1.00-1.03) appeared to be modestly weaker in Harbin and modestly stronger in Zhejiang (1.08 [1.04-1.13] and 1.10 [1.06-1.14], respectively). There was no notable variation in the strength of the association of RPG with CVD mortality (p for heterogeneity=0.5), MCE (p for heterogeneity=0.8) or ICH (p for heterogeneity=0.5) across study areas (Figures 7.19-7.21). When the strength of association was compared for events ascertained through different data sources there was modest variability for certain disease outcomes. For IS, the associations were slightly stronger for cases identified through both CDC disease reporting and HI data sources (1.06, CI 1.05-1.08) than through the CDC system only (1.04, 1.02-1.06) or through the HI system only (1.04, 1.03-1.05) (p for heterogeneity=0.006). Similar findings were seen for MOVD, with adjusted HRs of 1.06 (1.05-1.08), 1.04 (1.02-1.05) and 1.04 (1.03-1.05), respectively (p for heterogeneity=0.004). Conversely, the association of baseline RPG with ICH, was slightly weaker for events ascertained through both data sources (0.99, 0.94-1.05) than for events reported only in the CDC disease reporting system (1.01, 0.97-1.05) or the HI system (1.04, 1.03-1.05) (p for heterogeneity=0.03). There was no evidence of such heterogeneity for MCE (p for heterogeneity=0.9). When restricted to stroke events ascertained through the CDC disease

reporting system, there was similarly no difference in the strength of association between RPG and IS (p for heterogeneity=0.6) or ICH (p for heterogeneity=0.9) events with and without MR/CT scan confirmation.

As noted above, the association of RPG with IS and MOVD was stronger in Zhejiang where participants were asked to fast prior to blood sampling. After excluding participants from this area ($n=51\ 656$), the associations of usual RPG with incident CVDs were not altered (Table 7.3). The associations of RPG with CVDs also persisted after excluding participants with new onset of diabetes during follow-up ($n=12\ 048$) (Table 7.4) or those with screen-detected diabetes at baseline ($n=13\ 050$) (Table 7.5).

7.4 Discussion

These analyses represent the largest prospective investigation of the association of plasma glucose levels with risks of major CVDs in individuals without known diabetes in China, and one of the largest worldwide. They showed positive, log-linear associations between usual RPG levels and the risk of CVD mortality and major ischaemic CVDs—both IHD and IS—continuing down to a usual RPG level of 5.9 mmol/L, with no evidence of a threshold. It was estimated that each 1 mmol/L higher usual RPG was associated with 8-11% increased risk of these CVDs, which varied little across population subgroups. For ICH there was a significant, though more modest, association with RPG.

7.4.1 Plasma glucose and risk of cardiovascular disease

Several prospective studies of mostly Western populations have previously investigated the association of blood glucose levels with risks of CVDs among individuals without diabetes. Most of these studies have been based on FBG, and have relatively consistently shown a positive association with CVD risks within the pre-diabetes range (ie, 6.1-6.9 mmol/L).^{6, 102, 103, 306} At blood glucose levels below this range, however, there is conflicting evidence. In the APCSC of ~240 000 participants from 13 Asian and Australasian cohorts, there was a positive, log-linear association between usual FBG and incident IHD ($n=816$) and CVD

mortality (n=1661), which continued down to at least 4.9 mmol/L.¹⁰³ In contrast, two other large studies demonstrated J-shaped associations of fasting glucose levels with CVD risks.^{6, 102} In the KCPS of ~1.2 million Korean men and women without known diabetes, there was a J-shaped association of baseline FPG levels with both IHD (n=58 148) and IS (n=46 596), with the lowest risk at ~5.0 mmol/L, below which disease risk appeared to increase with decreasing FPG.¹⁰² In the ERFC, which included individual participant data on ~260 000 participants from 54 studies with ~11 000 IHD and ~1500 IS events, there was a J-shaped association of usual FBG with IHD risk, with the lowest risk at FPG levels of 3.9-5.6 mmol/L, but no significant association with IS risk (likely due to inadequate statistical power).⁶ In the KCPS, the J-shaped association persisted after exclusion of the first five years of follow-up suggesting little influence of reverse causality, but regression dilution bias was not adjusted for.¹⁰² Relatively small numbers of events among individuals with the lowest fasting glucose levels, particularly in the ERFC study,⁶ may have produced imprecise risk estimates at the lower end of the blood glucose distribution. Although studies have suggested hypoglycaemia may be associated with increased CVD risk, for example, through effects on catecholamine levels, coagulation, inflammation, endothelial function and the risk of cardiac arrhythmias, these relate to more extreme hypoglycaemia complicating treatment with hypoglycaemic agents.³⁰⁷

To-date, no prospective studies have reported on the association of blood glucose levels with risk of CVDs in mainland China. Two small prospective studies of Chinese in Taiwan, each with only 100-200 CVD events, have reported conflicting findings on the association of blood glucose levels with composite CVD outcomes.^{249, 308} In the Chin Shan Community Cardiovascular Cohort of ~2000 men and women aged ≥ 35 years without prior CVD from a single township, each 1 SD higher FPG was associated with a 25% increase in the risk of CVD (n=170) after full adjustment, with a comparable 29% increased risk per 1 SD increment in post-load plasma glucose.³⁰⁸ A study of ~16 500 adults without diabetes recruited from a single hospital, found no association when comparing the risk of CVD

mortality (n=95) in individuals with a baseline fasting (HR 0.84, 95% CI 0.47-1.51) or post-load (1.61, 0.82-2.99) plasma glucose in the highest, compared with the lowest, quintile.²⁴⁹ The null findings of this study possibly reflect the small number of CVD deaths included, short (average 3.5 years) follow-up period or relatively small differences in plasma glucose levels between quintiles (FPG: 4.7 mmol/L vs 5.5-7.0 mmol/L; post-load plasma glucose ≥ 7.4 mmol/L vs < 5.2 mmol/L). Furthermore, selection of participants from a hospital setting may limit generalisability to the general population.

Nearly all previous studies have investigated FBG, with very limited evidence available regarding the association of RBG levels with risk of CVD among individuals without diabetes. In a recent meta-analysis, based on published summary data from seven studies including ~50 000 individuals, there was no convincing evidence of an association between RBG levels and the risk of cardiac (n=314) (HR 1.02, 95% CI 0.98-1.07 per 1 mmol/L higher RBG), stroke (n=544) (1.11, 0.95-1.31) or cardiovascular (n=1782) (1.11, 1.00-1.24) mortality, and only weak evidence of a positive association with total CVD (n=2087) (1.21, 1.01-1.25).¹⁰⁴ The CKB includes much greater numbers of well-characterised CVD events than this earlier meta-analysis. In addition, the presented CKB analyses have taken into account potential biases due to reverse causality and regression dilution, so providing the most robust evidence to-date of a positive association of RPG with CVD risks among adults without prior diabetes diagnoses.

As in the CKB, previous studies have tended to find comparable strengths of association of blood glucose with different CVD subtypes.^{103, 104} In the APCSC the associations with CVD mortality, IHD and stroke were similar, with 19% (95% CI 15-22%), 23% (19-27%) and 21% (18-24%) increased risks, respectively, per 1 mmol/L higher usual FBG.¹⁰³ Due to relatively low rates of ICH in most Western populations, and limited availability of scanning technology in some, particularly earlier, studies, previous evidence on the association of blood glucose levels with HS has been limited, and largely restricted to the KCPS, which included ~19 000 HS (including ICH and SAH).¹⁰² In keeping with findings from the CKB, it

showed a more modest association of FPG with HS than with other CVD outcomes, with clearly elevated risks only in the highest FPG categories when compared with a FPG of 4.7-5.5 mmol/L (7.0-7.7 mmol/L: HR 1.15 in men and 1.12 in women; ≥ 7.8 mmol/L: 1.24 in men and 1.14 in women). In the CKB there was a stronger association for fatal, than for non-fatal, ICH, and similar patterns were seen, but to a lesser extent, for IHD and IS. This may reflect a survival effect or more severe disease in fatal cases, although this trend has not been reported previously.^{102, 103}

In the CKB, there was a suggestion of a stronger association of RPG with CVD risks in men than in women, which has not been described in previous large studies. For MOVD and IS, the RPG-associated risks appeared stronger in younger age groups, which was also seen for the association of FPG with TS in a large tabular data meta-analysis including over 370 000 individuals from 116 cohorts.¹⁴² The meta-analysis also found a similar trend with age in the association of FPG with IHD. This was not seen in the CKB, nor was there any clear evidence of effect modification by age in the associations of RPG with the risk of CVD mortality or ICH. The lack of effect modification likely reflects the lower number of CVD deaths and MCE and ICH events, such that there was inadequate power to detect such an effect in the CKB. Consistent with earlier studies,^{102, 103} there was no other clear, consistent evidence of effect modification of the association of RPG with CVD risks by other CVD risk factors in the CKB. There was, however, some evidence of variation in the strength of association of RPG with IS and MOVD across the 10 study areas. The modestly weaker association seen in Harbin is consistent with the modestly weaker association of self-reported diabetes with IS in this study area, although, as described (Section 6.4.2), the reasons for these lower risks are unclear. There was a modestly stronger association of RPG with MOVD and IS in Zhejiang, which may reflect the fact that participants in this study area were asked to fast prior to attending the study assessment centre. The relevance of the small differences in the strength of association according to data source and to scan-

status for stroke events are also unclear. However, given the modest nature of these differences and the number of tests undertaken, these could reflect the play of chance.

Repeat exposure measurements were collected at five year intervals in the CKB on a randomly selected 5% sample of the study population (Section 3.6). These data enable adjustment for regression dilution bias and, therefore, estimation of disease risks associated with usual, or longer-term average, RPG levels. As described (Section 4.3.5), there was considerable intra-individual variation in RPG levels in the CKB, and failure to adjust for this would have substantially underestimated RPG-associated CVD risks. As expected,¹⁷⁴ the lowest baseline RPG categories included a disproportionate number of participants with baseline RPG levels lower than their usual RPG level. The usual RPG levels for the lowest baseline RPG categories were thus similar, and were combined when plotting CVD risks associated with usual RPG levels. As a result of this, although for usual RPG the association was observed down to 5.9 mmol/L, for baseline RPG the log-linear association continued down to at least 4.3 mmol/L. These findings from the CKB thus add to the body of evidence suggesting the existence of a continuous, positive association between blood glucose levels and CVD risks extending throughout the glycaemic range, with no evidence of a threshold effect.

The increased risks of CVD at higher blood glucose levels in this and previous studies could merely reflect the effects of undiagnosed or future diabetes⁷⁸ on CVD risks.⁶ After exclusion of participants with plasma glucose levels suggestive of diabetes, or who developed diabetes during follow-up in the CKB, the association of blood glucose levels with CVD persisted, lending further strong support to a continuous association of RPG with CVD risks throughout the glycaemic range. This approach has previously been adopted only in small studies. In the British Regional Heart Study of ~7500 middle-aged men without diabetes followed-up for 17 years, TS risk was elevated only among men with the highest RPG levels (≥ 8.2 mmol/L), with an adjusted HR of 1.86 (95% CI 1.11-3.13).³⁰⁹ When 320 men diagnosed with diabetes during follow-up were excluded, the elevated risk of TS in this

group became non-significant (1.56, 0.83-2.91), although this possibly reflects the relatively small sample size.

Residual confounding, reflecting more adverse CVD risk profiles in individuals with higher RPG levels and also the lack of adjustment for lipid levels, could contribute to some of the demonstrated association of RPG with CVD. Additional evidence, however suggests a causal association. Higher blood glucose levels may contribute to CVD risk through damage to the vascular endothelium,^{92, 93} pro-inflammatory and pro-thrombotic effects of blood glucose,⁹²⁻⁹⁵ oxidative stress, glycosylation of vessel wall proteins and lipids,²⁴⁸ or through damage to vessels supplying the arterial walls.⁹⁸ Furthermore, there is reasonably good evidence from Mendelian randomisation studies that a causal association exists between higher blood glucose levels and CVD throughout the full glycaemic range.^{148, 149} However, randomised trials of glucose-lowering agents in individuals with pre-diabetes have, so far, been inconclusive in their effects on cardiovascular outcomes.³¹⁰⁻³¹³ The STOP-NIDDM trial in 1368 individuals with IGT showed a 49% (HR 0.51, 95% CI 0.28-0.95) relative risk reduction for major CVD events (n=47) in those taking acarbose, compared with placebo, over a 3-year period.³¹⁰ This contrasts, however, with the NAVIGATOR Study, which showed no benefit (0.94, 0.82 to 1.09) of nateglinide over placebo for CVD risk in ~9300 individuals with IGT at high risk of CVD.³¹³ The ORIGIN Trial of insulin glargine versus standard care in ~12 500 individuals with IFG, IGT or type 2 diabetes and with high CVD risk also showed no effect (1.02, 0.94-1.11) on the risk of major CVDs (n=2054) after 6-years of follow-up.³¹² Finally, the DREAM Trial of rosiglitazone versus placebo in over 5000 individuals with IFG, IGT or type 2 diabetes, produced a null finding (1.37, 0.97-1.94) (n=130).³¹¹ Inadequate follow-up or statistical power, and investigation of study populations with differing CVD risks may go some way to explaining the inconsistent findings, but the impact of glucose lowering in individuals with pre-diabetes remains unclear.

7.4.2 Limitations

Fasting and post-load blood glucose are arguably more robust glycaemic measures than RBG, which may be subject to greater inter- and intra-individual variation, particularly as a result of post-prandial increases. Adjusting for regression dilution bias in the present analyses would have accounted for some of this variability, and greater variability in RPG measurements, compared with other glycaemic measures, would likely, if anything, attenuate the association with CVD risks. Use of additional data collected on fasting time demonstrated that only a small proportion of variation in plasma glucose levels in the CKB was explained by fasting time (Section 4.3.3), and no consistent difference was found in the associations of plasma glucose with CVD mortality and major ischaemic CVDs across strata of fasting time. Furthermore, investigation of fasting-time adjusted plasma glucose did not appreciably alter the nature of the associations of plasma glucose with CVD (Appendix Note F.1). Moreover, it has been suggested that non-fasting glucose may be a more relevant exposure for CVD risk than fasting glucose, as we spend most of our time in a non-fasting state.¹⁴⁸ Thus, despite the recognised limitations,²²⁷ in large-scale population-based epidemiological studies such as this RPG would appear to be a reliable and practical glycaemic indicator.²²⁶

Whilst adjustment of RPG levels for regression dilution bias is a strength of the presented analyses, the approach adopted has limitations. Unlike the parametric methods^{186, 187} (which were not used in the presented analyses due to violation of their underlying assumptions, Section 4.3.5), there is no generally accepted approach for calculating CIs for regression dilution ratios using MacMahon's method.¹⁷⁴ As such, the regression dilution ratios were assumed to be calculated without error, and any error in their calculation could not be reflected in usual RPG-associated CVD risk estimates. However, given the large sample size included in the resurvey, the effect of this was likely minimal,¹⁷⁵ as was any effect of not accounting for fasting times in calculation of regression dilution ratios (Section 4.4.4). Clearly, confounding variables, in addition to the main exposure variable, may be

susceptible to intra-individual variation and should ideally be adjusted for regression dilution bias.³¹⁴ Correction for within-person variation in multi-variable models is possible using a parametric approach,¹⁹⁰ but is again less feasible with the non-parametric approach, and is further complicated when the variables concerned are correlated,³¹⁵ as herein. It was not, therefore, feasible to adjust confounding variables for within-person variation in the presented analyses, which may have contributed to residual confounding.³¹⁴ Failure to adjust for unmeasured known (eg, lipids), and unknown, confounding factors will likely also have contributed to some degree of residual confounding.

As described (Section 6.4.4), the “healthy volunteer effect”³⁰⁵ may have influenced the findings of these analyses. Furthermore, issues with regards to the validity of CVD outcomes, described in Chapter 6, are also clearly relevant to analyses of the associations of RPG. Misclassification of CVD events would likely be non-differential and unrelated to RPG levels, attenuating the strength of the associations of RPG with CVDs. Given the prospective nature of these analyses, they are clearly less susceptible to reverse causality than the cross-sectional analyses presented in Chapter 5, particularly since individuals with a history of IHD or stroke/TIA at baseline were excluded (although reliance on self-report for medical history may have resulted in misclassification). However, there likely remains some bias due to reverse causality; blood glucose levels are known to increase after stroke and IHD events^{316, 317} (although this effect is unlikely to be chronic), and some CVD medications may impact negatively on glucose tolerance.³¹⁸ Furthermore, due to the relatively short follow-up period, it was not feasible to routinely exclude the first few years of follow-up, although the comparability of risk estimates based on the first and second halves of the follow-up period suggest the impact of this on the risk of reverse causality was minimal.

7.4.3 Conclusion

The presented analyses provide clear evidence of an independent, positive log-linear relationship of RPG levels with risk of major CVDs in adults without diabetes in China. Although current CVD risk prediction models consider blood glucose only in a dichotomous form—as the presence or absence of diabetes^{319, 320}—these findings suggest consideration of plasma glucose levels as a continuous risk factor may be informative. Furthermore, the findings highlight the potential benefit of population-level approaches aimed at shifting the entire blood glucose distribution, particularly in light of the reported global upward trend in fasting glucose levels.⁷⁹

Table 7.1. Adjusted hazard ratios for incident cardiovascular diseases by baseline random plasma glucose

Baseline random plasma glucose (mmol/L) (Mean)	Cardiovascular disease mortality			Major occlusive vascular disease			Major coronary event			Ischaemic stroke			Intracerebral haemorrhage		
	No. of events	HR	95% CI	No. of events	HR	95% CI	No. of events	HR	95% CI	No. of events	HR	95% CI	No. of events	HR	95% CI
<4.3 (4.0)	355	1.00	0.90-1.11	996	1.00	0.94-1.07	180	1.00	0.86-1.16	837	1.00	0.93-1.07	272	1.00	0.89-1.13
4.3-5.2 (4.8)	1727	1.01	0.96-1.06	5952	1.07	1.04-1.10	883	1.00	0.94-1.07	5184	1.08	1.05-1.11	1242	1.01	0.95-1.07
5.3-5.7 (5.5)	1132	1.05	0.99-1.12	4113	1.11	1.08-1.14	530	0.95	0.87-1.03	3637	1.13	1.10-1.17	772	1.04	0.97-1.11
5.8-6.7 (6.2)	1643	1.10	1.05-1.16	5456	1.14	1.11-1.17	806	1.06	0.99-1.14	4750	1.15	1.12-1.18	1073	1.06	1.00-1.13
6.8-7.7 (7.2)	887	1.16	1.08-1.24	2893	1.24	1.19-1.28	437	1.14	1.04-1.25	2509	1.25	1.20-1.30	507	1.02	0.93-1.11
7.8-11.0 (8.8)	568	1.29	1.19-1.41	1718	1.30	1.24-1.37	288	1.30	1.16-1.46	1462	1.30	1.23-1.37	320	1.10	0.98-1.22
≥11.1 (15.8)	333	2.03	1.82-2.26	895	1.77	1.66-1.89	146	1.74	1.47-2.04	774	1.79	1.66-1.92	140	1.32	1.11-1.56
χ^2_1 (<i>p</i> value) for trend	152.1 (<0.001)			255.9 (<0.001)			57.2 (<0.001)			213.8 (<0.001)			10.8 (0.10)		
Per 1 mmol/L higher		1.06	1.05-1.07		1.04	1.04-1.05		1.06	1.04-1.07		1.04	1.04-1.05		1.03	1.01-1.04

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, physical activity and systolic blood pressure.
CI, confidence interval; HR, hazard ratio.

Figure 7.1. Adjusted hazard ratios for cardiovascular diseases by baseline random plasma glucose

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HR) are plotted against mean baseline random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs). Numbers above the CIs are the HRs and numbers below the CIs are of cardiovascular disease events.

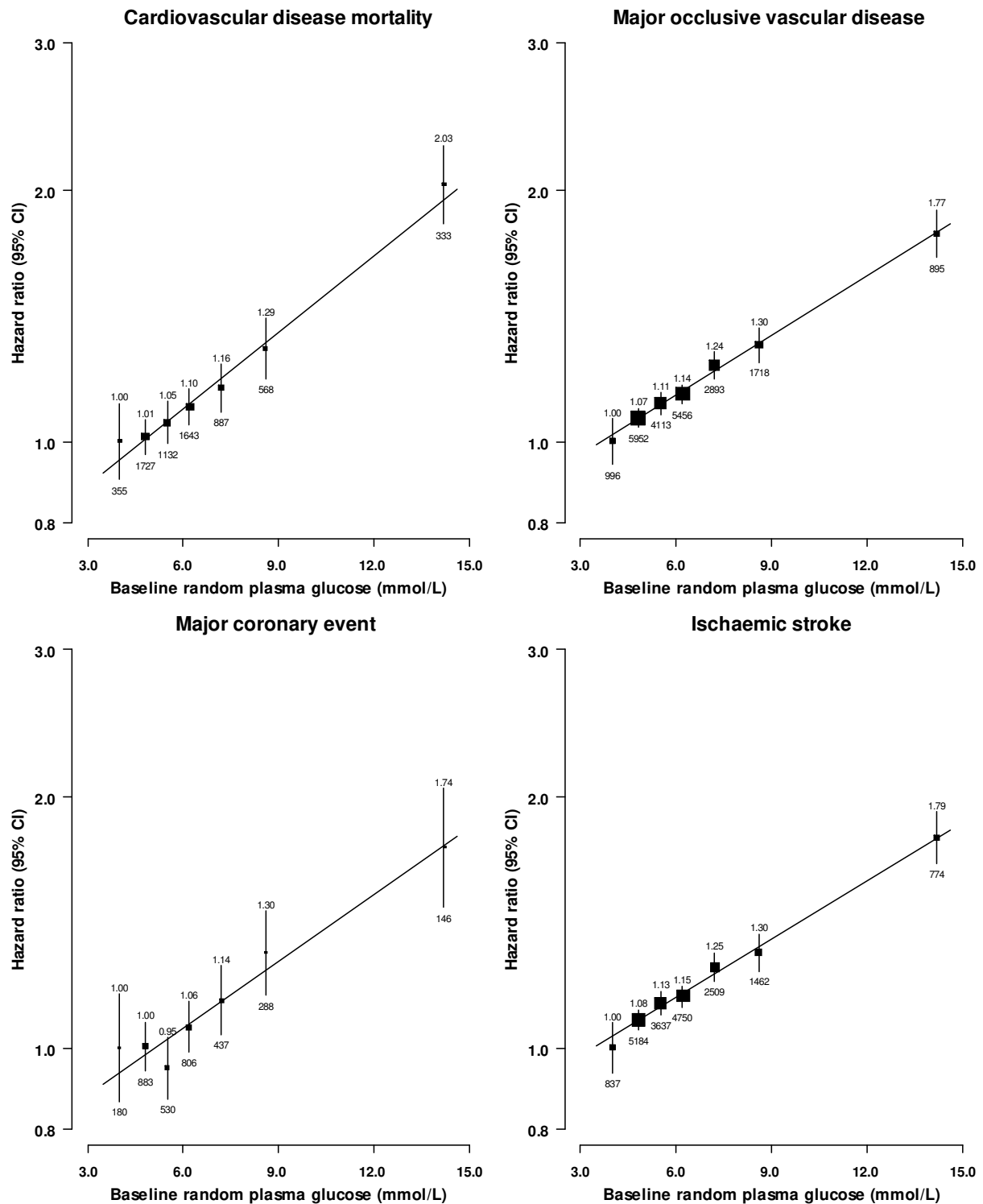


Figure 7.2. Adjusted hazard ratios for intracerebral haemorrhage by baseline random plasma glucose

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HR) are plotted against mean baseline random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs). Numbers above the CIs are the HRs and numbers below the CIs are of intracerebral haemorrhage events.

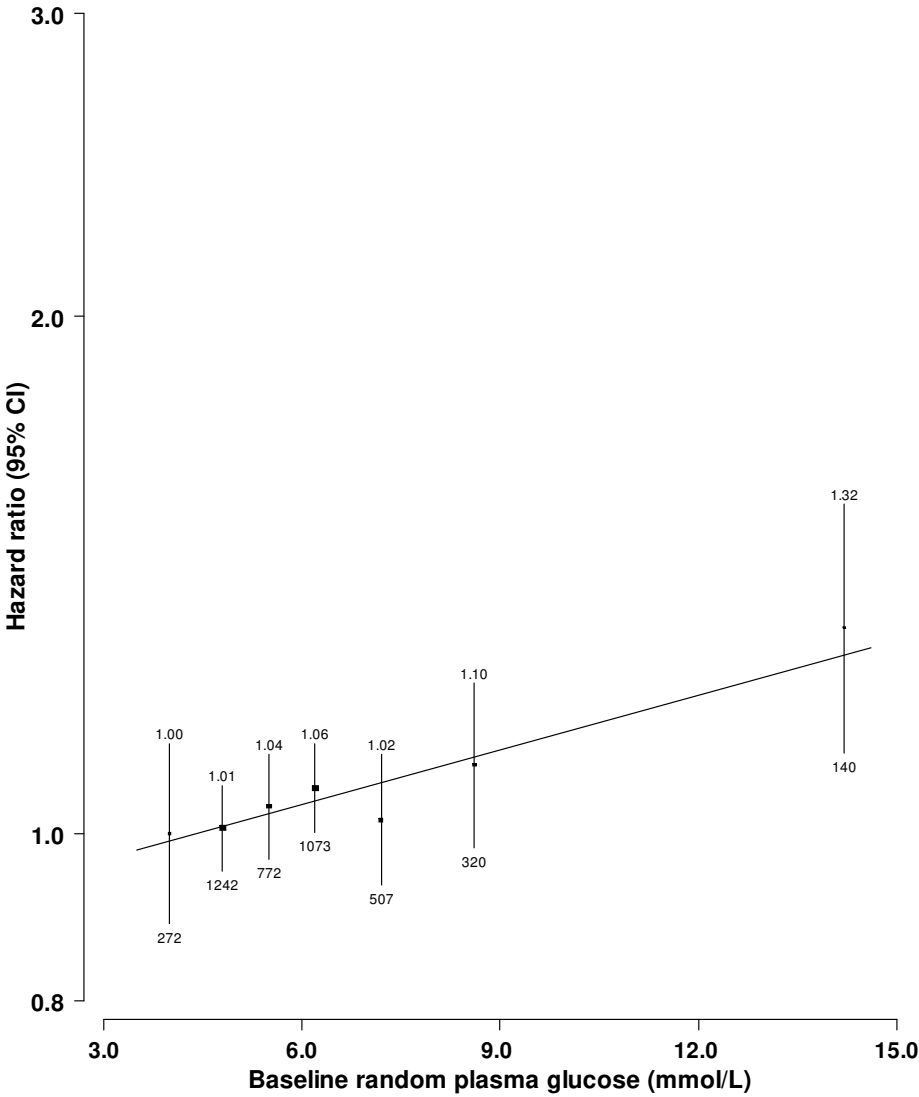


Figure 7.3. Adjusted hazard ratios for cardiovascular diseases by usual random plasma glucose

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HRs) are plotted against mean usual random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs). Numbers above the CIs are the HRs and numbers below the CIs are of cardiovascular disease events.

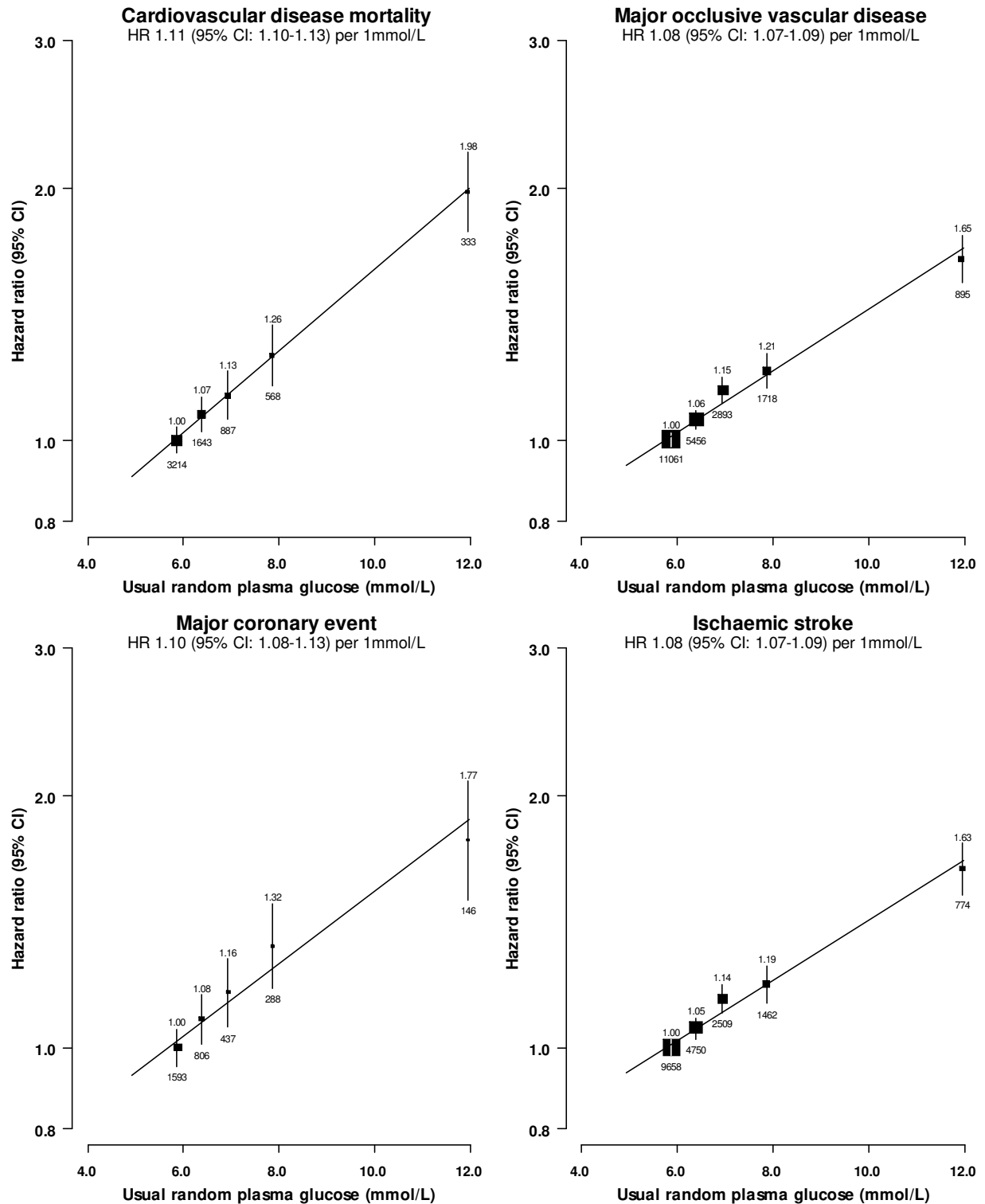


Figure 7.4. Age- and sex-specific adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease per 1 mmol/L higher usual random plasma glucose

Stratified by age, sex (where appropriate) and study area and adjusted for education, smoking, alcohol consumption, physical activity, systolic blood pressure. Squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs). The open diamond represents the overall HR and its 95% CI.

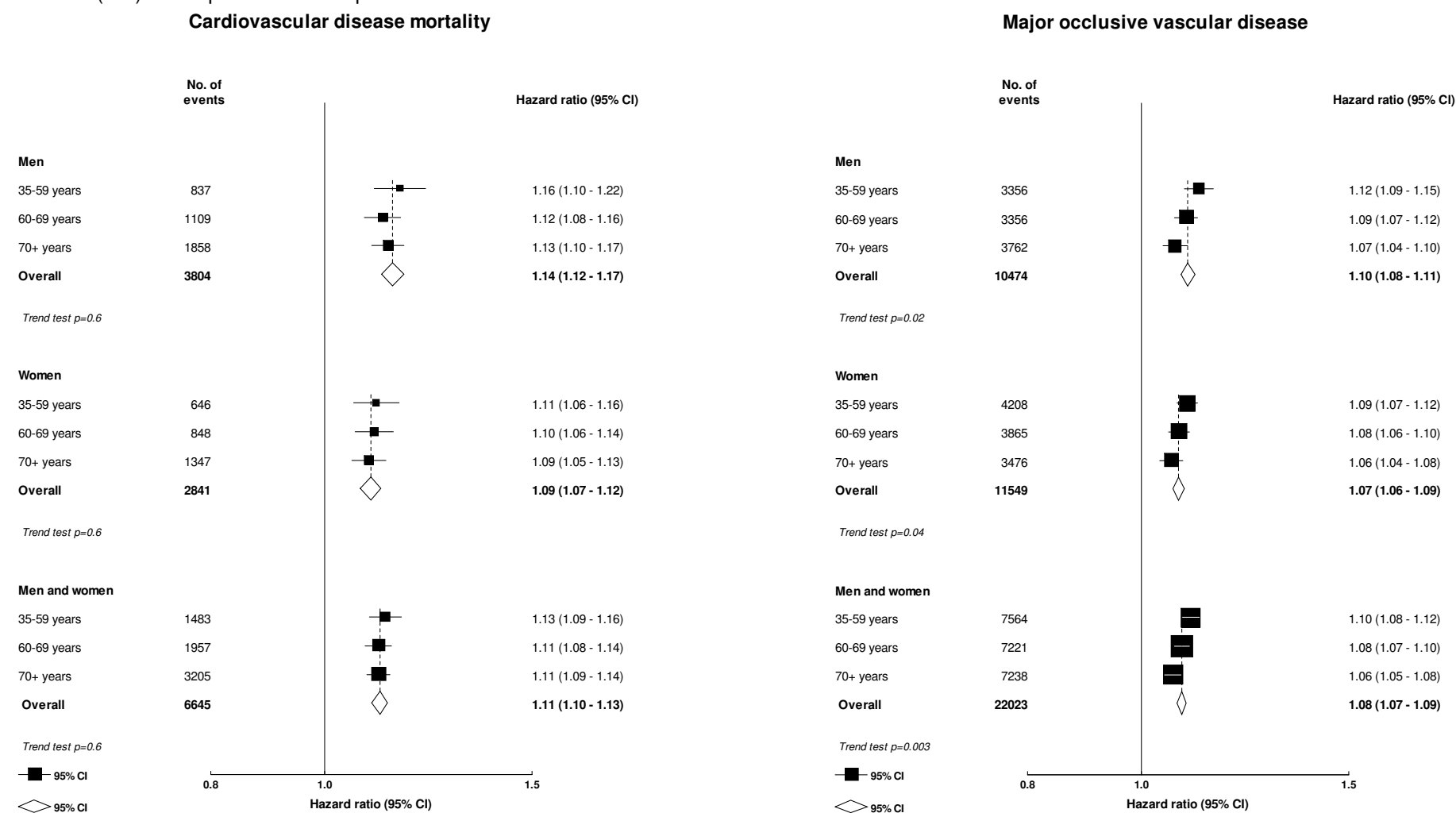


Figure 7.5. Adjusted hazard ratios for cardiovascular disease mortality per 1 mmol/L higher usual random plasma glucose in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

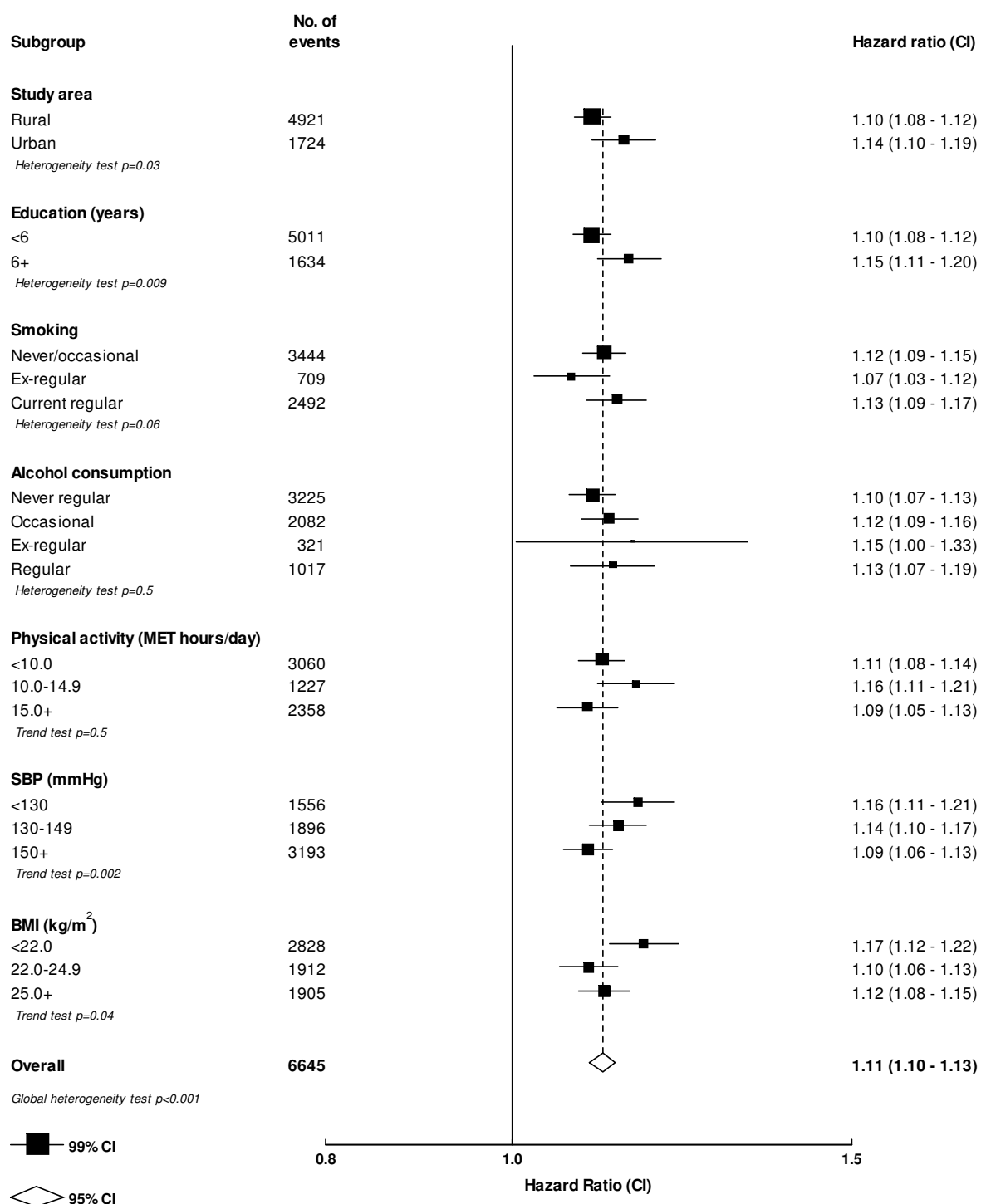


Table 7.2. Adjusted hazard ratios for incident cardiovascular diseases per 1 mmol/L higher usual random plasma glucose

Disease outcome	Events	Model A ^a		Model B ^b		Model C ^c	
		HR	95%CI	HR	95%CI	HR	95%CI
Cardiovascular disease mortality	6645	1.14	1.13-1.16	1.11	1.10-1.13	1.11	1.10-1.13
Major occlusive vascular disease	22023	1.11	1.10-1.12	1.08	1.07-1.09	1.07	1.06-1.08
Ischaemic heart disease							
Fatal myocardial infarction	1294	1.15	1.11-1.18	1.13	1.09-1.16	1.12	1.08-1.15
Non-fatal myocardial infarction	1147	1.07	1.03-1.12	1.05	1.01-1.10	1.04	1.00-1.09
Major coronary event	3270	1.12	1.10-1.15	1.10	1.08-1.13	1.10	1.07-1.12
Stroke							
<i>Ischaemic stroke</i>							
Fatal	446	1.18	1.13-1.24	1.15	1.10-1.21	1.15	1.10-1.21
Non-fatal	18707	1.10	1.09-1.11	1.08	1.07-1.09	1.07	1.06-1.08
Any	19153	1.10	1.09-1.11	1.08	1.07-1.09	1.07	1.06-1.08
<i>Intracerebral haemorrhage</i>							
Fatal	1947	1.14	1.11-1.17	1.10	1.07-1.13	1.11	1.08-1.14
Non-fatal	2379	1.02	0.99-1.06	0.98	0.95-1.02	0.98	0.94-1.02
Any	4326	1.08	1.06-1.11	1.05	1.02-1.07	1.05	1.02-1.07
<i>Total stroke</i>							
Fatal	2675	1.14	1.12-1.17	1.11	1.08-1.14	1.12	1.09-1.14
Non-fatal	22201	1.09	1.08-1.10	1.07	1.06-1.08	1.06	1.05-1.07
Any	24876	1.10	1.09-1.11	1.07	1.06-1.08	1.06	1.05-1.07

^aModel A, stratified by age, sex and study area. ^bModel B, additionally adjusted for education, smoking, alcohol, physical activity and systolic blood pressure. ^cModel C, additionally adjusted for waist-to-hip ratio.

CI, confidence interval; HR, hazard ratio.

Figure 7.6. Age- and sex-specific adjusted hazard ratios for major coronary event and ischaemic stroke per 1 mmol/L higher usual random plasma glucose

Stratified by age, sex (where appropriate) and study area and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs). The open diamond represents the overall HR and its 95% CI.

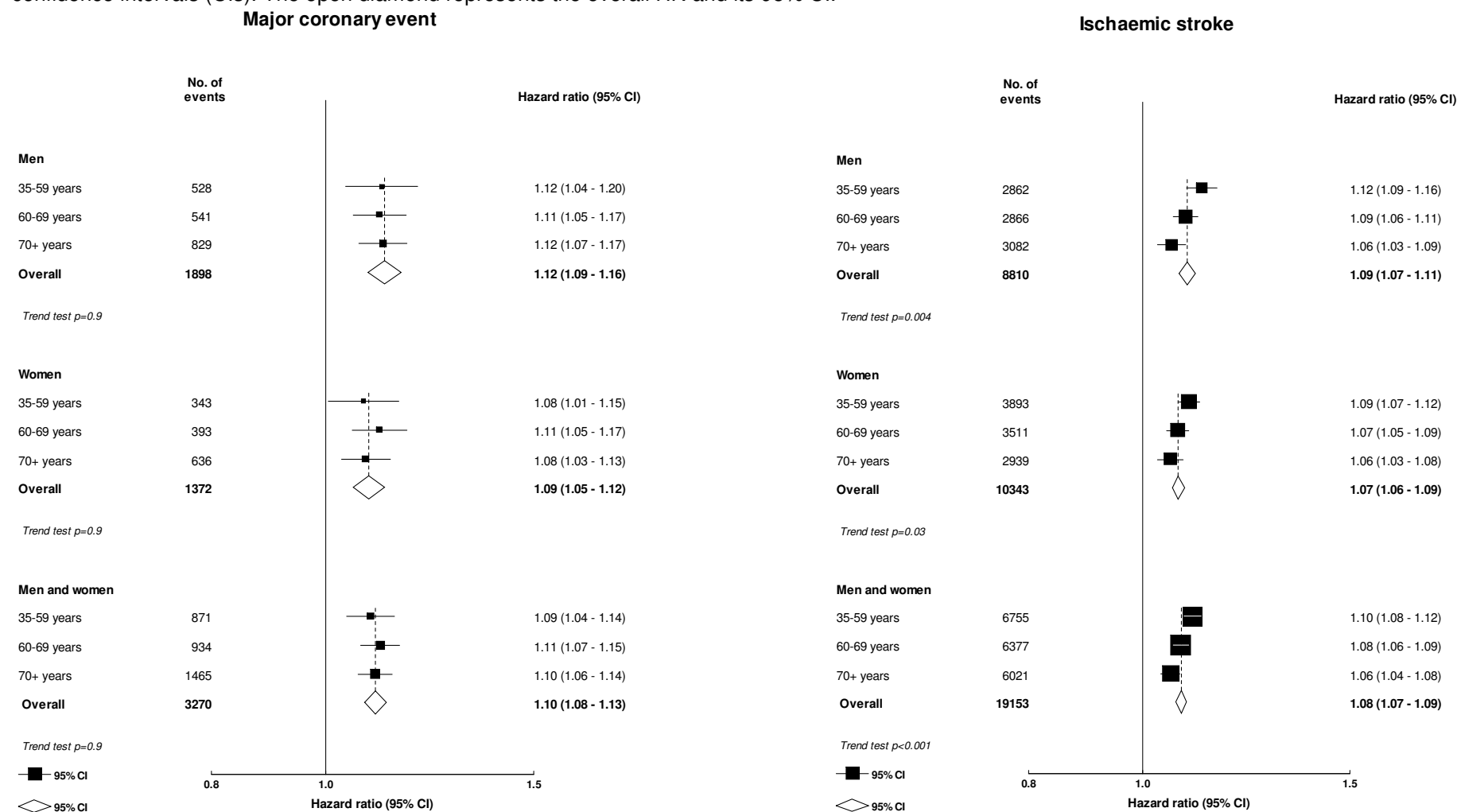


Figure 7.7. Adjusted hazard ratios for ischaemic stroke per 1 mmol/L higher usual random plasma glucose in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

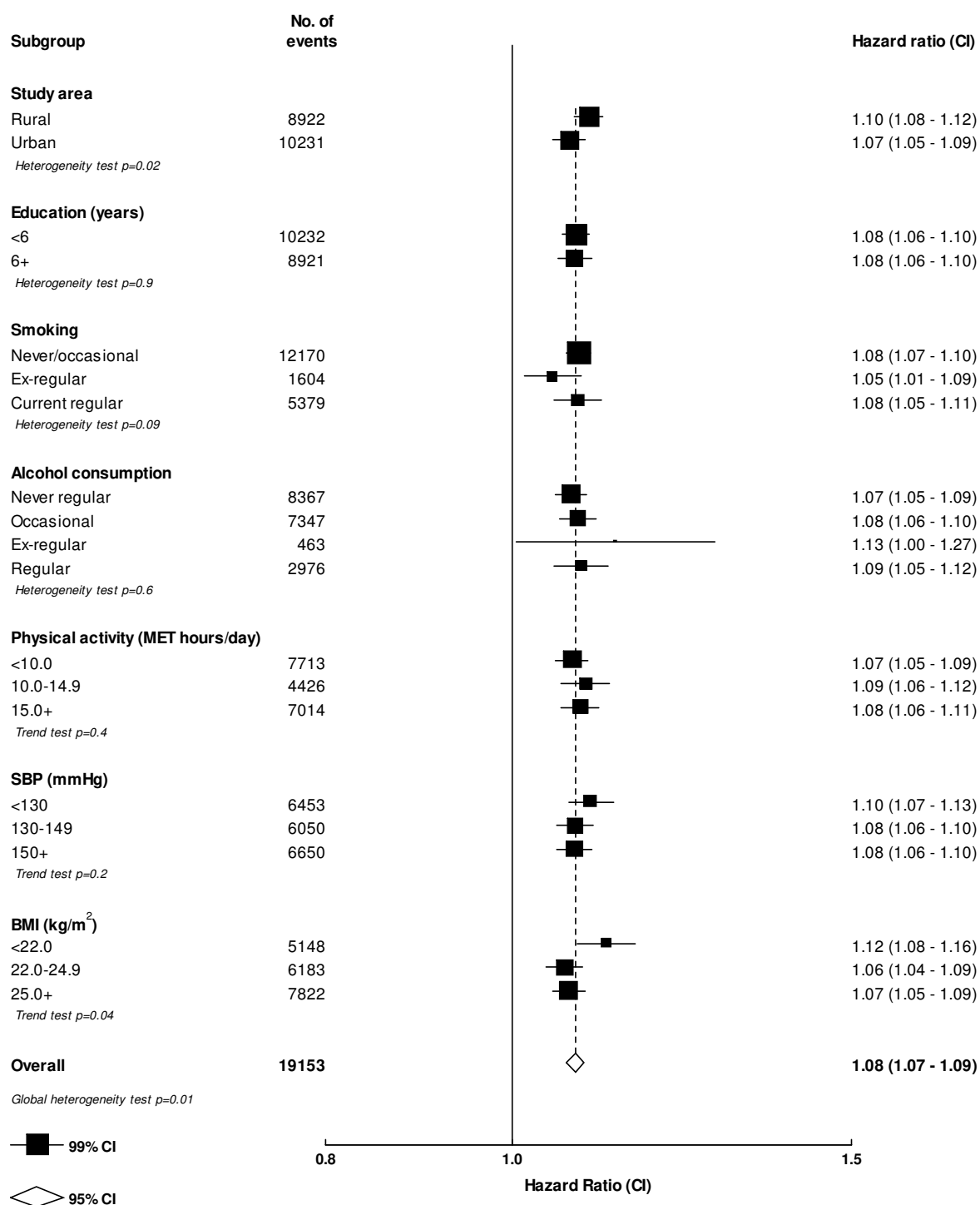


Figure 7.8. Adjusted hazard ratios for major occlusive vascular disease per 1 mmol/L higher usual random plasma glucose in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

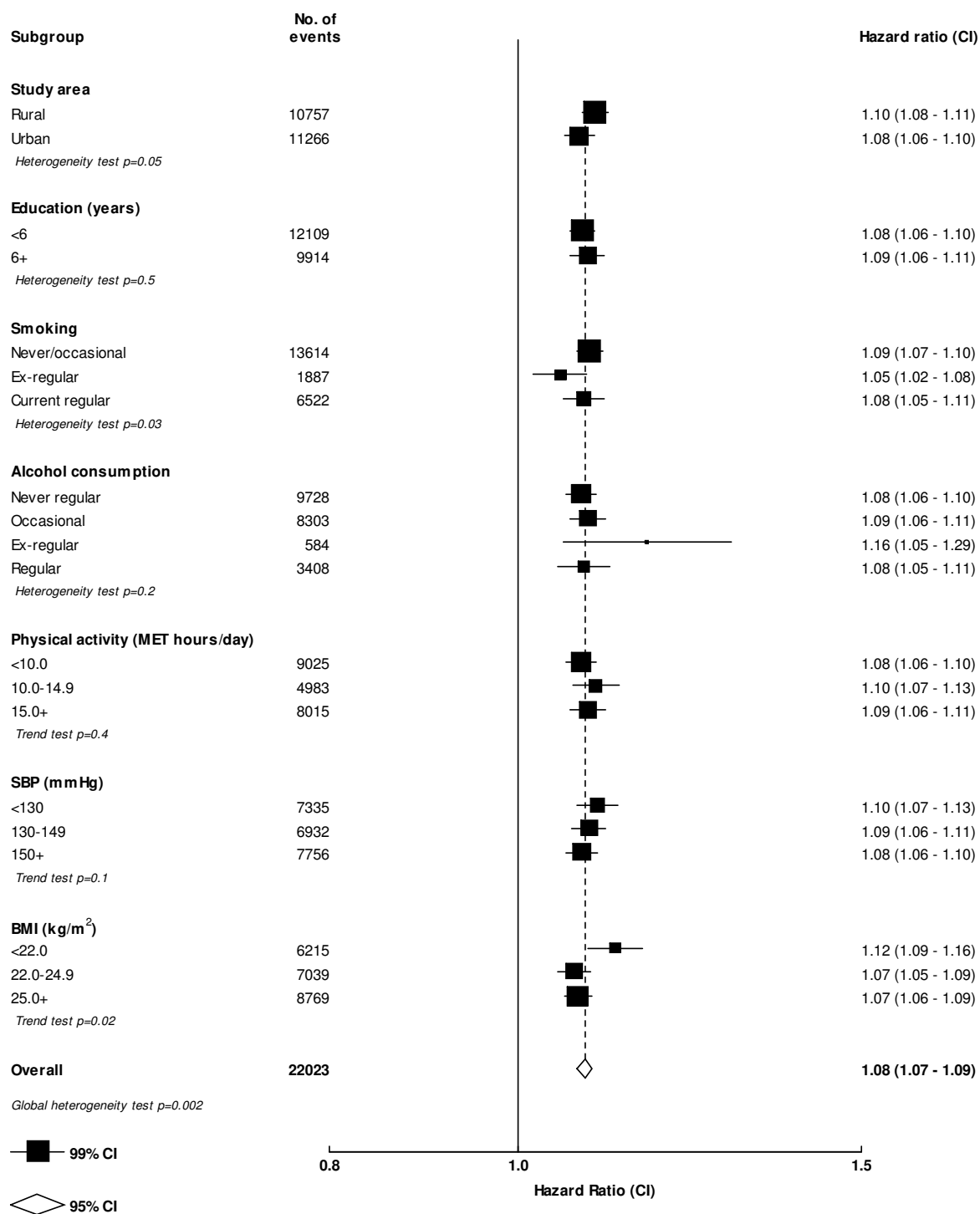


Figure 7.9. Adjusted hazard ratios for major coronary event per 1 mmol/L higher usual random plasma glucose in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

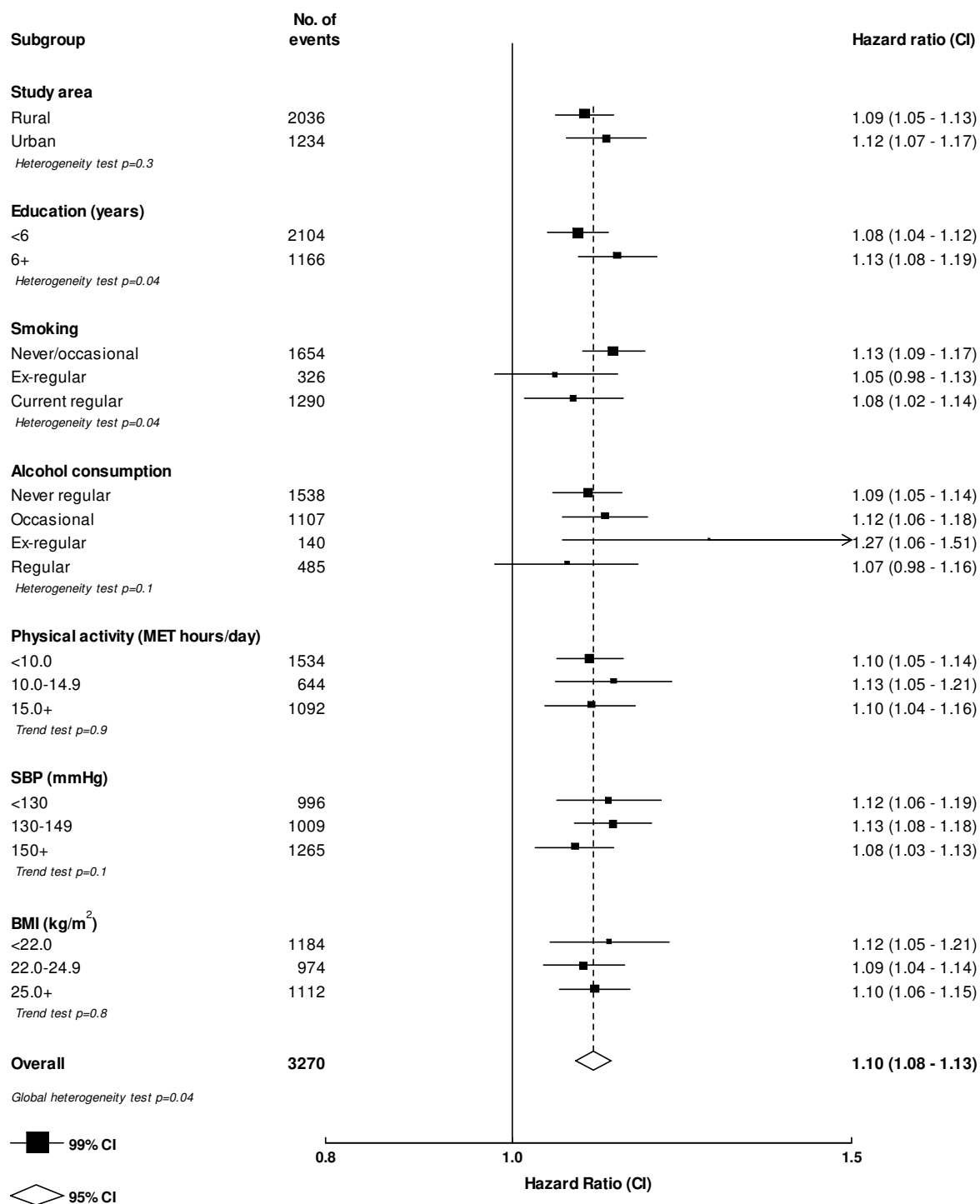


Figure 7.10. Adjusted hazard ratios for intracerebral haemorrhage by usual random plasma glucose

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HRs) are plotted against mean usual random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs). Numbers above the CIs are the HRs and numbers below the CIs are of intracerebral haemorrhage events.

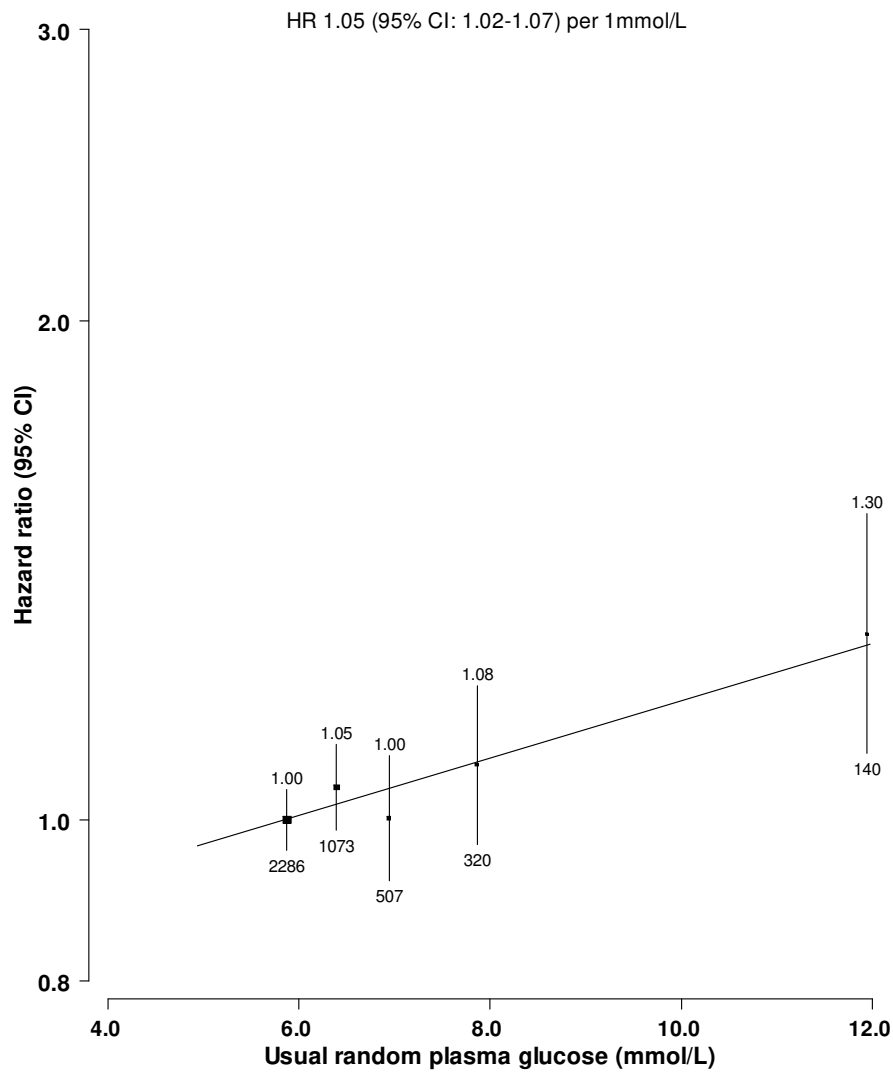


Figure 7.11. Age- and sex-specific adjusted hazard ratios for intracerebral haemorrhage per 1 mmol/L higher usual random plasma glucose

Stratified by age, sex (where appropriate) and study area and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs). The open diamond represents the overall HR and its 95% CI.

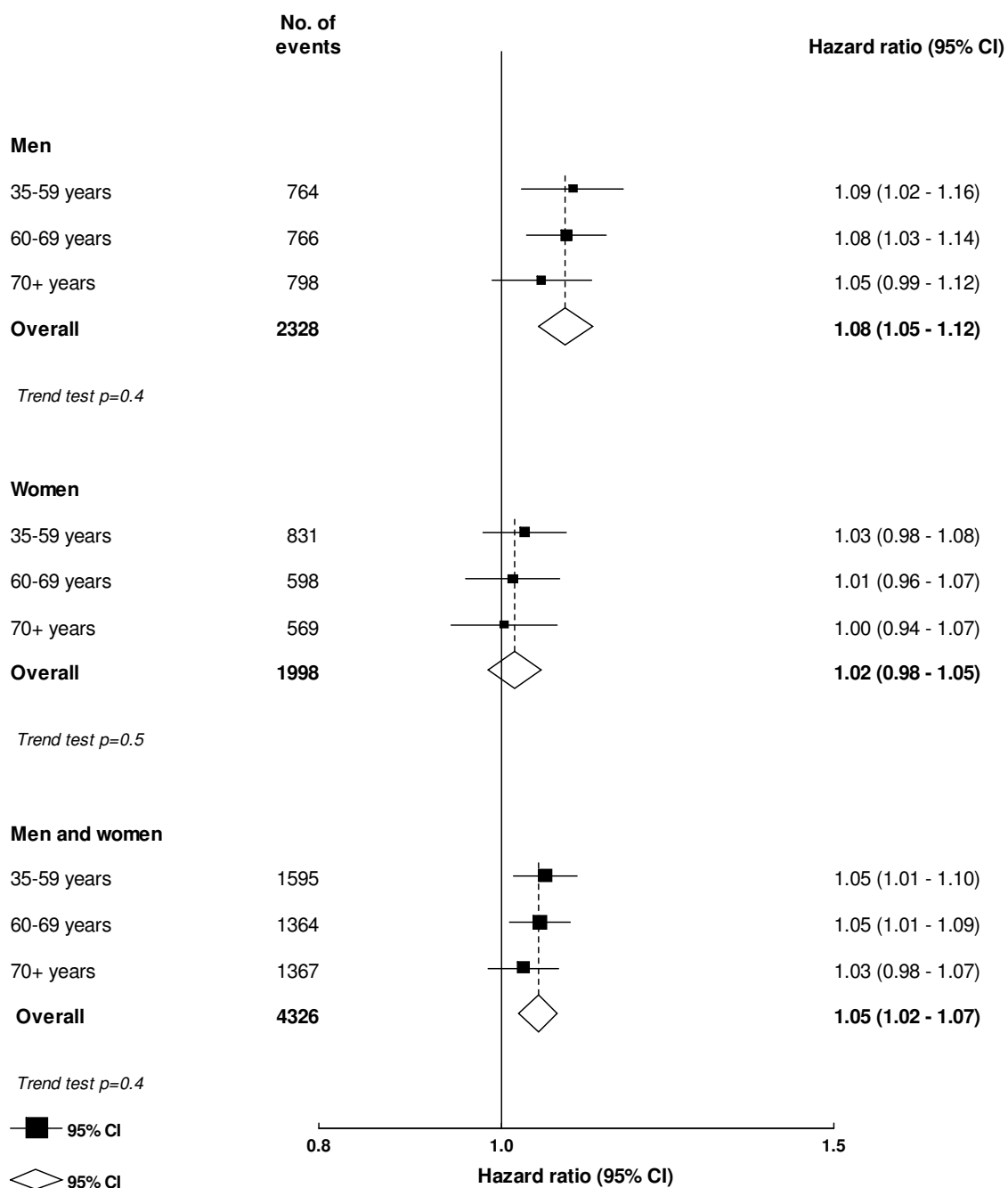


Figure 7.12. Adjusted hazard ratios for intracerebral haemorrhage per 1 mmol/L higher usual random plasma glucose in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

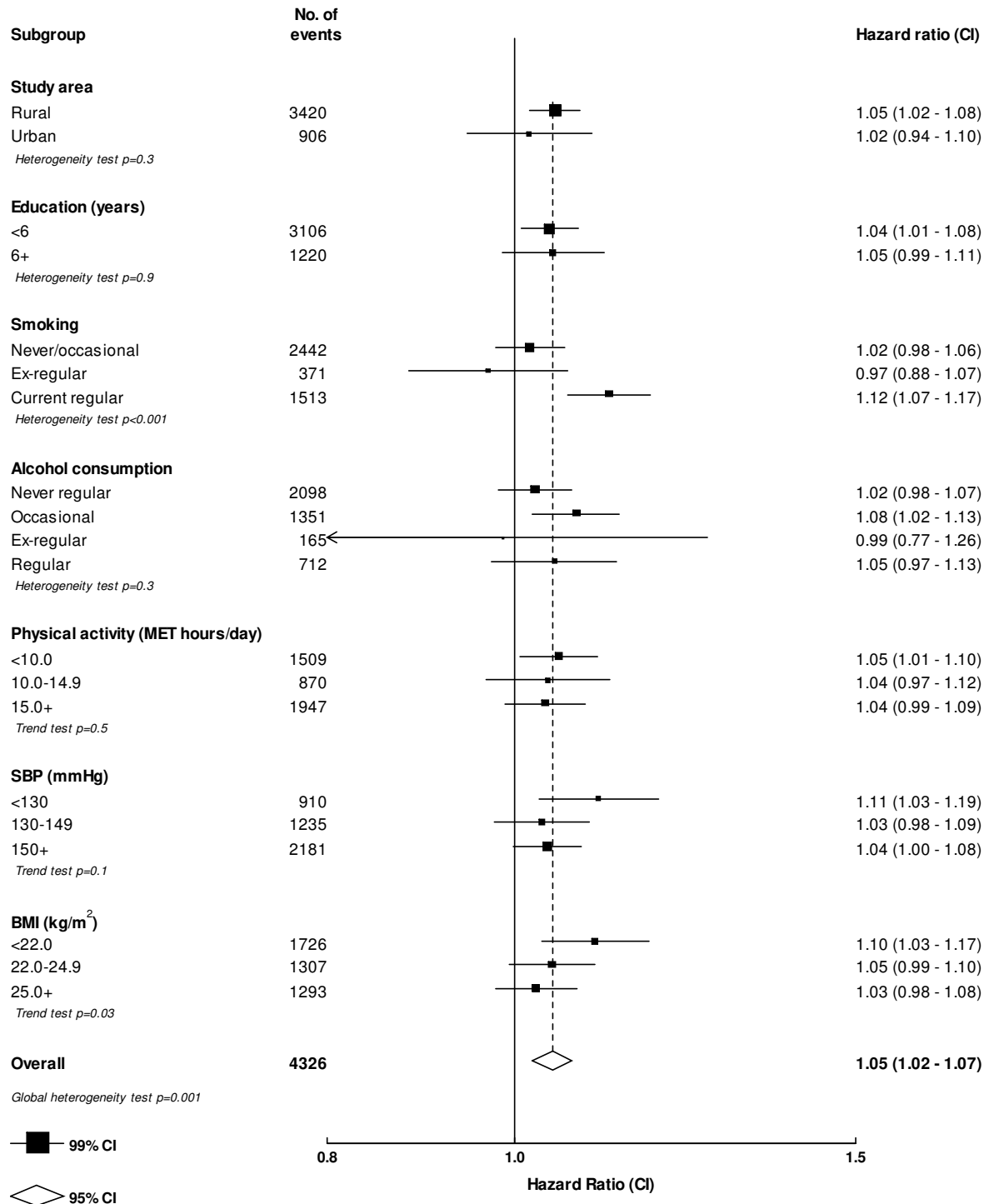


Figure 7.13. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease by baseline random plasma glucose, by fasting time

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HR) are plotted against mean baseline random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

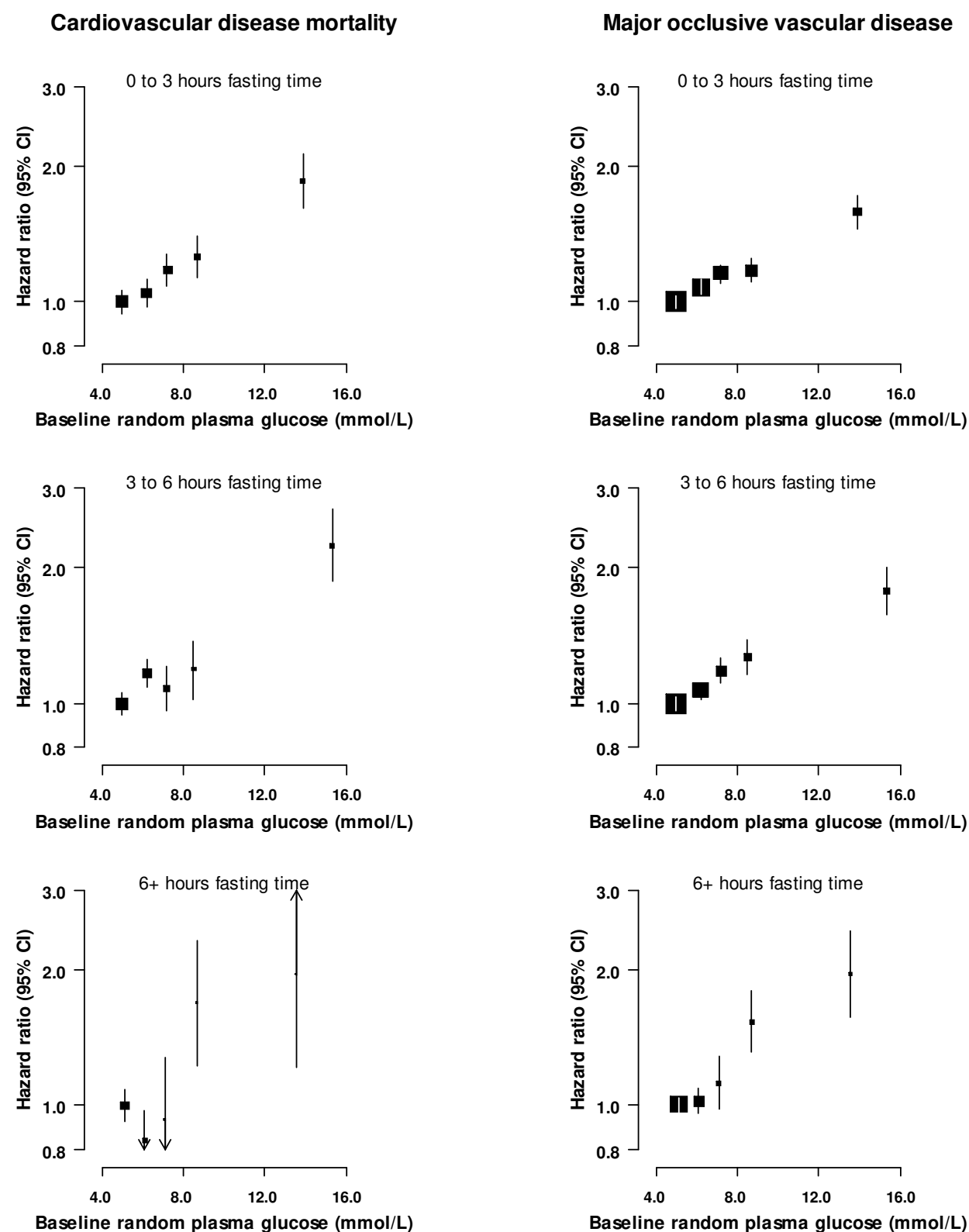


Figure 7.14. Adjusted hazard ratios for major coronary event and ischaemic stroke by baseline random plasma glucose, by fasting time

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HR) are plotted against mean baseline random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

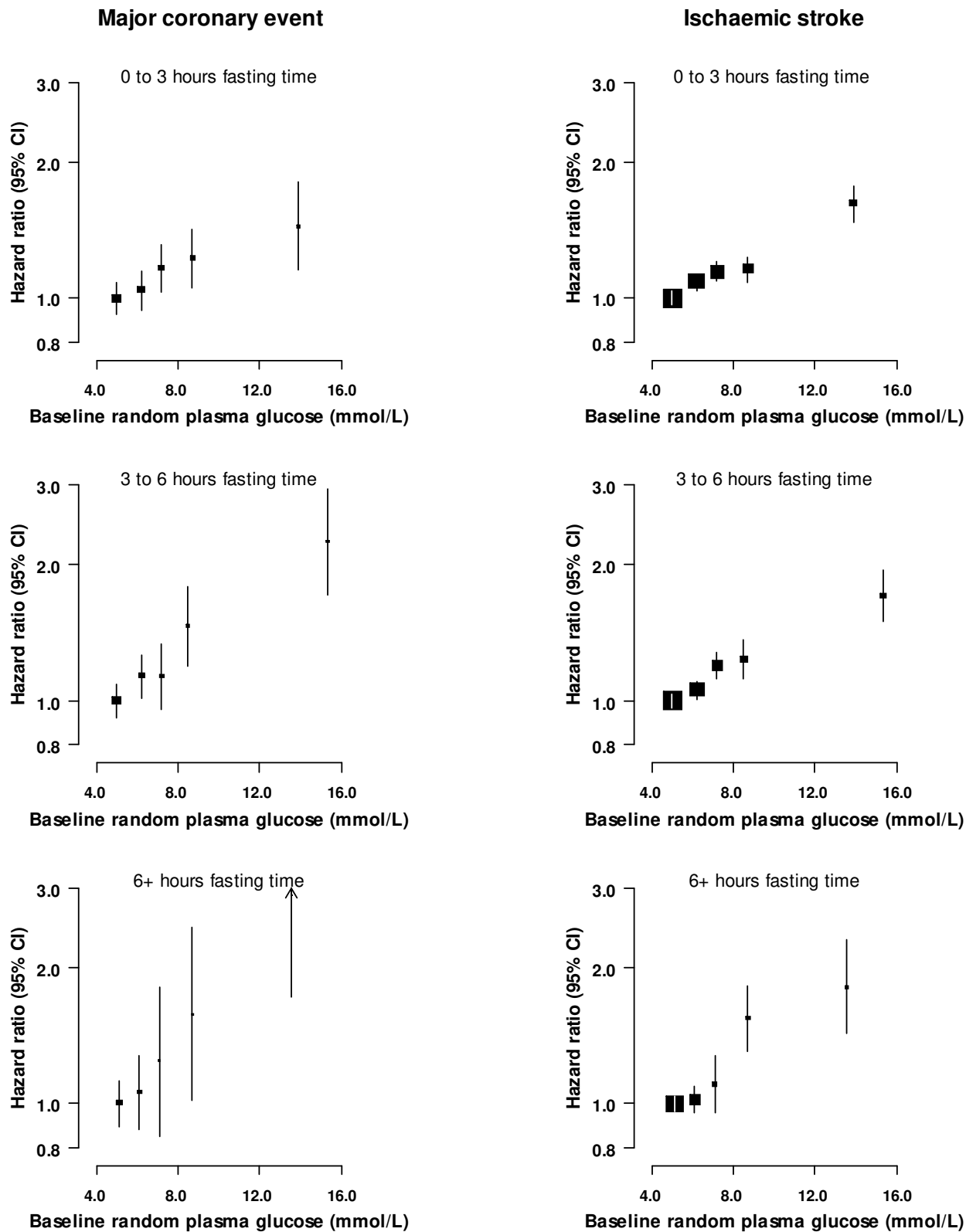


Figure 7.15. Adjusted hazard ratios for intracerebral haemorrhage by baseline random plasma glucose, by fasting time

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HR) are plotted against mean baseline random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

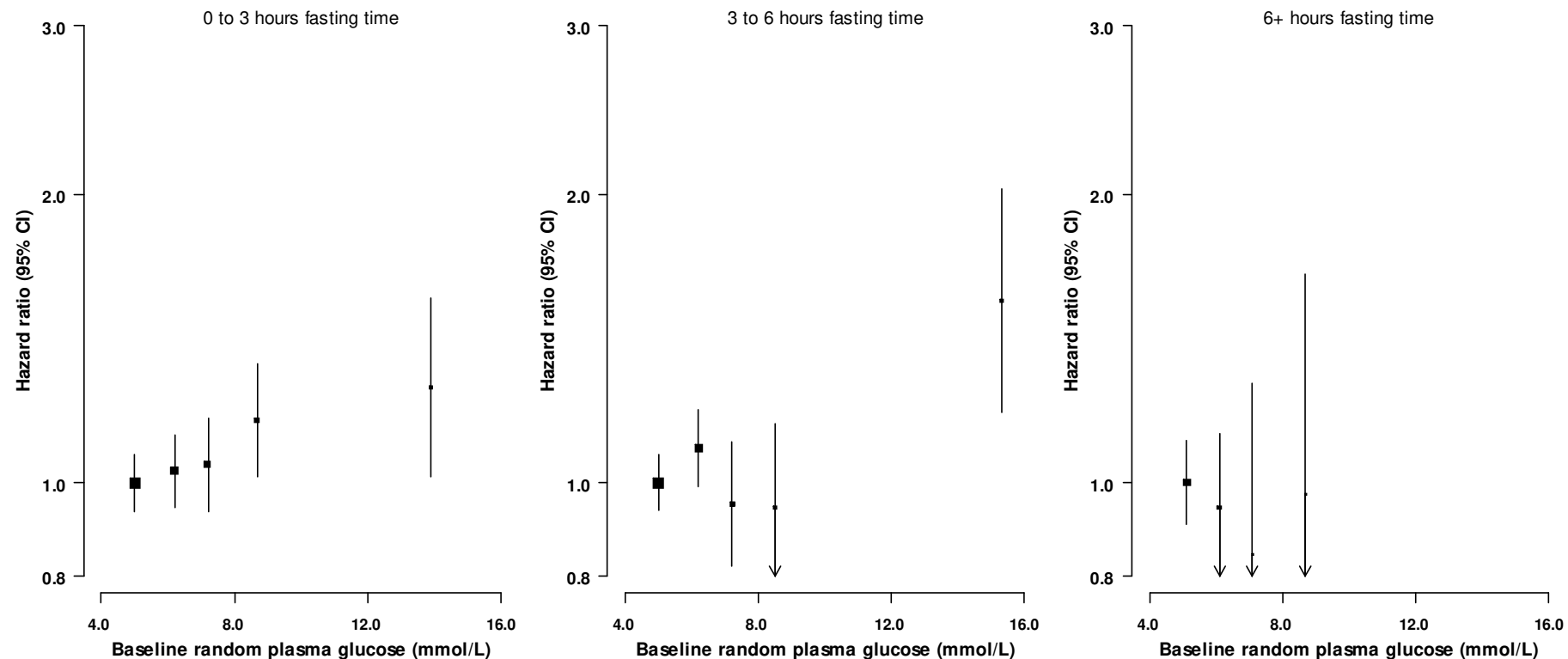


Figure 7.16. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease per 1 mmol/L higher baseline random plasma glucose progressively adjusted for potential confounding factors

Squares represent the hazard ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs).

BMI, body mass index; DBP, diastolic blood pressure; SBP, systolic blood pressure.

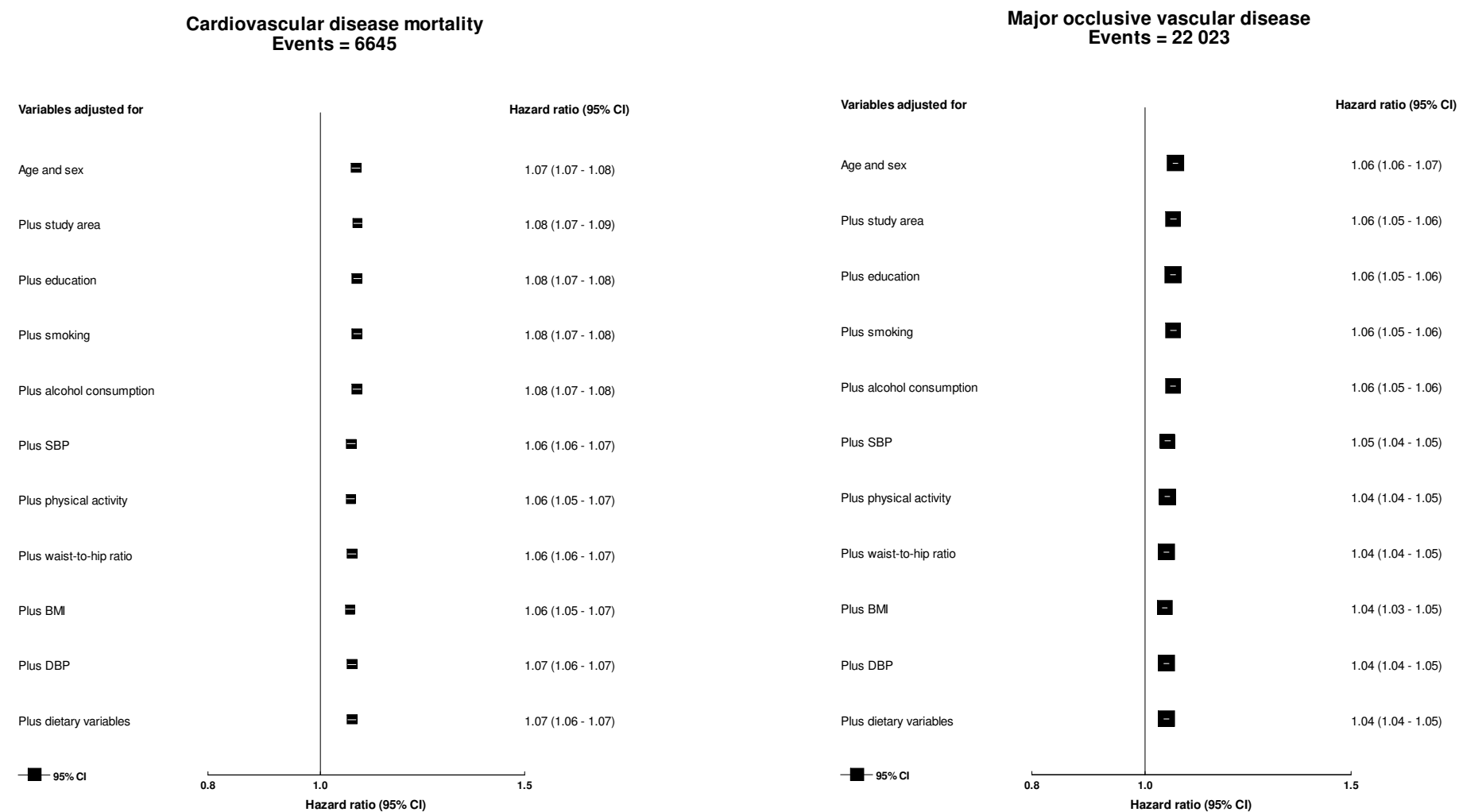


Figure 7.17. Adjusted hazard ratios for major coronary event and ischaemic stroke per 1 mmol/L higher baseline random plasma glucose progressively adjusted for potential confounding factors

Squares represent the hazard ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs).

BMI, body mass index; DBP, diastolic blood pressure; SBP, systolic blood pressure.

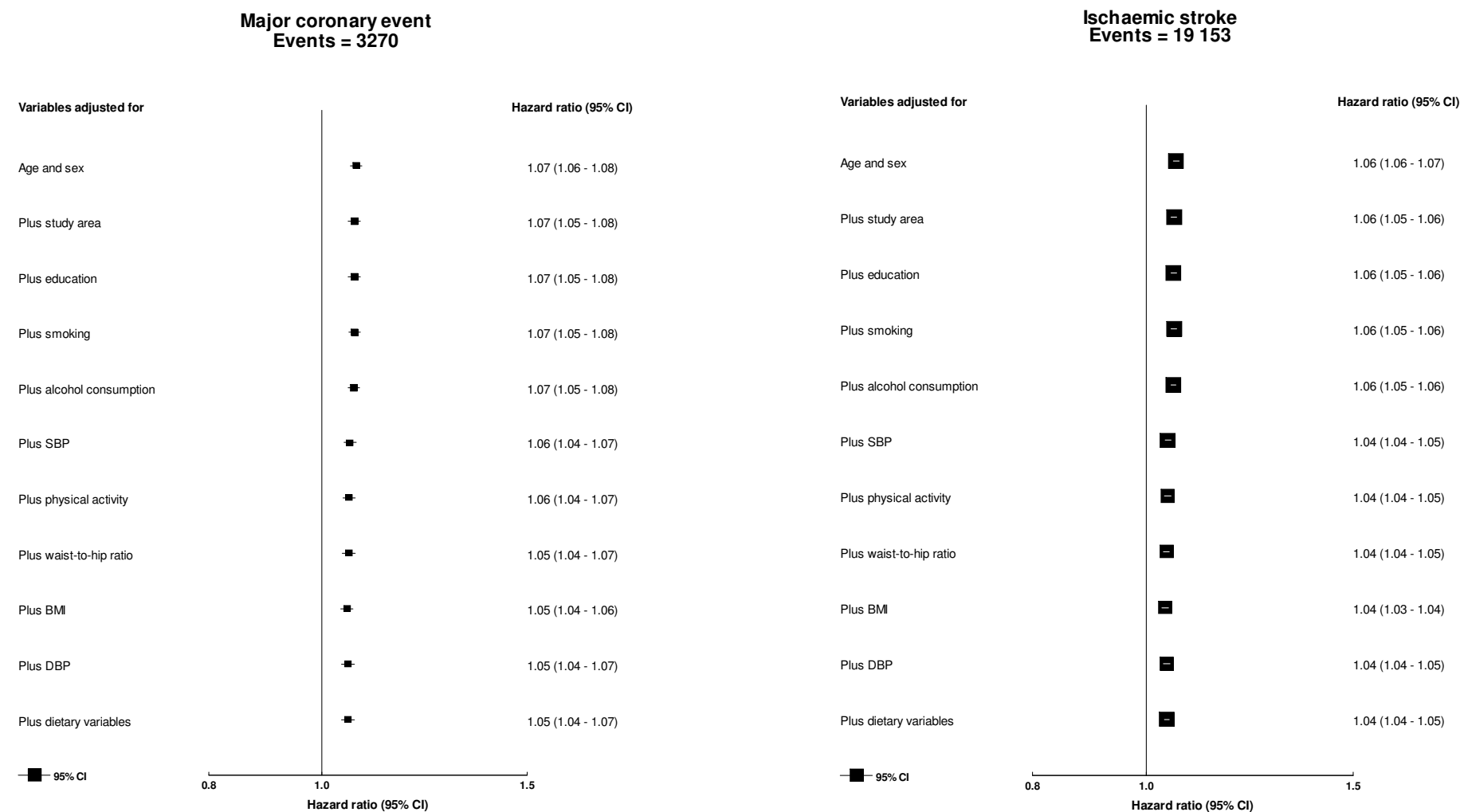


Figure 7.18. Adjusted hazard ratios for intracerebral haemorrhage per 1 mmol/L higher baseline random plasma glucose progressively adjusted for potential confounding factors

Squares represent the hazard ratio with area inversely proportional to the variance. Horizontal lines represent the corresponding 95% confidence intervals (CIs).

BMI, body mass index; DBP, diastolic blood pressure; SBP, systolic blood pressure.

Events = 4326

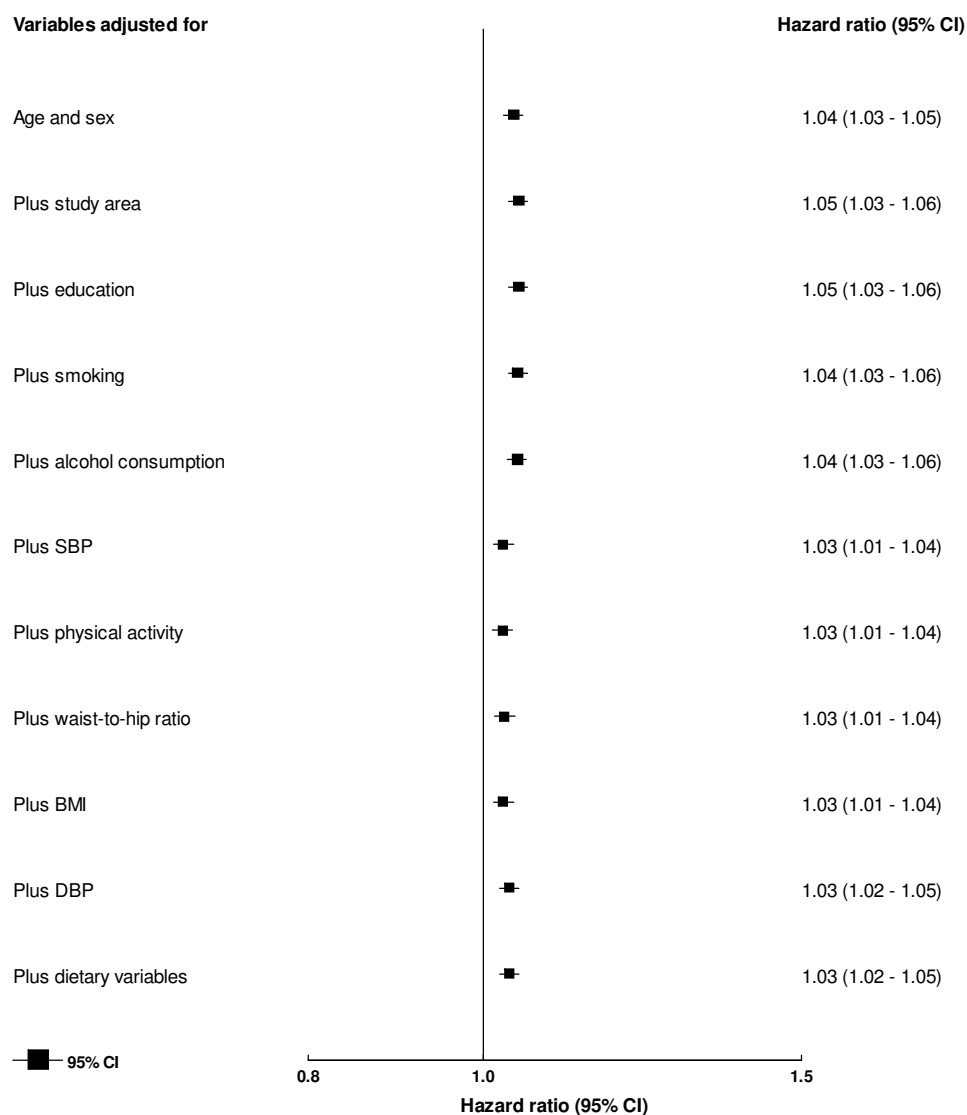


Figure 7.19. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease per 1 mmol/L higher baseline random plasma glucose, by study area

Stratified by age and sex and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall HR. The open diamond represents the overall HR and its 95% CI.

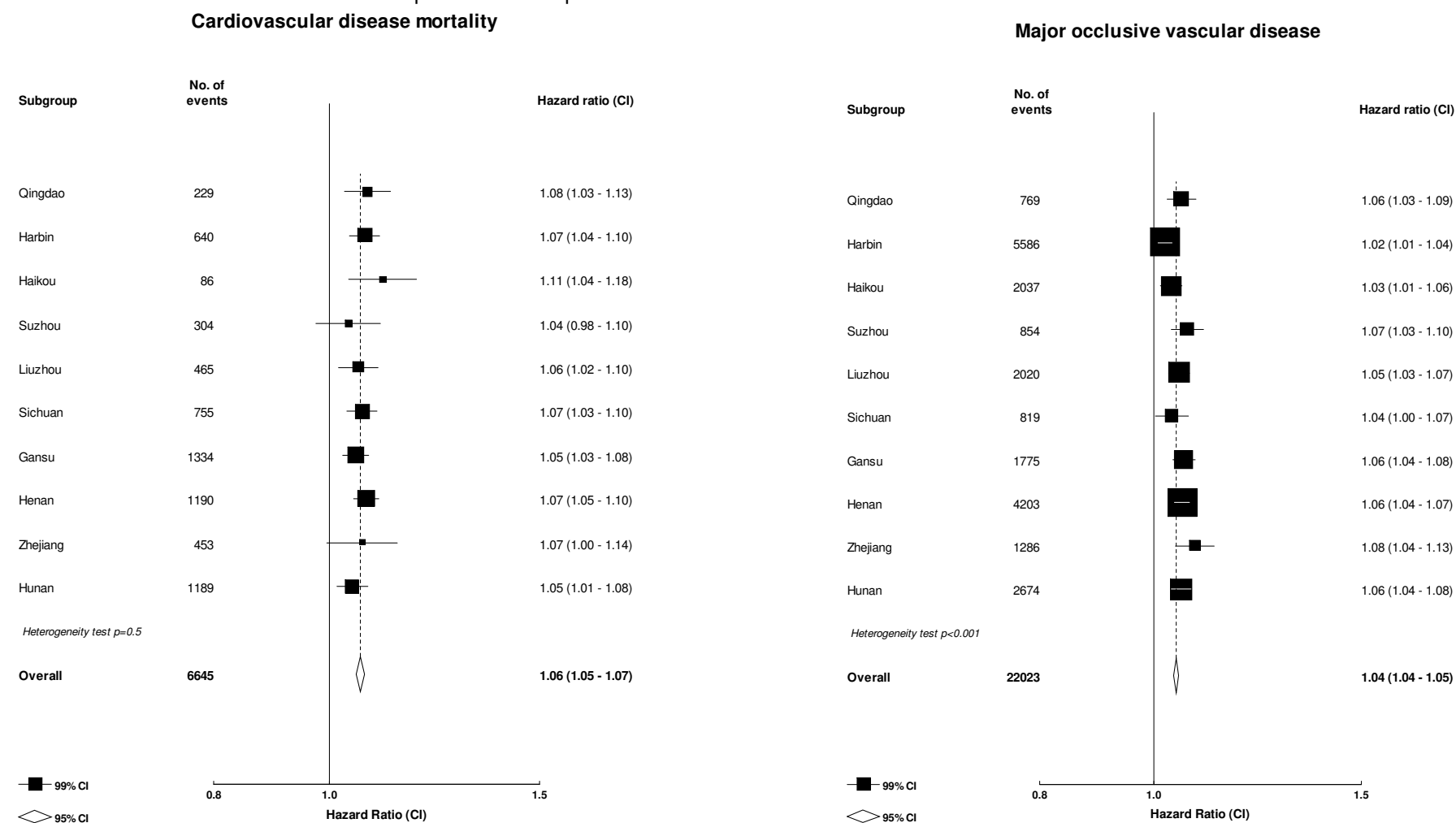


Figure 7.20. Adjusted hazard ratios for major coronary event and ischaemic stroke per 1 mmol/L higher baseline random plasma glucose, by study area

Stratified by age and sex and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall HR. The open diamond represents the overall HR and its 95% CI.

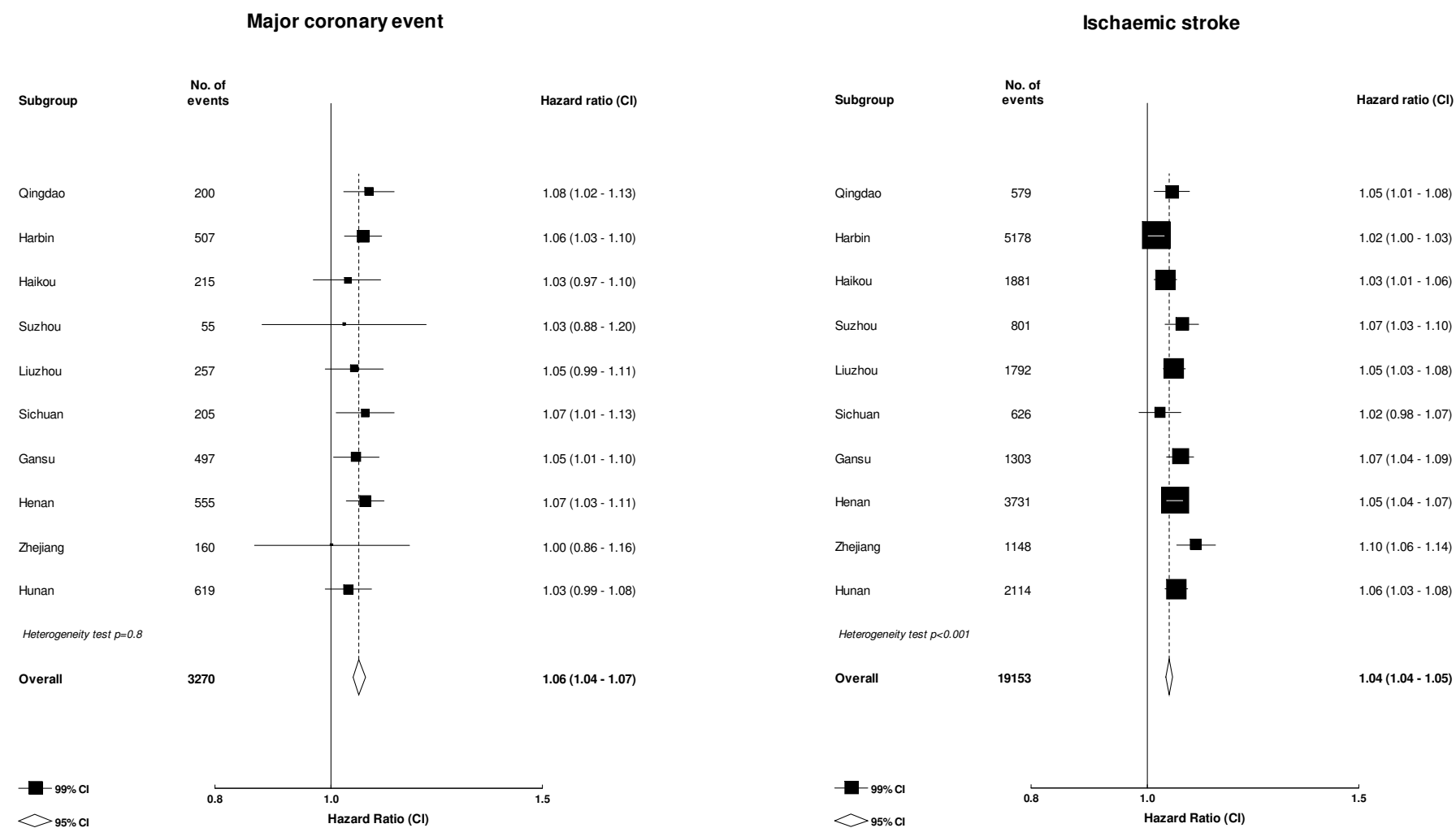


Figure 7.21. Adjusted hazard ratios for intracerebral haemorrhage per 1 mmol/L higher baseline random plasma glucose, by study area

Stratified by age and sex and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall HR. The open diamond represents the overall HR and its 95% CI.

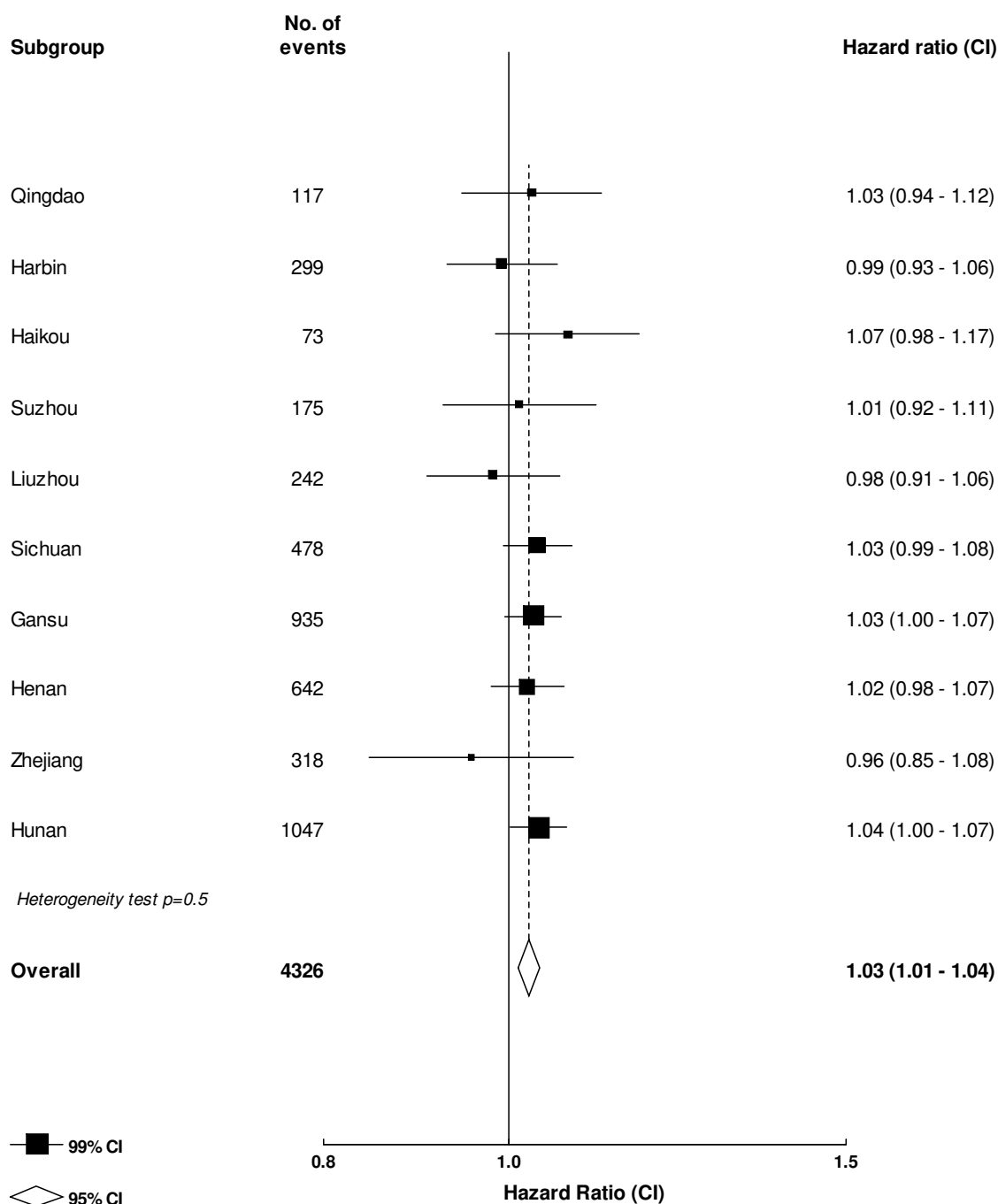


Table 7.3. Adjusted hazard ratios for incident cardiovascular diseases by baseline random plasma glucose excluding individuals recruited from Zhejiang

Baseline random plasma glucose (mmol/L) (Mean)	Cardiovascular disease mortality			Major occlusive vascular disease			Major coronary event			Ischaemic stroke			Intracerebral haemorrhage		
	No. of events	HR	95% CI	No. of events	HR	95% CI	No. of events	HR	95% CI	No. of events	HR	95% CI	No. of events	HR	95% CI
<4.3 (4.0)	336	1.00	0.90-1.11	951	1.00	0.94-1.07	170	1.00	0.86-1.17	800	1.00	0.93-1.07	252	1.00	0.88-1.13
4.3-5.2 (4.8)	1544	1.01	0.96-1.06	5456	1.07	1.04-1.10	813	1.02	0.95-1.09	4749	1.08	1.05-1.11	1097	1.02	0.97-1.09
5.3-5.7 (5.5)	1029	1.05	0.99-1.12	3814	1.11	1.07-1.14	493	0.96	0.88-1.05	3371	1.12	1.09-1.16	715	1.09	1.01-1.17
5.8-6.7 (6.2)	1551	1.11	1.05-1.16	5165	1.13	1.10-1.16	773	1.08	1.01-1.16	4488	1.14	1.10-1.17	1005	1.09	1.02-1.16
6.8-7.7 (7.2)	855	1.16	1.08-1.24	2813	1.23	1.19-1.28	432	1.18	1.07-1.30	2434	1.24	1.19-1.29	487	1.04	0.95-1.14
7.8-11.0 (8.8)	550	1.29	1.19-1.40	1663	1.29	1.23-1.35	285	1.33	1.19-1.50	1409	1.28	1.21-1.35	313	1.13	1.01-1.27
≥11.1 (15.8)	327	2.04	1.83-2.28	875	1.77	1.65-1.89	144	1.77	1.50-2.09	754	1.77	1.64-1.90	139	1.38	1.16-1.63
χ^2_1 (<i>p</i> value) for trend		149.1	(<0.001)		240.2	(<0.001)		59.4	(<0.001)		195.2	(<0.001)		13.4	(0.04)
Per 1 mmol/L higher		1.06	1.05-1.07		1.04	1.04-1.05		1.06	1.04-1.07		1.04	1.04-1.05		1.03	1.01-1.04

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, physical activity and systolic blood pressure.
CI, confidence interval; HR, hazard ratio.

Table 7.4. Adjusted hazard ratios for incident cardiovascular diseases by baseline random plasma glucose excluding individuals with new-onset of diabetes during follow-up

Baseline random plasma glucose (mmol/L) (Mean)	Cardiovascular disease mortality			Major occlusive vascular disease			Major coronary event			Ischaemic stroke			Intracerebral haemorrhage		
	No. of events	HR	95% CI	No. of events	HR	95% CI	No. of events	HR	95% CI	No. of events	HR	95% CI	No. of events	HR	95% CI
<4.3 (4.0)	351	1.00	0.90-1.11	976	1.00	0.94-1.07	178	1.00	0.86-1.16	819	1.00	0.93-1.07	266	1.00	0.89-1.13
4.3-5.2 (4.8)	1708	1.02	0.97-1.07	5808	1.07	1.04-1.10	865	1.00	0.93-1.07	5054	1.08	1.05-1.11	1225	1.02	0.96-1.08
5.3-5.7 (5.5)	1118	1.06	1.00-1.13	3956	1.10	1.07-1.14	517	0.94	0.86-1.03	3489	1.12	1.08-1.16	756	1.05	0.98-1.12
5.8-6.7 (6.2)	1618	1.11	1.06-1.17	5161	1.12	1.09-1.15	780	1.05	0.98-1.13	4475	1.13	1.10-1.16	1048	1.08	1.02-1.15
6.8-7.7 (7.2)	866	1.18	1.10-1.26	2695	1.21	1.16-1.26	416	1.13	1.02-1.24	2323	1.22	1.17-1.27	493	1.04	0.95-1.14
7.8-11.0 (8.8)	536	1.32	1.22-1.44	1479	1.24	1.17-1.30	266	1.30	1.15-1.47	1238	1.22	1.15-1.29	310	1.16	1.04-1.30
≥11.1 (15.8)	271	2.27	2.01-2.56	567	1.52	1.40-1.65	110	1.78	1.48-2.15	473	1.48	1.35-1.62	111	1.47	1.22-1.78
χ^2_1 (<i>p</i> value) for trend		169.1 (<0.001)			103.0 (<0.001)			51.5 (<0.001)			71.5 (<0.001)			17.4 (0.008)	
Per 1 mmol/L higher		1.07	1.06-1.08		1.03	1.03-1.04		1.06	1.05-1.08		1.03	1.02-1.04		1.03	1.02-1.05

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, physical activity and systolic blood pressure.
CI, confidence interval; HR, hazard ratio.

Table 7.5. Adjusted hazard ratios for incident cardiovascular diseases by baseline random plasma glucose excluding individuals with screen-detected diabetes at baseline

Baseline random plasma glucose (mmol/L) (Mean)	Cardiovascular disease mortality			Major occlusive vascular disease			Major coronary event			Ischaemic stroke			Intracerebral haemorrhage		
	No. of events	HR	95% CI	No. of events	HR	95% CI	No. of events	HR	95% CI	No. of events	HR	95% CI	No. of events	HR	95% CI
<4.3 (4.0)	355	1.00	0.90-1.11	996	1.00	0.94-1.07	180	1.00	0.86-1.16	837	1.00	0.93-1.07	272	1.00	0.89-1.13
4.3-5.2 (4.8)	1727	1.01	0.97-1.06	5952	1.07	1.04-1.09	883	1.00	0.94-1.07	5184	1.08	1.05-1.11	1242	1.00	0.95-1.06
5.3-5.7 (5.5)	1132	1.05	0.99-1.12	4113	1.11	1.07-1.14	530	0.95	0.87-1.03	3637	1.13	1.09-1.16	772	1.03	0.96-1.11
5.8-6.7 (6.2)	1643	1.10	1.05-1.16	5456	1.13	1.10-1.16	806	1.06	0.99-1.14	4750	1.15	1.11-1.18	1073	1.06	1.00-1.12
6.8-7.7 (7.2)	863	1.17	1.09-1.25	2768	1.23	1.19-1.28	421	1.14	1.04-1.26	2399	1.25	1.20-1.30	490	1.01	0.93-1.11
7.8-11.0 (8.8)	488	1.26	1.15-1.37	1382	1.23	1.16-1.29	237	1.23	1.08-1.40	1169	1.22	1.15-1.29	288	1.10	0.98-1.24
χ^2_1 (<i>p</i> value) for trend		25.7 (<0.001)			61.9 (<0.001)			16.1 (0.007)			51.5 (<0.001)			3.1 (0.7)	
Per 1 mmol/L higher		1.04	1.03-1.06		1.05	1.03-1.06		1.06	1.03-1.09		1.04	1.03-1.06		1.02	0.99-1.04

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, physical activity and systolic blood pressure.

Screen-detected diabetes: no self-reported diabetes with a random plasma glucose level ≥ 7.0 mmol/L if time since last food ≥ 8 hours, or ≥ 11.1 mmol/L if time since last food <8 hours, or a fasting plasma glucose level ≥ 7.0 mmol/L on subsequent testing.

CI, confidence interval; HR, hazard ratio.

Chapter 8. Assessing the causal association between random plasma glucose and cardiovascular disease: a Mendelian randomisation study

8.1 Background

As described in Chapter 2, several studies in predominantly Western populations have investigated the association of blood glucose levels with CVD risk in the general population.^{6, 83, 102-105, 142, 144} These have generally shown positive associations, but with some inconsistencies, particularly in the shape of the association at the lower end of the blood glucose distribution and, to some extent, in the strength of association with specific CVD subtypes.^{6, 102, 103} Analyses of data from the CKB have shown independent, continuous, positive associations of RPG with the risk of major ischaemic CVDs and CVD mortality, and a significant, though more modest, association with the risk of ICH (Chapter 7). The continuous, independent nature of the association among individuals without prior CVD (reducing the risk of reverse causality) is suggestive of a causal association, at least with ischaemic CVDs and CVD mortality. However, such observational epidemiological analyses are susceptible to residual confounding, by measured and unmeasured known, and also unknown, confounding factors, and to the effects of biases, measurement error and reverse causality, preventing conclusions regarding the causal nature of the associations. RCTs overcome many of these limitations, and provide the strongest evidence for establishing causality. However, RCTs of glucose-lowering interventions in individuals with diabetes^{270, 274, 276-280, 321} and pre-diabetes³¹⁰⁻³¹³ have, so far, been inconclusive with regards to CVD risk, and are complicated by the tendency for such interventions to also influence other CVD risk factors.^{270, 276} Furthermore, studies of pharmacological interventions to lower blood glucose levels may not be feasible in individuals with blood glucose levels in the “normoglycaemic” range. In such situations,

Mendelian randomisation can provide a useful tool for investigating the likely causal nature of associations.

Mendelian randomisation relies on Mendel's laws of independent assortment and segregation, as a result of which, genes influencing different traits assort into gametes and are inherited independently of each other.³²² Thus, although somewhat of an oversimplification, if a genetic variant directly influences an exposure and is also associated with a disease, but is not associated with other risk factors for the disease, the association is likely to be via the exposure of interest, suggesting a causal relationship.¹⁴⁵ Due to random allocation of genetic variants at conception, a variant would not be expected to be associated with the environmental and lifestyle factors which typically confound associations in observational epidemiology.^{145, 146} Furthermore, since an individual's genotype is assigned at conception, reverse causality will not influence the association of genotype with disease, and it can provide an indication of lifelong exposure levels. Genetic variants are also less susceptible to measurement error, which may impede findings from observational epidemiological studies.¹⁴⁵ Mendelian randomisation has thus been increasingly used as a tool for investigating the likely causal nature of relationships between exposures and diseases. GWAS, and meta-analyses of these, have identified multiple genetic variants associated with blood glucose, including in individuals without diabetes.^{80, 81, 323-326} This knowledge provides opportunity to investigate the likelihood of a causal relationship of blood glucose with CVDs using a Mendelian randomisation approach.

Based on ~90 000 genotyped individuals in the CKB, this chapter aims to assess the nature of the relationship between RPG levels and ischaemic CVDs. The objectives of this chapter were to a) Examine the association of plasma glucose-related SNPs with RPG levels in the CKB; b) Develop a GRS for RPG; c) Examine the association of plasma glucose-related SNPs and of the RPG GRS with ischaemic CVDs in the CKB; and d) Use a Mendelian randomisation approach to assess the likely causal nature of the relationship between RPG and ischaemic CVDs.

8.2 Methods

8.2.1 Population

Overall, 95 680 CKB participants (including 66 outside the 30-79 years age range included in this thesis) were randomly selected from the entire CKB cohort for genotyping.

8.2.2 Genotyping

DNA was extracted from 800 µl of buffy coat using a magnetic bead purification method (KingFisher™ Flex Magnetic Particle Processors; Thermo Fisher Scientific Inc., USA), and genotyping was undertaken using a customised genotype array containing a panel of 384 SNPs (GoldenGate Genotyping Assay; Illumina® Inc., USA). SNPs were selected for inclusion on the array based on a range of criteria (Section 3.5), including SNPs with demonstrated associations with outcomes of interest, and functional variants and SNPs with pleiotropic effects on multiple outcomes of interest or with particular biological significance. The selected SNPs included 56 type 2 diabetes-related genetic variants (Table B.1). Quality control procedures resulted in exclusion of 2483 samples due to gender mismatch and other potential linkage errors, low sample call rates (<98%) and high heterozygosity frequency. 93 131 participants (37 677 men and 55 454 women) remained after applying these criteria (Figure G.1). One autosomal SNP was excluded due to significant divergence from Hardy Weinberg equilibrium (HWE) and SNPs with a low call rate (<95%) were excluded. Genotyping concordance was between 98.7% and 100% for 2063 pairs of duplicate samples. These quality control procedures were undertaken by Dr Robin Walters, Senior Scientist in Genetic Epidemiology and Statistical Genetics in CTSU.

SNPs were selected for inclusion in analyses examining the relationship of RPG with CVD if they had demonstrated associations, at a genome-wide significance level ($p < 5 \times 10^{-8}$), with fasting or post-load plasma glucose in MAGIC (a meta-analysis of data from GWAS investigating glycaemic and metabolic traits),^{80, 81, 326} or if they had demonstrated associations with blood glucose in East Asian populations in published literature. A total of

10 SNPs were identified for inclusion: rs10811661 (*CDKN2A/B*), rs10830963 (*MTNR1B*), rs13266634 (*SLC30A8*), rs1552224 (*CENTD2*), rs340874 (*PROX1*), rs4607517 (*GCK*), rs780094 (*GCKR*), rs7901695 (*TCF7L2*) and rs11708067 (*ADCY5*), from MAGIC and rs1470579 (*IGF2BP2*), shown to be associated with FPG in a Han Chinese population-based cohort including 1908 individuals without diabetes.³²⁷

8.2.3 Exposure definitions

Plasma glucose levels were handled as described previously (Sections 4.2.2, 5.2.1 and 7.2.1). The small number of recorded plasma glucose levels above the maximum reading (27.8 mmol/L) on the SureStep Plus glucometer were imputed as 27.8 mmol/L for these analyses (Section 4.2.2).

8.2.4 Outcome definitions

Information on CVD events was obtained from the CDC disease reporting system, the national HI system and from death certificates (Section 3.7). The CVD outcomes investigated were MCE (non-fatal MI [ICD-10 I21-23] or fatal IHD [I20-25]), IS (I63) and MOVD (IS, non-fatal MI or fatal IHD). Detailed definitions of these disease outcomes are presented in Table E.1. CVD mortality was not examined due to smaller event numbers and the composite nature of this outcome. ICH was also not included in view of the less clear association with RPG levels in observational analyses and the relatively small number of events (Chapter 7).

8.2.5 Participant exclusions

In addition to the participant exclusions applied in the quality control procedure (Section 8.2.2) and exclusion of individuals <30 years or >79 years of age at baseline (who were not included in other analyses in this thesis), participants were excluded from analyses if they had missing data for RPG or BMI, or extreme or implausible data for SBP, DBP, height, waist circumference, hip circumference, waist-to-hip ratio or BMI at baseline (n=1142) (Figure G.1). Participants with self-reported diabetes at baseline were also excluded

(n=2914) since hypoglycaemic medication or lifestyle interventions initiated in response to diagnosis would be expected to have influenced RPG levels. After applying these criteria, 89 075 participants (36 112 men and 52 963 women) remained for inclusion in the current analyses.

Participants with self-reported doctor-diagnosed IHD or stroke/TIA at baseline or an incident major vascular event (fatal I00-02, I05-13, I15, I20, I24-28, I30-52, I62, I65-68, I69.2, I69.8, I70-74, I77-89 I95, I97-99 and fatal or non-fatal I21-23, I60, I61, I63, I64, I69.0, I69.1, I69.3, I69.4) recorded during follow-up were additionally excluded from the control group.

8.2.6 Statistical analyses

Analyses examining the associations of measured baseline and usual RPG levels with the risk of incident MOVD, MCE and IS, as described in Chapter 7, were repeated in the 89 075 participants included in the presented genetic analyses.

Deviation from HWE was assessed for the 10 included SNPs in control participants, using the X^2 statistic. One SNP (rs13266634, *SLC30A8*) had a p value <0.001; this was not excluded from the main analyses, but additional analyses were conducted excluding this SNP. Linear regression analyses, adjusting for age, sex and study area, were undertaken to calculate per allele and per genotype effects on baseline RPG for each of the 10 included SNPs. The reference alleles were assigned so that all SNPs had the same direction of effect on RPG. A RPG allele count GRS was developed by summing the number of glucose-increasing alleles for each individual; this assumes the SNPs have equal and additive effects on RPG and that there are no interactions between included variants.¹⁴⁵ In addition to the relative simplicity of this approach, simulations have shown allele count GRSs provide unbiased estimates of associations with little loss of power, when compared with weighted GRSs.¹⁴⁵ Individuals with missing data for SNPs were excluded from GRS analyses. Linear regression analyses, adjusted for age, sex and study area, were used to estimate the association of the GRS (per glucose-increasing allele and categorised into <9, 9-10, 11-12

and >12 glucose-increasing alleles, to ensure reasonable numbers of participants in each group) with baseline RPG.

Cox proportional hazards models, stratified by age-at-risk, sex and study area, were used to estimate HRs for MOVD, MCE and IS per glucose-increasing allele for each SNP individually and for the RPG GRS and, using the floating absolute risk method,²⁴¹ across categories of the RPG GRS (<9, 9-10, 11-12 and >12). The predicted association of the RPG GRS with risk of MOVD, MCE and IS was estimated using a “triangulation” or ratio of coefficients approach (Figure 8.1).¹⁴⁵ The observed increase in baseline RPG per glucose-increasing allele was multiplied by the regression coefficient for the observed increase in risk of MOVD, MCE or IS per 1 mmol/L increase in usual RPG. Variance estimates for these effect sizes were calculated as $\beta_1^2\sigma_2^2 + \beta_2^2\sigma_1^2 + \sigma_1^2\sigma_2^2$, where β_1 and β_2 and σ_1^2 and σ_2^2 are regression coefficients and variance, respectively, for the association of the GRS with RPG and the association of RPG with MOVD, MCE or IS.³²⁸

To investigate potential pleiotropic effects of the glucose-related SNPs, ie, effects which may influence CVD risk other than through RPG, the prevalence and mean values of baseline characteristics were calculated across genotypes at each locus, standardised by 5-year age groups, sex and study area. An online power calculator^{329, 330} was used to estimate the power to detect a genetic association between RPG and CVDs in this study. Using event and participant numbers in the CKB, power was calculated based on the observed HRs for MOVD, MCE and IS per 1 SD (1.9 mmol/L) higher usual RPG in the CKB and a range of HRs around these estimates (equivalent to HRs of between 1.04 and 1.24 per 1 mmol/L higher usual RPG), the coefficient of determination for the association between the RPG GRS and RPG (ie, the amount of variance in RPG explained by the GRS, or r^2) and a 0.05 significance level.

SAS version 9.3 (SAS Institute Inc., Cary, North Carolina, USA) was used to conduct all statistical analyses. A macro developed within CTSU was used to undertake the Cox

proportional hazards analyses. Line graphs and forest plots were produced using R version 2.13.1 (R Foundation for Statistical Computing, Vienna, Austria).

8.3 Results

Among the 89 075 CKB participants included in these analyses, the mean (SD) age was 51.1 (10.6) years and 59.5% were women. Mean (SD) baseline RPG was 5.9 (2.0) mmol/L.

8.3.1 Random plasma glucose and risk of cardiovascular diseases: observational analyses

As described in Chapter 7, there was a positive, log-linear association between RPG levels and incident MOVD, MCE and IS in the 467 508 CKB participants without self-reported diabetes at baseline. After adjusting for the effects of age, sex, study area, education, smoking, alcohol, physical activity and SBP, each 1 mmol/L higher usual RPG was associated with an 8% (HR 1.08, 1.07-1.09) increased risk of MOVD and 10% (1.10, 1.08-1.13) and 8% (1.08, 1.07-1.09) increased risks of MCE and IS, respectively (Table 8.1). During the 7-year follow-up period, 4697 MOVD events, 761 MCE and 4054 IS were recorded among the 89 075 CKB participants included in these genetic analyses. The associations of RPG with incident CVDs in this randomly-selected subgroup were similar to those in whole study population, with HRs per 1 mmol/L higher usual RPG of 1.07 (1.05-1.09) for MOVD, 1.09 (1.04-1.14) for MCE and 1.07 (1.04-1.09) for IS (Table G.1).

8.3.2 Genotype and random plasma glucose

The associations of the 10 individual SNPs with baseline RPG levels are presented in Table 8.2. Several SNPs showed strongly significant associations with RPG levels, in particular rs10811661 (*CDKN2A/B*) ($p=2.4\times10^{-26}$), rs13266634 (*SLC30A8*) ($p=1.2\times10^{-9}$), rs1470579 (*IGF2BP2*) ($p=1.5\times10^{-8}$) and rs4607517 (*GCK*) ($p=1.0\times10^{-12}$). The proportion of variation in RPG levels explained by the individual SNPs was, however, very small; rs10811661 (*CDKN2A/B*) had the largest effect, but explained only 0.1% of variation in RPG.

Figure 8.2 shows the distribution of the RPG GRS, produced by summing the number of RPG-increasing alleles. The GRS was normally distributed, with a mean (SD) of 9.6 (1.9). It was strongly significantly associated with baseline RPG levels ($p < 1 \times 10^{-30}$), with increasing numbers of alleles associated with higher baseline RPG levels (Figure 8.3). Each additional glucose-increasing allele was associated with a 0.04 (95% CI 0.04-0.05) mmol/L increase in baseline RPG, with a 0.1 mmol/L increment between each RPG GRS category (<9, 9-10, 11-12 and >12 glucose-increasing alleles) (Figure 8.4). Despite a large F-statistic (158), the RPG GRS explained just 0.2% of the variation in RPG.

There were no apparent clinically significant associations of other measured CVD risk factors with the included SNPs (Table G.2). Some statistically significant differences were demonstrated, but these likely reflect the large sample size and multiple testing.

8.3.3 Genotype and risk of cardiovascular diseases

When glucose-related SNPs were examined individually, one SNP (rs1470579, *IGF2BP2*) was weakly associated with the risk of MOVD, after adjustment for age, sex and study area (HR 1.05, 95% CI 1.01-1.10) (Table 8.3). Rs1470579 (*IGF2BP2*) was also associated with the risk of IS (1.06, 1.01-1.11). The remaining 9 SNPs were not associated with the risk of MOVD or IS, and no individual SNPs were associated with MCE. The RPG GRS showed no significant association with incident MOVD, MCE or IS, with HRs of 1.00 (95% CI 0.98-1.01), 0.98 (0.94-1.02) and 1.00 (0.98-1.02), respectively, per additional glucose-increasing allele and the risk of MOVD, MCE and IS varied little across RPG GRS categories (p for trend=0.4, 0.07 and 0.6, respectively) (Figures 8.5 and 8.6).

8.3.4 Random plasma glucose and cardiovascular diseases: genetic analyses

Assuming a causal relationship between RPG and ischaemic CVDs, based on the 0.04 mmol/L increase in RPG per glucose-increasing allele, and the estimated HR of 1.08 (95%CI 1.07-1.09) per 1 mmol/L higher usual RPG from the observational analyses, the expected HR for MOVD per glucose-increasing allele was calculated as 1.0034 (1.0027-

1.0042) (Table 8.4). Similarly, the expected HRs for MCE and IS per glucose-increasing allele would be 1.0042 (1.0029-1.0056) and 1.0033 (1.0025-1.0041). These expected associations of the RPG GRS with ischaemic CVDs are consistent with the observed associations of the RPG GRS, although neither expected nor observed associations differed clearly from 1.0.

Based on the observed HR for MOVD per 1 mmol/L higher usual RPG (1.08), it was estimated that there was 6.4% power to detect an association between genetically-determined RPG and MOVD in the genotyped CKB population. For MCE and IS (based on HRs of 1.10 and 1.08, respectively) the power to detect an association was estimated to be 4.2% and 6.0%, respectively. Additional power estimates, based on HRs of between 1.04 and 1.24 for the associations of usual RPG with the CVDs investigated were consistently less than 25% (Table G.3).

Due to the risk of bias inherent in the use of internally-weighted GRSs and the availability of external weights based only on the associations of genotypes with fasting or post-load plasma glucose, and not RPG as measured in the CKB, the results of analyses based on internally- and externally-weighted GRSs are not presented. The findings of these analyses were, however, very similar to those presented for the allele count GRS. Similarly, analyses excluding rs13266634 (*SLC30A8*), which showed evidence of deviation from HWE, did not alter the findings.

8.4 Discussion

The presented analyses based on data from ~90 000 genotyped CKB participants, demonstrated that multiple SNPs previously shown to be associated with FPG or with post-load plasma glucose levels,^{80, 81, 323-326} and a GRS based on 10 such SNPs, were significantly associated with RPG in the CKB, although the proportion of variation in RPG levels explained by the genetic variants was very small. No association was found between the glucose-related SNPs, or the GRS based on these, and CVD risks, likely due to

insufficient statistical power. As such, it is not possible to draw any reliable conclusions from these analyses about the likely causal nature of the relationship between measured RPG levels and ischaemic CVDs in the CKB.

Despite lack of clear evidence from Mendelian randomisation analyses in the CKB, a number of factors are suggestive of a causal relationship between blood glucose and ischaemic CVDs. These include the positive association of blood glucose with CVD risk demonstrated in several large, prospective, observational studies (including the CKB),^{6, 102, 103} plausible biological mechanisms²⁴⁸ and the findings of long-term follow-up of some,²⁷⁸⁻²⁸⁰ although not all,^{276, 321} RCTs of intensive glycaemic control in diabetes. Several previous Mendelian randomisation studies have examined the relationship of blood glucose levels with manifest^{148, 150, 331} or subclinical¹⁴⁹ CVD. Data from several Copenhagen-based cohort studies, including a total of ~80 500 participants (of whom a small number were known to have diabetes) and ~14 000 IHD events and ~600 MI, found a positive association of five glucose-related alleles with RPG levels and with the risk of MI and IHD.¹⁴⁸ Using a two-stage least-squares regression approach, they demonstrated a likely causal relationship between RPG levels and IHD; each 1 mmol/L higher genetically-determined RPG was associated with IHD and MI HRs of 1.25 (95% CI 1.03-1.52) and 1.69 (1.28-2.23), respectively, slightly stronger than the association of measured usual RPG levels with IHD (1.18, 1.15-1.22) and MI (1.09, 1.07-1.11) in the studies. A causal relationship between blood glucose and CVD was also found in a study combining data on glucose-related genetic variants from MAGIC (~133 000 individuals without diabetes) with data from the CARDIOGRAMplusC4D Metabochip study, an IHD GWAS meta-analysis including ~64 000 IHD cases and ~131 000 controls.³³¹ Overall, 24 SNPs associated with FPG in MAGIC, and with no known pleiotropic effects on other CVD risk factors, were included in the analyses. Their associations with IHD were ascertained from CARDIOGRAMplusC4D data, which suggested each 1 mmol/L higher genetically-determined FPG in individuals without diabetes was associated with a 15% increased odds (OR 1.15, 95% CI 1.00-1.32) of IHD,

suggestive of a causal role for FPG in IHD, although the association was only marginally statistically significant. Another study using data from MAGIC and CARDIOGRAMplusC4D, investigated the association of 30 FPG-related SNPs (17 of which were included in the previous study) with the risk of IHD.¹⁵⁰ A similar OR for IHD, although not statistically significant, was found (1.18, 0.97-1.42 per 1 mmol/L higher genetically-determined FPG), which was attenuated slightly (1.16, 0.94-1.42) after adjusting for the SNPs' effects on other CVD risk factors (hypercholesterolemia, hypertension and obesity). This study was restricted by limited statistical power (5.6%).¹⁵⁰ An additional study used data from the ARIC study to examine the association of a FPG GRS (based on five SNPs) with CIMT, a measure of subclinical atherosclerosis, in 7260 individuals without diabetes.¹⁴⁹ The internally-weighted FPG GRS was significantly associated with CIMT ($p=0.009$), with a 0.0048 mm increase per 1 SD (1.4) increment in the GRS. Using a ratio of coefficients approach, based on the observed 0.0103 mm increase in CIMT per 1 SD (0.5 mmol/L) higher FPG, it was estimated that each 1 SD increment in the FPG GRS would be expected to be associated with a 0.0018 mm increase in CIMT, lower than the observed association and consistent with a causal relationship. Although use of an internally-weighted GRS may have introduced bias, the authors reported that an externally-weighted GRS produced similar findings.

The lack of evidence of a causal association in the CKB contrasts with these previous findings. This likely reflects the lack of power in the CKB analyses, contributed to by the low proportion of variance in RPG explained by the SNPs and by the GRS. Two of the described previous studies reported that 2.9%¹⁴⁹ and 4.8%³³¹ of variation in glucose levels was explained by their investigated genotypes. In the CKB, only 0.2% of variation in RPG could be explained by the RPG GRS. The 10 SNPs included in the RPG GRS in the CKB were selected based on previous demonstrated associations with fasting or post-load plasma glucose, so their relationship with RPG may be less strong, although analyses limited to the 17 877 participants who reported not having eaten for at least 8 hours prior to blood

sampling produced similar null findings. Moreover, most of the SNPs were selected based on studies in populations of predominantly European descent,^{148-150, 331} and their relevance for Chinese populations, in whom such data are limited, is less clear. Although the magnitude of effects of the included genetic variants on RPG in the CKB were similar to their effects on FPG or post-load plasma glucose in MAGIC, the proportion of variance in glucose levels explained by the variants was considerably lower in the CKB.^{80, 81}

8.4.1. Strengths and limitations

Despite the null findings of the presented analyses, the study design has several strengths. An important assumption of Mendelian randomisation studies is that the genotypes investigated are associated with the outcome of interest only through the exposure, and not through other risk factors for the disease, that is, they have no pleiotropic effects. Some relevant genes have been reported to be associated with potentially important biological mechanisms, for example *CDKN2A/B* may have direct effects on the vasculature or on inflammation to influence IHD risk,^{26, 332} *GCKR* has been associated with elevated lipid levels³³³ and *TCF7L2* has been associated with blood pressure and lipids.³³⁴ Other relevant baseline characteristics in the CKB did not, however, appear to be associated with genotype, although data are not currently available for lipid levels. Furthermore, use of multiple SNPs, likely influencing plasma glucose levels through different mechanisms, should reduce the risk of pleiotropy. This is the first study to investigate the causality of this association in a Chinese population, complementing previous studies in Caucasian populations,^{148-150, 331} the findings of which may not be generalisable to populations with different genetic architectures.

Population stratification may produce spurious genotype associations due to inclusion of population subgroups with differing allele frequencies and exposure or disease distributions. Adjustment for genetic principal components can adjust for population stratification, but this was not possible in the present analyses as they are currently under

development. However, inclusion of study area in the statistical models should, to some extent, have accounted for this. Furthermore, there was little variation in allele frequencies across the 10 study areas, despite statistically significant heterogeneity (Table G.4). LD, that is the non-random association of alleles at different loci due to their physical proximity, may act as a further source of confounding in Mendelian randomisation studies if a genetic variant under investigation is in LD with another variant associated with the disease of interest through an alternative mechanism. The null findings in these analyses, and absence of any clear association between genetically-determined RPG and CVD would suggest this did not importantly influence the presented analyses. In addition, data from Phase 3 of the 1000 Genomes project from the Han Chinese in Beijing and the Southern Han Chinese populations,³³⁵ showed no evidence of LD between the 10 investigated SNPs ($r^2 \leq 0.002$ where relevant).

Although SNPs were selected for inclusion in these analyses based on studies in populations of predominantly European descent, many have also been shown to be associated with blood glucose levels in Chinese^{327, 336-338} or other East Asian³²⁴ populations. SNPs were not selected for inclusion on the genotyping array on the basis of known associations with blood glucose levels; rather, a proportion of the 384 SNPs were selected due to known associations with type 2 diabetes. Although there is overlap in the genotypes associated with diabetes and with plasma glucose levels in individuals without diabetes, the included SNPs may not be those that would have been selected in order to examine associations with glucose levels.

The major limitation of these analyses is clearly the lack of statistical power to detect an association between genetically-determined RPG and CVD. Despite reasonable F statistics for the RPG GRS and for several individual SNPs, weak instrument bias (due to the very low amount of variability in RPG explained) may have influenced the presented findings. Another consideration is HWE, which refers to the balance of genotype and allele frequencies in a population and relates to the idea that, in a situation of random mating,

genetic variation in a population will remain constant between generations. Multiple factors can cause deviation from HWE, but in the context of Mendelian randomisation, testing for HWE provides a means of assessing missingness of genotype data which can produce false null findings.³³⁹ One of the 10 included SNPs (rs13266634, *SLC30A8*) showed evidence of deviation from HWE, but exclusion of this SNP did not alter the findings. Developmental compensation (or canalisation), the phenomenon whereby individuals compensate for their disadvantageous genotype through biological adaptation, may also bias Mendelian randomisation findings towards the null by weakening the association of the genetic variant with the disease of interest. However, given the relatively weak effect of the investigated genetic variants on RPG levels, it seems less likely that canalisation, which usually relates to more dramatic genetic effects,³⁴⁰ would be a major concern.

8.4.2. Conclusion

The presented analyses used a Mendelian randomisation approach to examine the causal nature of the relationship between RPG and ischaemic CVDs. They showed no significant association between genetically-determined RPG and the risk of ischaemic CVDs. However, since the study had limited statistical power, it is not possible to determine whether this is due to a non-causal relationship, or an inability to detect a causal relationship. Genome-wide genotyping using a custom-designed 800 000 SNP array, including the 384 SNPs originally investigated, is currently underway on a further subset of ~100 000 CKB participants. Once completed, it will be of value to repeat the presented analyses in the entire genotyped population, and including SNPs with stronger effects on RPG, when, combined with longer follow-up and a greater number of CVD events, there may be adequate power to demonstrate, or refute, the existence of a causal relationship. Furthermore, it may provide opportunity to identify novel RPG-associated SNPs to inform the construction of an improved RPG GRS.

Figure 8.1. “Triangulation” approach for calculation of predicted association of the random plasma glucose genetic risk score with cardiovascular disease risk

β_1 , regression coefficient of the random plasma glucose genetic risk score-random plasma glucose association; β_2 , regression coefficient of the random plasma glucose-CVD association. β_3 , observed regression coefficient of the random plasma glucose genetic risk score-CVD association.

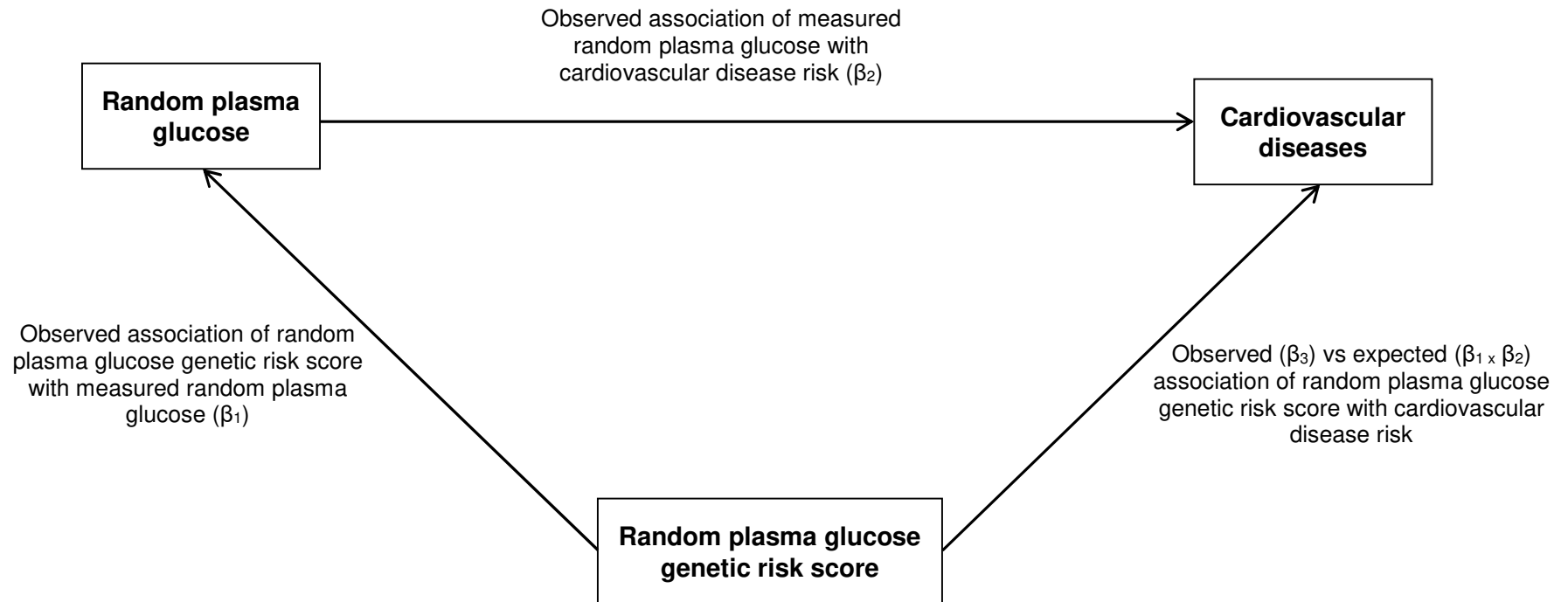


Table 8.1. Adjusted hazard ratios for incident cardiovascular diseases per 1 mmol/L higher baseline and usual random plasma glucose (n=467 508)

Disease outcome	No. of events	Baseline random plasma glucose		Usual random plasma glucose	
		HR	95% CI	HR	95% CI
Major occlusive vascular disease	22023	1.04	1.04-1.05	1.08	1.07-1.09
Major coronary event	3270	1.06	1.04-1.07	1.10	1.08-1.13
Ischaemic stroke	19153	1.04	1.04-1.05	1.08	1.07-1.09

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, physical activity and systolic blood pressure.

CI, confidence interval; HR, hazard ratio.

Table 8.2. Association of previously published plasma glucose-related SNPs with baseline random plasma glucose

SNP	Nearby gene	No. of participants	Alleles		EAF	Per allele effect ^a			% variance explained	F statistic
			Effect	Other		β	(SE)	p value		
rs10811661	<i>CDKN2A/B</i>	89007	C	T	0.541	0.097	(0.009)	2.4×10^{-26}	0.1	112.88
rs10830963	<i>MTNR1B</i>	89051	C	G	0.427	0.005	(0.009)	5.6×10^{-1}	0.0008	0.34
rs11708067	<i>ADCY5</i>	89071	A	G	0.996	0.145	(0.076)	5.8×10^{-2}	0.004	3.61
rs13266634	<i>SLC30A8</i>	88511	T	C	0.536	0.052	(0.008)	1.2×10^{-9}	0.04	36.99
rs1470579	<i>IGF2BP2</i>	89052	A	C	0.257	0.058	(0.010)	1.5×10^{-8}	0.03	31.82
rs1552224	<i>CENTD2</i>	89074	A	C	0.916	0.051	(0.016)	1.6×10^{-3}	0.01	9.91
rs340874	<i>PROX1</i>	89052	T	C	0.390	0.008	(0.009)	4.1×10^{-1}	0.002	0.67
rs4607517	<i>GCK</i>	89058	A	G	0.211	0.079	(0.011)	1.0×10^{-12}	0.04	50.85
rs780094	<i>GCKR</i>	89064	T	C	0.487	-0.014	(0.009)	1.2×10^{-1}	0.006	2.53
rs7901695	<i>TCF7L2</i>	88346	C	T	0.031	0.067	(0.026)	1.0×10^{-2}	0.007	6.62

^aAdjusted for age, sex and study area.

EAF, effect allele frequency; SNP, single nucleotide polymorphism.

Figure 8.2. Distribution of random plasma glucose genetic risk score

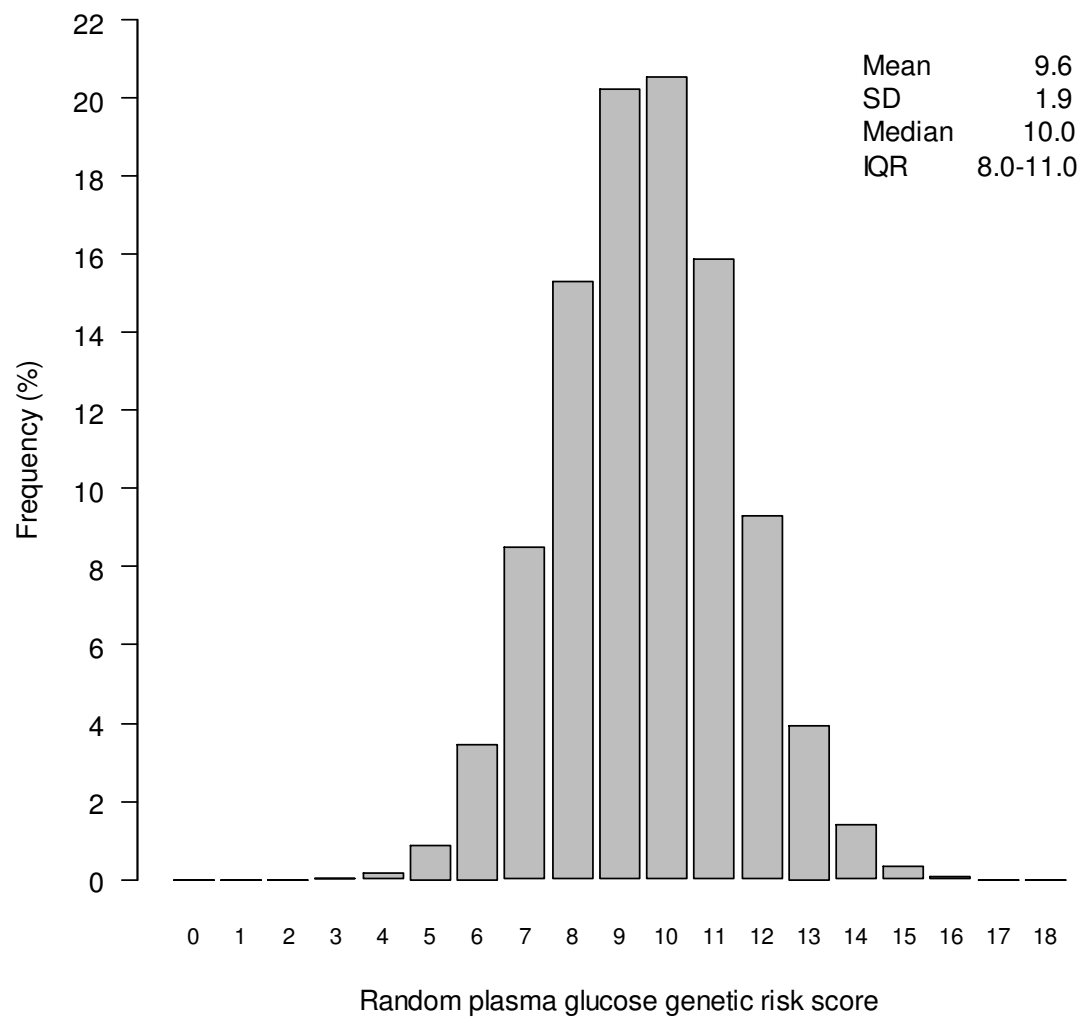


Figure 8.3. Mean baseline random plasma glucose levels by random plasma glucose genetic risk score

Adjusted for age, sex and study area.
Squares represent mean random plasma glucose with area inversely proportional to the variance.
Vertical lines represent the corresponding 95% confidence intervals.

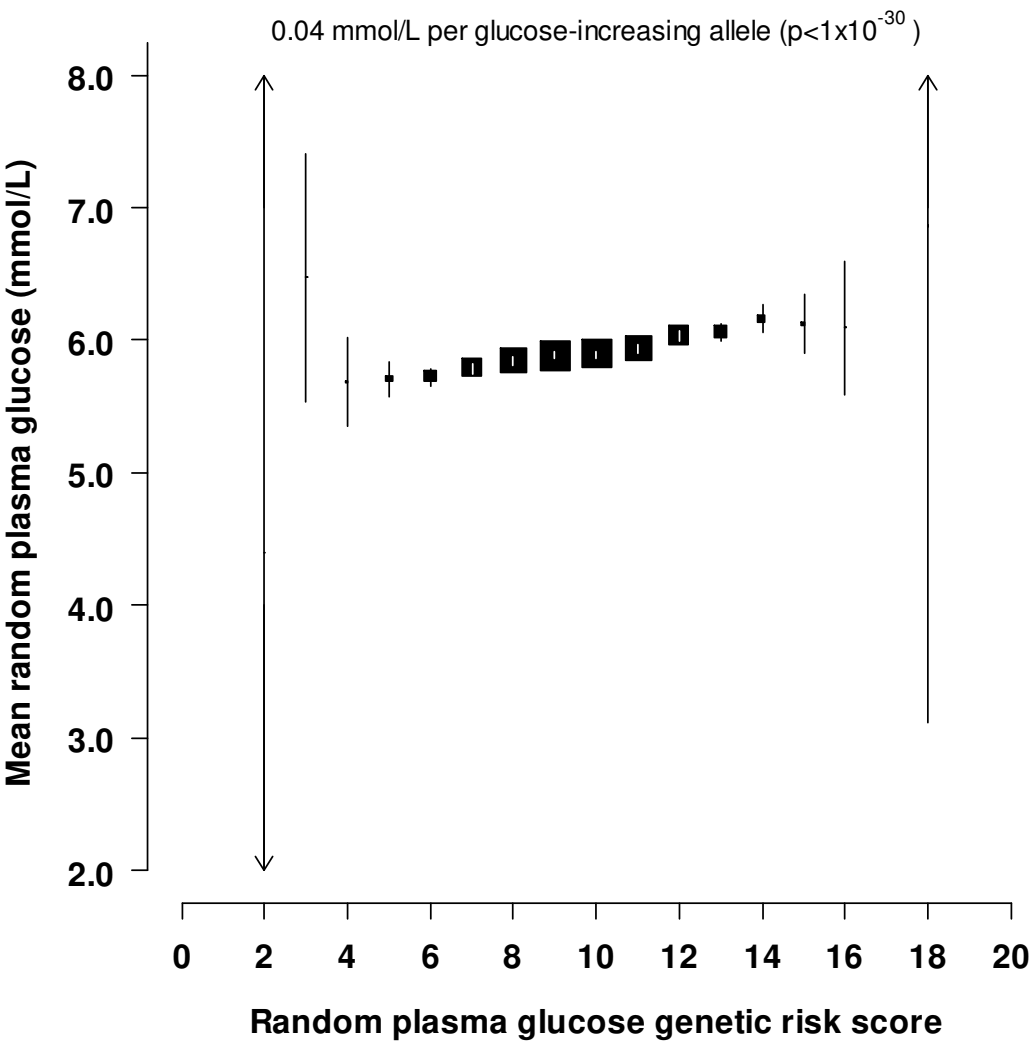


Figure 8.4. Mean baseline random plasma glucose by categories of random plasma glucose genetic risk score

Adjusted for age, sex and study area. Mean baseline random plasma glucose (RPG) plotted against mean number of glucose-increasing alleles in each RPG genetic risk score (GRS) category. Squares represent mean RPG with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals. Numbers below the squares represent RPG GRS categories.

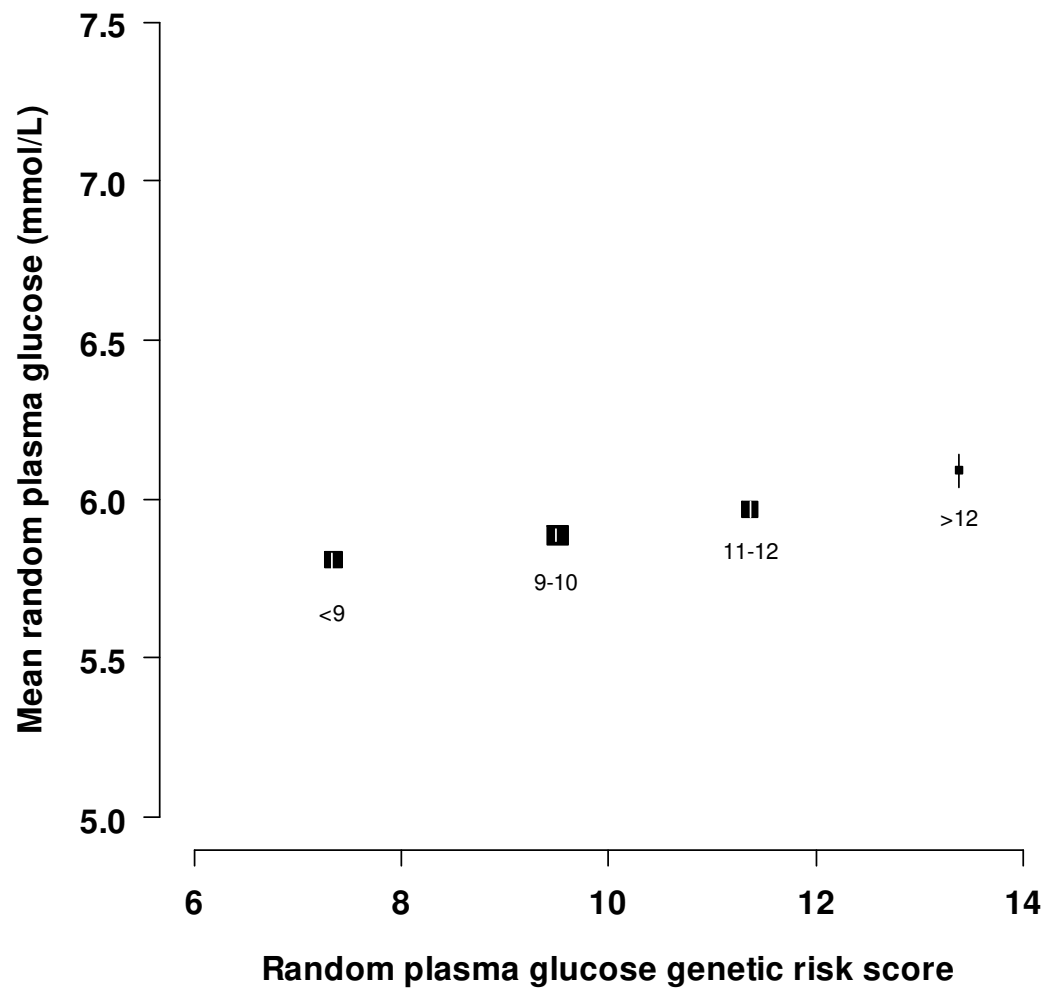


Table 8.3. Adjusted hazard ratios for incident cardiovascular disease per glucose-increasing allele

SNP	Nearby gene	Alleles		Major occlusive vascular disease			Major coronary event			Ischaemic stroke		
		Effect	Other	Events	HR ^{a,b}	95% CI	Events	HR ^{a,b}	95% CI	Events	HR ^{a,b}	95% CI
rs10811661	<i>CDKN2A/B</i>	C	T	4694	1.00	0.96-1.05	761	0.94	0.85-1.04	4051	1.01	0.97-1.06
rs10830963	<i>MTNR1B</i>	C	G	4696	1.00	0.96-1.04	760	0.99	0.89-1.09	4054	1.00	0.96-1.05
rs11708067	<i>ADCY5</i>	A	G	4697	0.98	0.71-1.35	761	1.11	0.50-2.49	4054	0.94	0.67-1.31
rs13266634	<i>SLC30A8</i>	T	C	4678	0.97	0.93-1.01	759	0.91	0.83-1.00	4037	0.97	0.94-1.02
rs1470579	<i>IGF2BP2</i>	A	C	4695	1.05	1.01-1.10	761	1.04	0.93-1.17	4052	1.06	1.01-1.11
rs1552224	<i>CENTD2</i>	A	C	4697	1.01	0.94-1.09	761	1.13	0.94-1.36	4054	0.99	0.92-1.07
rs340874	<i>PROX1</i>	T	C	4697	0.99	0.95-1.03	761	1.01	0.91-1.12	4054	0.99	0.94-1.03
rs4607517	<i>GCK</i>	A	G	4696	1.03	0.98-1.08	761	1.09	0.96-1.23	4053	1.02	0.97-1.07
rs780094	<i>GCKR</i>	T	C	4694	0.97	0.93-1.01	761	0.90	0.81-0.99	4051	1.02	0.97-1.06
rs7901695	<i>TCF7L2</i>	C	T	4670	0.95	0.85-1.07	756	0.99	0.75-1.32	4031	0.93	0.81-1.05

^aPer effect allele. ^bStratified by age, sex and study area.

CI, confidence interval; HR, hazard ratio; SNP, single nucleotide polymorphism.

Figure 8.5. Adjusted hazard ratios for major occlusive vascular disease and major coronary event by categories of random plasma glucose genetic risk score

Stratified by age, sex and study area. Hazard ratios (HR) are plotted against mean number of glucose-increasing alleles in each random plasma glucose genetic risk score category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs). Numbers above the CIs are the HRs and numbers below the CIs are of cardiovascular disease events. Random plasma glucose genetic risk score categories: <9, 9-10, 11-12, >12.

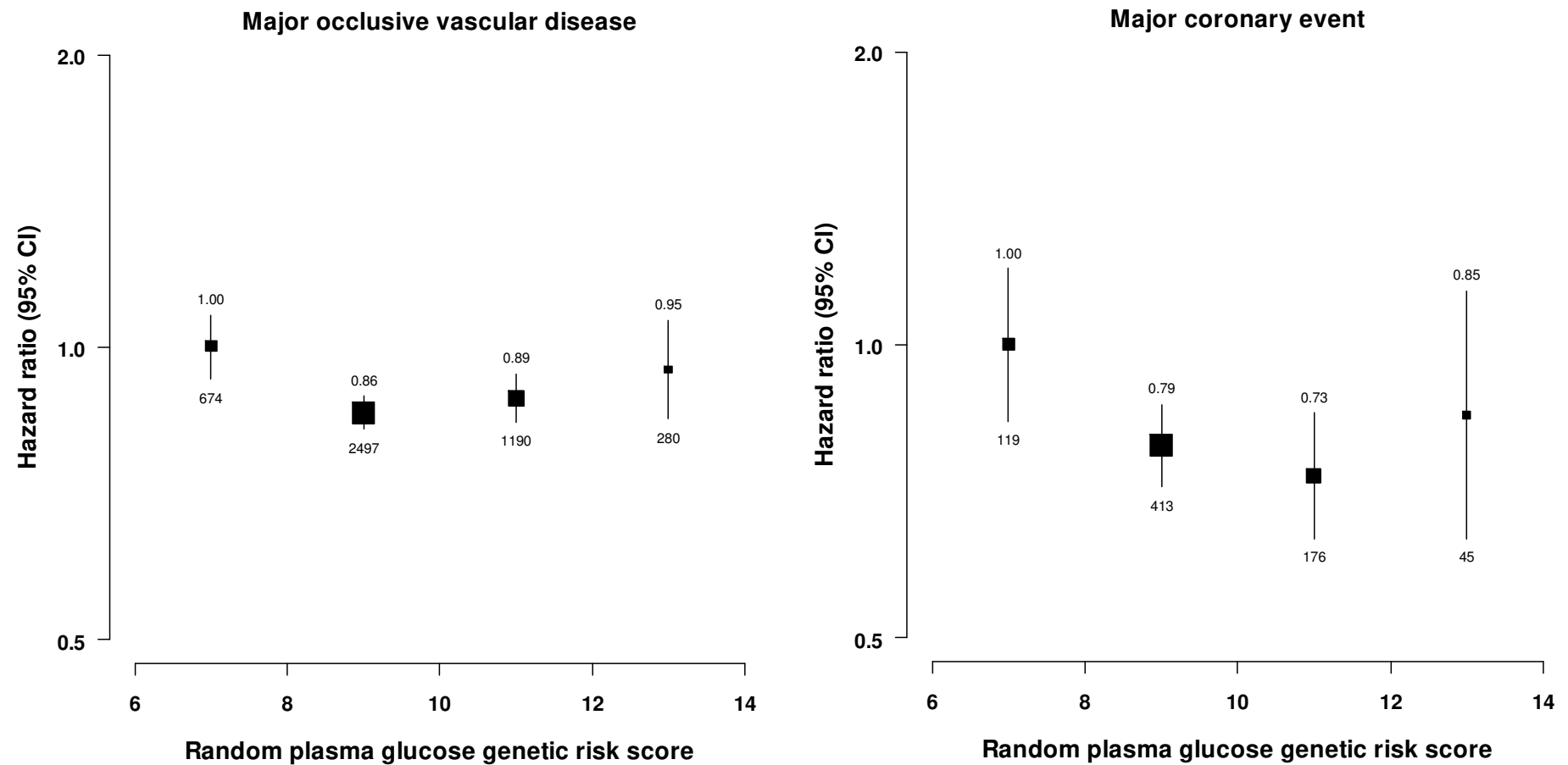


Figure 8.6. Adjusted hazard ratios for ischaemic stroke by categories of random plasma glucose genetic risk score

Stratified by age, sex and study area. Hazard ratios (HR) are plotted against mean number of glucose-increasing alleles in each random plasma glucose genetic risk score category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs). Numbers above the CIs are the HRs and numbers below the CIs are of ischaemic stroke events. Random plasma glucose genetic risk score categories: <9, 9-10, 11-12, >12.

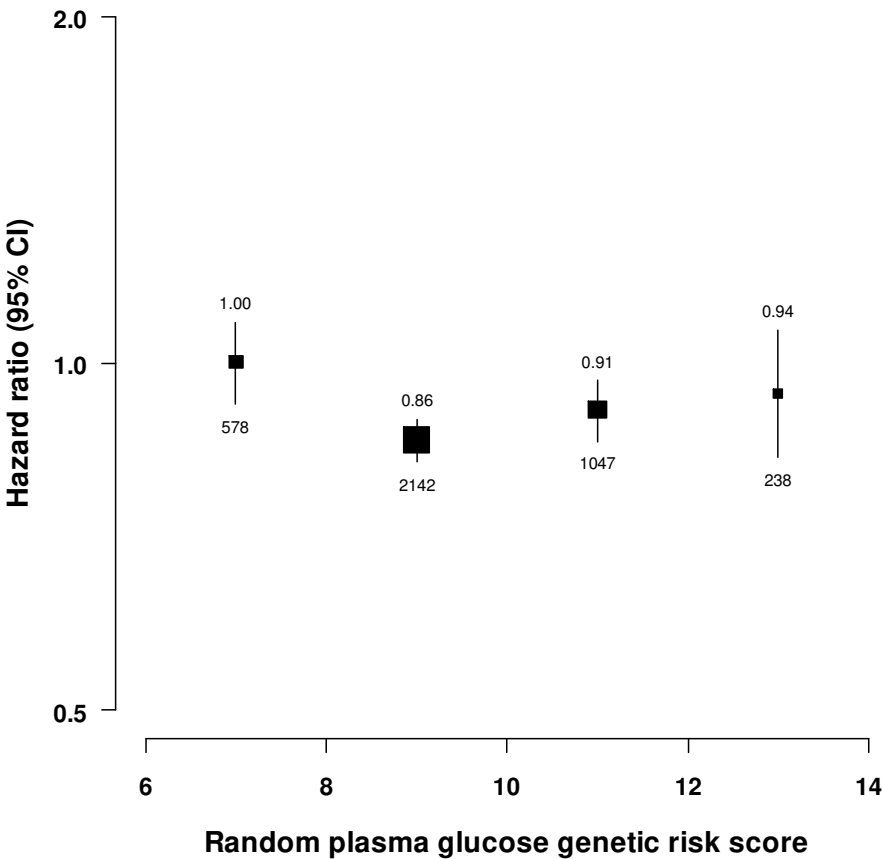


Table 8.4. Observed and Mendelian randomisation expected hazard ratios for incident cardiovascular diseases per glucose-increasing allele in the random plasma glucose genetic risk score

Disease outcome	Events	Observed		Mendelian randomisation expected		p value ^c
		HR ^{a,b}	95% CI	HR ^a	95% CI	
Major occlusive vascular disease	4641	1.00	0.98-1.01	1.0034	1.0027-1.0042	0.7
Major coronary event	753	0.98	0.94-1.02	1.0042	1.0029-1.0056	0.2
Ischaemic stroke	4005	1.00	0.98-1.02	1.0033	1.0025-1.0041	0.7

^aHazard ratio per glucose-increasing allele. ^bStratified by age, sex and study area; ^cP value for difference between observed and Mendelian randomisation expected associations.
CI, confidence interval; HR, hazard ratio.

Chapter 9. Discussion

As set out in the main objectives (Section 2.6), this thesis has a) described the prevalence of diabetes and the distribution of plasma glucose levels in individuals without known diabetes within the CKB population; b) examined the association of diabetes and, among individuals without known diabetes, of plasma glucose with other CVD risk factors; c) assessed the association of diabetes, and of RPG in individuals without known diabetes, with prevalent IHD and stroke/TIA and how these associations are modified by other factors; d) characterised the association of diabetes, and of usual RPG in individuals without known diabetes, with the risk of incident CVDs, both overall and in certain specific population subgroups; and e) explored the association of blood glucose-related genetic variants with CVDs in individuals without known diabetes and used a Mendelian randomisation approach to investigate the likely causal nature of any association. This chapter summarises the main findings and how these compare with published literature, discusses the likely causal nature of the investigated associations and the strengths and limitations of the study, and considers the potential clinical and public health implications of the findings as well as possible areas for future research leading on from these.

9.1 Summary of main findings

9.1.1 Distribution and correlates of diabetes and plasma glucose

Overall, 3.2% of CKB participants reported a history of doctor-diagnosed diabetes at baseline, with a high level of agreement between baseline and resurvey (>85%) (Chapter 4). A further 2.8% of participants had screen-detected diabetes. The prevalence of diabetes was higher in urban than in rural areas, and was associated inversely with physical activity and positively with age, blood pressure and adiposity. Self-reported diabetes prevalence was higher in ex-regular smokers and alcohol drinkers, whilst screen-detected diabetes prevalence was higher in ex-regular smokers only.

Among 487 552 CKB participants without known diabetes at baseline mean (SD) RPG was 5.9 (1.9) mmol/L. RPG increased with age and was modestly higher in urban, than in rural, areas. There was marked intra-individual variation in RPG levels, with a regression dilution ratio of 0.56. RPG levels were positively associated with blood pressure and with measures of adiposity, but varied little by participants' socioeconomic and lifestyle characteristics.

9.1.2 Diabetes, plasma glucose and prevalent cardiovascular disease

At baseline, 3.0% of participants reported a history of IHD and 1.7% a history of stroke/TIA (Chapter 5). After adjusting for potential confounding factors, there was an approximately 2-fold increase in the odds of prevalent IHD (OR 2.16, 95% CI 2.05-2.28) and stroke/TIA (1.88, 1.75-2.02) among those with self-reported diabetes. There was some evidence of a stronger association of self-reported diabetes with prevalent IHD in individuals with higher BMI and a longer duration of diabetes, and with prevalent stroke/TIA in women, more physically active individuals and ex-regular alcohol drinkers. Screen-detected diabetes was not clearly associated with prevalent IHD or stroke/TIA.

Among participants without self-reported diabetes at baseline, there was a positive association between baseline RPG and IHD and stroke/TIA prevalence after adjustment for potential confounding factors (Chapter 5). At RPG <11.1 mmol/L, there was an apparent log-linear association, with each 1 mmol/L higher baseline RPG associated with an adjusted OR of 1.04 (95% CI 1.02-1.05) for prevalent IHD and 1.05 (1.03-1.07) for prevalent stroke/TIA. The association of RPG with prevalent IHD appeared to be stronger at lower SBP levels, whilst for prevalent stroke/TIA it was stronger in rural areas. For both IHD and stroke/TIA, the associations were similar for fasting and non-fasting plasma glucose.

9.1.3 Diabetes, plasma glucose and incident cardiovascular diseases

Self-reported diabetes was associated with 1.5- to 2.5-fold increased risks of CVD mortality and incident ischaemic CVDs after adjustment for potential confounders in individuals without prior CVD (Chapter 6). There was also a significant, though more modest, increased

risk of ICH. For CVD mortality, the association was somewhat stronger in women, never-regular smokers and never-regular drinkers, while for MOVD and IS it was stronger in rural areas and in less well educated participants. The risk of all major ischaemic CVDs increased with longer duration of diabetes diagnosis and with a greater number of additional presenting cardiovascular risk factors. Screen-detected diabetes was associated with significant, but less extreme, excess CVD risks.

Among 467 508 study participants without self-reported diabetes, IHD or stroke/TIA at baseline there was a positive log-linear association between RPG and the risk of incident CVDs with no evidence of a threshold down to a usual RPG of at least 5.9 mmol/L (Chapter 7). It was estimated that each 1 mmol/L higher usual RPG was associated with ~10% increased risk of CVD mortality and major ischaemic CVDs. A weaker association was seen for ICH. There was some evidence of a stronger association of RPG with CVD in men than in women, and of a stronger association with MOVD and IS in younger age groups. These associations persisted after exclusion of individuals newly diagnosed with diabetes during the follow-up period.

9.1.4 The causal relevance of plasma glucose in cardiovascular disease

A GRS comprising 10 plasma glucose-related SNPs exhibited a weak but highly significant association ($p < 1 \times 10^{-30}$) with RPG in individuals with no prior history of diabetes, with each additional plasma glucose-increasing allele associated with 0.04 mmol/L higher baseline RPG (Chapter 8). This RPG GRS showed no significant association with the risk of MOVD, MCE or IS, but power to detect such an association was low (<6.5%).

9.2 Consistency with published literature

Detailed discussion of the consistency of the findings of this thesis with published literature has been presented in individual chapters, and is therefore only summarised here.

9.2.1 Distribution and correlates of diabetes and plasma glucose

The CKB was not designed to be nationally representative. Despite this, the prevalence of diabetes (5.9%), based on self-report and plasma glucose measurement, in the CKB was broadly similar to that in a contemporaneous nationally representative and similarly-aged study population (5.5%).¹⁸⁰ It was higher, however, than the 2.7% prevalence in another similarly-timed study of adults aged ≥ 20 years.¹⁸¹ Two more recent surveys have found higher diabetes prevalence in the adult Chinese population (9.7%¹⁷⁹ among adults aged ≥ 20 years and 11.6%⁸⁷ among adults aged ≥ 18 years), largely reflecting higher screen-detected diabetes prevalence.^{87, 179} As well as use of different diagnostic methods and the non-representative nature of the CKB, these differences likely reflect temporal trends in diabetes prevalence in China. The described correlates of diabetes in the CKB were consistent with those described in other studies in both Chinese^{87, 89, 179-181, 192, 194, 195, 198-200, 202, 203, 206} and non-Chinese populations,^{65, 202, 205, 206} and are in accordance with current knowledge of the aetiology and pathophysiology of diabetes.^{197, 201, 204, 205}

The distribution of RPG has not previously been described in Chinese populations, nor in large-scale studies elsewhere. There are similarly limited previous descriptions of the correlates of RPG, but the associations with age,^{82, 83, 218, 220-222} sex,^{220, 221} adiposity^{82-84, 103, 223} and SBP^{82-84, 103} are consistent with those in previous studies of alternative blood glucose measures. Intra-individual variation in RPG in the CKB was slightly less marked than, but not inconsistent with, the one previous report.²²⁴ As would be expected, these were more extreme than previously reported intra-individual variation in FPG.^{6, 103}

9.2.2 Diabetes and cardiovascular disease

Previous cross-sectional investigations in China have largely found approximately 2.5-fold increased odds of IHD in individuals with diabetes, including the nationally representative CNDMDS¹¹³ and a study in rural Beijing.¹¹⁰ These studies were much smaller than the CKB, and the moderately lower diabetes-associated odds of IHD in the CKB (OR 2.16) may reflect differences in definitions of diabetes or IHD,^{110, 113} or more extensive adjustment for

potential confounding factors in the CKB.¹¹⁰ The estimated odds of IHD and stroke/TIA associated with self-reported diabetes in the CKB are generally more consistent with those from previous prospective studies of Chinese¹³⁰⁻¹³² and Western populations.^{6, 90, 133} Previous Chinese studies have found 1.5- to 3- fold increased risks of incident IHD associated with diabetes,¹³⁰⁻¹³² but these studies were again much smaller than the CKB¹³⁰⁻¹³² and frequently under-adjusted for confounding.^{131, 132} The largest studies of diabetes with incident IHD have been in predominantly Western populations. The ERFC individual participant data meta-analysis, including ~530 000 participants, showed an independent doubling of incident IHD risk associated with diabetes (combined self-reported and screen-detected),⁶ whilst the APCSC individual participant data meta-analysis of ~160 000 participants showed about a doubling of the risk of IHD mortality associated with predominantly self-reported diabetes.⁹¹ A study of 1.9 million adults in the UK, based on routine health records, found a more modest 1.5-fold increased risk of MI in diabetes.⁹⁰ The associations of self-reported diabetes with incident IHD in the CKB, with an adjusted HR for MCE of 2.44 (95% CI 2.18-2.73), are thus largely comparable with these previous prospective studies, as are estimated associations with CVD mortality.⁹¹

Previous cross-sectional^{110, 113, 120} and case-control studies^{118, 121-126} have generated inconsistent findings on associations of diabetes with stroke. Prospective studies have reasonably consistently found ~1.5- to 2.5-fold greater risks of TS or IS associated with diabetes.^{6, 90, 127, 129-132, 134-138} Thus, despite the cross-sectional design, the odds of prevalent stroke/TIA associated with self-reported diabetes in the CKB are reasonably consistent with these. Diabetes-associated IS risks in previous prospective studies have tended to be modestly higher than the 68% increased risk shown in the CKB,^{6, 90, 129, 131, 132, 136} which may reflect residual confounding in some previous studies,^{131, 132, 136} or possibly some misclassification of IS events in the CKB. There is more limited evidence on the association of diabetes with ICH.^{6, 90, 91, 129, 131, 136} Two large studies, the ERFC meta-analysis⁶ and the CALIBER study of 1.9 million UK adults,⁹⁰ found modest elevated risks, with adjusted HRs

of 1.56 (95% CI 1.19-2.05)⁶ and 1.28 (1.02-1.62),⁹⁰ respectively. The CKB is the largest study—in terms of ICH event numbers—to examine this association, providing important additional evidence of a modest, but significant, elevated risk in self-reported diabetes.

Previous studies have demonstrated higher diabetes-associated risks of IHD in women than in men.^{6, 133} For stroke, this sex-difference is less clear.^{6, 90, 138} With the exception of CVD mortality, no such sex-difference was evident in the CKB. The reasons for the variable findings are uncertain, but there is clearly scope for confounding in these sex comparisons. There is relatively limited evidence of effect modification by other factors in previous literature, although several studies have shown a stronger association in younger age groups,^{6, 90, 128, 130, 136} as seen for MOVD and IS in the CKB.

9.2.3 Plasma glucose and cardiovascular disease

The role of plasma glucose levels in CVD risk in individuals without diabetes is an area of continuing controversy. Adequate power to demonstrate an association is reliant upon large sample sizes, but even the largest studies to-date reveal mixed findings.^{6, 102-104} The ERFC meta-analysis⁶ and the KCPS¹⁰² both found J-shaped associations of baseline FPG levels with IHD risk, with the lowest risk at ~5.0 mmol/L. A similar J-shaped association with IS was seen in the KCPS.¹⁰² In contrast, the APCSC meta-analysis showed a positive, log-linear association of FBG with both incident IHD and TS that continued down to a usual FPG of 4.9 mmol/L (equivalent to a lower FBG than a baseline FBG of 5.0 mmol/L).¹⁰³ The relevance of the contrasting findings at the lower end of the FBG range is not entirely clear, as discussed in Section 7.4.1, but may include the effects of relatively small event numbers producing less precise risk estimates. The findings of this thesis—the only study to investigate this association in any detail in a Chinese population—are clearly more consistent with those of the APCSC,¹⁰³ possibly, in part, reflecting adjustment for regression dilution bias. The more modest association of RPG with ICH, than with ischaemic CVDs, in the CKB is consistent with findings from the KCPS, the only other study with adequate power to investigate this association.¹⁰²

No large individual studies have previously investigated the association of RBG levels with CVD risk. Einarson et al's meta-analysis combined existing published data on the association, showing a positive association with the risk of total CVDs (HR 1.21, 95% CI 1.01-1.25 per 1 mmol/L higher RBG), but not with CVD subtypes, likely reflecting inadequate power.¹⁰⁴ Although not adjusted for regression dilution bias, this rather imprecise risk estimate is reasonably consistent with RPG-associated CVD risks in the CKB.

9.2.4 The causal relevance of plasma glucose in cardiovascular disease

Three studies, all in Caucasian populations, have reported Mendelian randomisation analyses examining the causal relevance of blood glucose levels in CVD.¹⁴⁸⁻¹⁵⁰ Two studies, one investigating the association of RPG with IHD and including individuals with known diabetes,¹⁴⁸ and the other investigating the association of FPG with CIMT in individuals without known diabetes,¹⁴⁹ found significant associations between genetically-determined blood glucose levels and CVDs, consistent with the positive associations between measured glucose levels and the CVD outcomes and with a causal relationship. Comparison with the Mendelian randomisation analyses in this thesis is difficult, since the analyses presented herein were inadequately powered. Lack of power also limited investigation of the causal nature of the relationship of FBG levels with IHD risk in another study combining data from consortia of GWAS, again in Caucasian populations.¹⁵⁰ No other studies have examined this association in a Chinese population.

9.3 Causality

As described, limitations inherent in observational epidemiological studies limit causal inferences. Bradford Hill proposed a number of characteristics which might be considered when attempting to interpret the likely causal nature of associations between exposures and diseases,³⁴¹ and it is of value to review the presented analyses in the context of these.

1.) *Strength* of association was suggested to infer causality since stronger associations are less readily explained by confounding or bias. The magnitude of the associations of

diabetes with CVDs, and particularly with CVD mortality and ischaemic CVDs, favours a causal association. The identified 8-11% increased risks for these CVDs per 1 mmol/L higher usual RPG are comparably modest. However, when considered in the context of the full glycaemic range they represent a reasonably strong effect. The comparably weaker associations of diabetes and of RPG with ICH could arguably be more easily explained by residual confounding and bias. However, associations of both diabetes and RPG with all investigated CVDs persisted after additional adjustment for other potential confounders.

2.) *Consistency* refers to the idea that associations observed repeatedly in different populations and circumstances are more likely to be causal. Within the CKB, there was reasonable consistency in the associations of diabetes and of RPG with CVDs across the 10 diverse study areas, and the associations largely persisted in other subgroup analyses. As described, associations of diabetes in the CKB are reasonably consistent with previous studies in predominantly Western populations,^{6, 91, 103} and there are some previous similar descriptions of associations of blood glucose with CVDs.^{103, 104}

3.) *Specificity* suggests that if an exposure is associated only with one disease, or is more strongly associated with one disease than with others, then that association is more likely to be causal. This criterion cannot be fully assessed in this thesis, which focussed only on CVD, although the observed weaker associations with ICH, which has a different pathophysiology, goes some way towards fulfilling the specificity criterion. Diabetes is known to be associated with a range of complications affecting multiple systems,^{61, 127, 342} and blood glucose levels in individuals without diabetes have been associated with a range of causes of morbidity and mortality.³⁴² Whilst this might suggest against a causal relationship, the relevance of this criterion has been questioned.³⁴³

4.) *Temporality* observes that for a relationship to be causal the exposure must precede the disease. The cross-sectional analyses herein lack such temporal clarity, but the prospective analyses more clearly meet this criterion, particularly as they excluded individuals with prior CVD. However, self-reporting of personal medical history and the

current inability to exclude the first few years of follow-up increased the risk of reverse causality.

5.) *Biological gradient* in the association of an exposure with a disease supports a causal relationship. This criterion was clearly met in the association of RPG with CVD mortality and ischaemic CVDs and, although less clearly, ICH. Although less relevant to dichotomous exposures, the gradient of effect with duration of diabetes diagnosis could be considered as evidence of a biological gradient.

6.) *Plausibility* of an association given knowledge of the pathophysiology of the disease under investigation supports causality. Although the precise details of the mechanisms involved remain elusive, biological processes by which higher blood glucose levels promote and contribute to the atherosclerotic process have been demonstrated⁹²⁻⁹⁷ (Chapter 1). The biological bases of the associations of both exposures with ICH are, however, less clear.

7.) *Coherence* is similar to plausibility, and suggests that the conclusion that a causal relationship exists should not contradict knowledge of the natural history or biology of the disease. This would appear to hold for the associations of diabetes and of RPG with ischaemic CVDs, at least.

8.) *Experiment* has been interpreted in a variety of ways,^{343, 344} but Bradford Hill suggested that if action taken in response to an observed association prevents the disease, the association is more likely to be causal. The effect of reversal of diabetes to a non-diabetic state on CVD risk has not been studied. Some studies have examined the impact of prevention or delay of diabetes onset in individuals with pre-diabetes on CVD risk, with varied findings.^{345, 346} The Da Qing Diabetes Prevention Study RCT examined the impact of a lifestyle intervention (addressing diet or physical activity or both), compared with usual care, on diabetes incidence and cardiovascular mortality in almost 600 Chinese adults with IGT.³⁴⁶ Following a six-year intervention period, all participants received usual medical care and were followed up, on average, for a further 17 years, when the risks of diabetes (HR 0.55, 95% CI 0.40-0.76) and of CVD mortality (0.59, 0.36-0.96) were found to be significantly lower in the intervention group. This contrasts, however, with the findings of a

meta-analysis of 10 RCTs of interventions (lifestyle or pharmacological) aimed at preventing progression from pre-diabetes to diabetes (including earlier findings from the Da Qing Diabetes Prevention Study).³⁴⁵ Despite effectively delaying or preventing diabetes (0.83, 0.80–0.86), this meta-analysis found no beneficial effects of such interventions on the risk of CVD mortality (1.04, 0.61–1.78), a non-significant reduction in the risk of MI (0.59, 0.23–1.50) and a just significant reduction in the risk of stroke (0.76, 0.58–0.99). Relatively short follow-up in the included studies, as well as the relatively low risk of CVDs in individuals with pre-diabetes, likely contributed to these largely null findings. These same factors have precluded meaningful investigation of the effects of interventions to prevent diabetes in approximately 2800 individuals with IGT on the risk of CVD in the 15 year follow-up of the Diabetes Prevention Program.³⁴⁷ Participants in this trial were randomised to an intensive lifestyle intervention, metformin or placebo for 3 years, followed by a 1 year transition period when all participants were offered an intensive lifestyle intervention, followed by on-going quarterly lifestyle sessions. Both interventions were associated with significant reductions in diabetes incidence when compared with placebo,³⁴⁷ and, after 10 years' follow-up, intervention and control groups had comparable levels of various cardiovascular risk factors despite lower use of cardiovascular medications in the group initially randomised to the lifestyle intervention.³⁴⁸ It will be of considerable interest to observe whether longer follow-up of participants in this study confirms the encouraging findings of the Da Qing Diabetes Prevention Study. Mendelian randomisation studies, described as nature's RCT,¹⁴⁷ have shown findings consistent with a causal relationship between diabetes and CVDs.^{150, 331, 349} A handful of RCTs have examined the effects of glucose lowering in individuals with pre-diabetes on CVD risks, but have again been inconclusive.³¹⁰⁻³¹² Although no conclusions regarding the causal nature of the relationship of blood glucose with CVD can be reached from the Mendelian randomisation analyses in this thesis, similar investigations in different study populations support a causal association throughout the glycaemic range.^{148, 149}

9.) *Analogy* suggests that a causal relationship is more likely if a similar exposure-outcome association is deemed causal. There are no obvious relevant analogous associations for the association of diabetes, or of RPG, with CVDs. However, the importance of this criterion has been questioned, particularly given its subjective nature.³⁴³

Causality can never be proven,³⁴⁴ and Bradford Hill was clear that these criteria should not be used as a checklist to determine causality, but rather as a guide to the points to consider.^{341, 343} Overall, there is considerable evidence in support of a causal relationship between diabetes and ischaemic CVDs, with some, less strong, evidence in favour of a causal relationship with ICH. The findings of this thesis, in combination with the broader body of evidence, are also suggestive of a causal relationship between RPG levels and ischaemic CVDs below the diabetes threshold, although the nature of the association with ICH is again less clear.

9.4 Strengths and limitations of the study

Many strengths and limitations of the study and their likely impact on the findings presented have been discussed in previous chapters, and are only summarised here. The size of the CKB is a major strength, enabling robust and precise CVD risk estimates. Indeed, the CKB is probably the only study to-date with adequate power to robustly investigate the association of RPG with CVDs. Furthermore, the large number of well-phenotyped, incident CVD events ensured adequate power to investigate associations with CVD subtypes, including uniquely precise estimates of the associations of both diabetes and RPG with risk of ICH. Selection of study participants from the general population and unrelated to glycaemic or CVD status, should have reduced the risk of selection bias. Loss to follow-up was extremely low (0.5%). As a result of this, although participants lost to follow-up differed from the study population as a whole in various ways—they were more likely to live in urban areas, to be more highly educated and less physically active, to have higher BMI, but lower SBP, and to have prior diabetes—any resulting bias would be negligible.

Although the cross-sectional analyses lacked temporal clarity, the main prospective analyses were more robust in this regards, and less susceptible to reverse causality. Exclusion of individuals with a history of IHD or stroke/TIA at baseline further reduced the risk of reverse causality, although self-reporting of personal medical history increased the risk of misclassification, which is of particular concern given the relatively poor agreement in self-reporting of prior IHD and stroke/TIA at baseline and resurvey and the inability to validate self-reported medical history. Exclusion of incident CVD events ascertained during the first few years of follow-up, to ensure exclusion of individuals with pre-clinical or undiagnosed CVDs, would have provided an additional measure to prevent reverse causality, but the relatively short follow-up period currently available in the CKB precluded this. However, the comparability of CVD HRs in the first and second halves of follow-up provides some reassurance that the observed associations are not the result of reverse causality.

The diversity of the data collected in the CKB enabled comprehensive adjustment for confounding factors, although the current lack of data on lipids is an important limitation. Data collected in resurveys not only provided insight into the reproducibility of self-reported personal medical history, but also enabled assessment of, and adjustment for, intra-individual variation in RPG levels. Ideally, similar adjustments would have been made for within-person variation in confounding variables, but, as described (Section 7.4.2), this was not possible, likely contributing to some residual confounding of the association of RPG with CVD. Adjustment of RPG for fasting time may have improved the accuracy of regression dilution ratio estimates, but any such effect would have been minimal. Collection of data on disease outcomes through a number of pre-existing data sources reduced the risk of selection and recall biases, as well as limiting the effects of biases inherent in any individual data source. The diversity of the follow-up data enabled exclusion of individuals who were newly diagnosed with diabetes from analyses examining the association of RPG with CVD risks, considerably increasing the insight gained.

Although the study had many strengths, consideration of the limitations is also important for informed and balanced interpretation of the findings. The overall participation rate in the CKB was relatively low (~30%). This may have limited generalisability of the findings, but was unlikely to have introduced bias, since non-participation would be unlikely to be related to both the exposure and to other factors influencing CVD risk. Unfortunately, however, comparison of responders and non-responders is not possible. The analyses may have been susceptible to the “healthy volunteer” effect,³⁰⁵ potentially underestimating disease risks, and perhaps particularly diabetes-associated CVD risks since individuals with less well-controlled diabetes of longer duration, say, may have been less likely, or able, to participate. A potential criticism of the CKB more generally could be the lack of representativeness of the study population with regards to the adult Chinese population more generally. The importance of representativeness in epidemiological studies has been the source of considerable debate, and it is generally agreed that this depends upon the study aims.^{167, 350-353} Studies aiming to estimate the frequency of an exposure or disease in a population clearly require a representative population sample.^{167, 350, 353} The same is not generally true, however, of studies attempting to characterise the relationships of exposures with diseases.^{167, 350, 351, 353} Such studies do not require representative population samples, but instead are reliant on sufficiently diverse populations in terms of exposure, and limited loss to follow-up,³⁵⁰ requirements clearly met in the CKB. Thus, the lack of representativeness would not be expected to adversely impact on the aim of this thesis to determine the relationships of diabetes, and of blood glucose levels in individuals without known diabetes, with CVDs in Chinese adults and likely in East Asian populations more generally. Wider generalisability of the findings is less clear, although the consistency, at least of the associations of diabetes with CVD risks, with large studies in Western populations suggests the findings may be more widely generalisable.

For a range of frequently practical reasons, RPG measured on a point-of-care glucometer was the selected glycaemic measure in the CKB. As described (Sections 4.4.5 and 7.4.2),

in a general population study this is a practical, relevant and interesting exposure. However, as a means of diagnosing diabetes, even within the context of a large-scale population-based epidemiological study, and taking account of self-reported data on fasting time, it has limitations. Most importantly, this approach will most likely have led to a degree of misclassification of screen-detected diabetes (Section 4.4.2), underestimating associated CVD risks. The focus on self-reported diabetes in this thesis, in part, reflects these concerns.

The extent, diversity and detail of the follow-up data in the CKB are, as described, notable strengths of the study. However, although large numbers of CVD events were recorded, the smaller numbers of CVD deaths, MCE and ICH limited the power of some subgroup analyses. Furthermore, given the nature of the data sources used, under-reporting of less severe CVD events was likely. This contrasts with the self-reported CVD outcomes investigated in the cross-sectional analyses, which would clearly not include fatal events. Although preliminary validation of CVD diagnoses recorded during follow-up suggested a relatively high level of accuracy for stroke and for IHD overall, the accuracy of IHD subtypes was less good. Inaccuracies in recording of CVD diagnoses would most likely bias risk estimates towards the null; the on-going validation work will be important for better understanding this. Furthermore, although subtyping data were available for the majority of stroke events, the extent to which this was based on the findings of CT/MR scans is currently unclear for events recorded in the HI system.

The availability of genotyping data in the CKB represents a tremendous opportunity for investigation of genetic determinants of disease and gene-environment interactions, and for understanding of causal relationships. Unfortunately, there was inadequate power to demonstrate an association between genetically-determined RPG and the risk of CVD with the currently available data. However, once genome-wide data on an additional ~100 000 study participants are available, there may be adequate power to warrant repetition of these

analyses including additional genetic variants more strongly associated with plasma glucose levels.

9.5 Clinical and public health implications

Recent decades have seen marked and rapid increases in the prevalence of diabetes in China.^{65, 85-87} Despite this, there has been limited understanding of the consequences, including the cardiovascular consequences, of diabetes in this population. The 1.5- to 2.5-fold increased risks of CVD mortality and ischaemic CVDs associated with self-reported diabetes, and the more modest, but significant, 24% increased risk of ICH in the CKB now demonstrate the substantial impact of diabetes on CVDs in adult Chinese. Assuming a causal relationship, and applying age-specific population attributable fractions for CVD mortality in the CKB to diabetes prevalence, derived from the most recent nationally representative survey,⁸⁷ and the number of age-specific CVD deaths in mainland China,²³⁶ it was estimated that 489 676 (95% CI 335 777-681 202) cardiovascular deaths could be attributed to diabetes in China in 2010. Given the prediction that diabetes prevalence will continue to rise in China, with particularly marked increases expected in young adults,⁶⁵ this burden is likely to increase still further. This clearly highlights the importance of effective prevention of diabetes. It is generally accepted that economic development and associated urbanisation and mechanisation have played important roles in the rapid rise in diabetes prevalence in China over recent decades.^{86, 88, 89} Although these factors themselves are not feasible targets for diabetes prevention, the accompanying lifestyle changes contributing to escalating rates of obesity and diabetes are much more so. Interestingly, although the average caloric intake has remained stable, there have been reductions in levels of physical activity—particularly occupational physical activity—and changes in dietary composition, notably increased intake of refined grains, dietary fats, red meat and sugar-sweetened beverages, in the Chinese population over recent decades.⁸⁹ Initiatives addressing these population-specific insights will be essential for effective primary prevention of diabetes, with the Da Qing Diabetes Prevention Study providing valuable evidence of the potential

effectiveness of such interventions.³⁴⁶ In addition, tackling the high levels of tobacco smoking among men in China³⁵⁴ would be expected to contribute to prevention of diabetes in the population.⁷⁶ Furthermore, and clearly related to this, the findings of this thesis highlight the need for effective CVD prevention in individuals who already have diabetes, through a multifaceted approach addressing all components of cardiovascular risk.

Prevention of CVDs in individuals with diabetes might be considered a “high-risk” approach, as described by Rose.³⁵⁵ Rose also highlighted the importance of the “population strategy” for disease prevention, whereby the entire distribution of a risk factor in a population is shifted in a favourable direction, with likely small benefits for the individual, but much greater potential benefits at the population level. The positive log-linear association between RPG levels and risk of CVDs in individuals without known diabetes in the CKB highlights the potential benefit for CVD prevention of shifting the entire population blood glucose distribution to the left, perhaps, for example, through lifestyle interventions to tackle obesity. Assuming a causal relationship, based on the observed associations, each 1 mmol/L lower RPG would be expected to reduce CVD mortality by 11%, and incident MOVD, MCE, IS and ICH by 8%, 10%, 8% and 5%, respectively. Although comparably small effects when compared to the burden of CVDs attributable to diabetes, at a population level the impact would be considerable, particularly given China’s high levels of cardiovascular morbidity and mortality.

9.6 Future research

Future research should address various limitations of the presented analyses. Although the CKB currently includes more CVD events than many previous studies, longer follow-up and the accompanying additional CVD events would, in particular, valuably increase the precision of associations with CVD mortality, MCE and ICH, and provide opportunity for more detailed subgroup analyses. Furthermore, it would permit exclusion of CVD events recorded during the first few years of follow-up to minimise any residual impact of reverse causality. The average duration of diabetes diagnoses in the CKB was relatively short (5.7

years) and longer follow-up would also enable more detailed investigation of the relevance of diabetes duration to CVD risk. Measurement of blood lipid levels is currently underway on a subset of the CKB population; these data will provide a valuable opportunity to examine confounding and effect modification by lipids in the associations of diabetes and of RPG, with CVD risks. Furthermore, the genome-wide genotyping of up to 100 000 participants, using a custom-designed, 800 000 SNP array, will hopefully provide adequate power, through inclusion of additional participants and cardiovascular events, and possibly also additional SNPs with stronger effects on RPG, to enable investigation of the causal relevance of RPG, and additionally of diabetes, for CVDs.

Findings of the current stroke validation process will additionally provide opportunity to build on the findings of this thesis, particularly with regards to investigation of the associations of diabetes and of RPG with IS subtypes, including lacunar infarcts and small vessel disease, for which diabetes may be a particular risk factor,³⁵⁶ and which are thought to be particularly prominent in Chinese populations.⁵⁶ Future analyses would also valuably investigate the association of diabetes, and of RPG, with other cardiovascular and non-cardiovascular diseases, particularly peripheral vascular diseases, heart failure, aneurysmal conditions such as abdominal aortic aneurysm and SAH,⁹⁰ chronic liver and renal disease, retinopathy and various cancers.^{127, 342, 357} Analyses examining diabetes-associated cause-specific mortality risks in the CKB have shown associations with several of these conditions, with a particularly strong, greater than 10-fold increased risk of mortality due to chronic renal disease found in individuals with self-reported diabetes. In addition, it will be important to understand how the different complications of diabetes, including renal complications, influence the relationship of diabetes with CVDs.

The validity of the screen-detected diabetes definition used in this thesis is unclear. An exercise to attempt to better understand this, for example through laboratory-based measurement of FPG or HbA1c in the subset of the study population included in future resurveys would be useful. Finally, genotyping and measurement of blood-based

biomarkers, including in the context of proteomics and metabolomics, in the CKB will provide opportunity for investigation of mechanisms underlying the associations of diabetes, and of RPG, with CVDs and other conditions.

9.7 Conclusion

Despite the greater than 10% estimated prevalence of diabetes in the adult population of China,⁸⁷ there has been surprisingly limited understanding of the CVD consequences of diabetes in this population. These analyses, based on a diverse population of over 0.5 million Chinese adults, have shown that diabetes is associated with significantly elevated risks of CVDs among adults in mainland China, similar in magnitude to those observed in Western populations, and independent of other CVD risk factors.^{6,90} Furthermore, this study provides evidence of an independent, positive log-linear relationship of RPG with risk of major CVDs throughout the full glycaemic range in Chinese adults without prior diabetes. Further research is needed to determine the likely causal nature of both these associations in Chinese populations, and the overall impact of diabetes on a wider range of cardiovascular and non-cardiovascular diseases.

References

1. GBD 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age–sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015;385(9963):117-71.
2. World Health Organization. The global burden of disease: 2004 update. Geneva: World Health Organization; 2008 [29th April 2015]. Available from: http://www.who.int/healthinfo/global_burden_disease/en/.
3. Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, Aboyans V, Abraham J, Ackerman I, Aggarwal R, Ahn SY, Ali MK, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Bahalim AN, Barker-Collo S, Barrero LH, Bartels DH, Basanez M-G, Baxter A, Bell ML, Benjamin EJ, Bennett D, Bernabe E, Bhalla K, Bhandari B, Bikbov B, Bin Abdulhak A, Birbeck G, Black JA, Blencowe H, Blore JD, Blyth F, Bolliger I, Bonaventure A, Boufous S, Bourne R, Boussinesq M, Braithwaite T, Brayne C, Bridgett L, Brooker S, Brooks P, Brugha TS, Bryan-Hancock C, Bucello C, Buchbinder R, Buckle G, Budke CM, Burch M, Burney P, Burstein R, Calabria B, Campbell B, Canter CE, Carabin H, Carapetis J, Carmona L, Cella C, Charlson F, Chen H, Cheng AT-A, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahiya M, Dahodwala N, Damsere-Derry J, Danaei G, Davis A, De Leo D, Degenhardt L, Dellavalle R, Delossantos A, Denenberg J, Derrett S, Des Jarlais DC, Dharmaratne SD, Dherani M, Diaz-Torne C, Dolk H, Dorsey ER, Driscoll T, Duber H, Ebel B, Edmond K, Elbaz A, Ali SE, Erskine H, Erwin PJ, Espindola P, Ewoigbokhan SE, Farzadfar F, Feigin V, Felson DT, Ferrari A, Ferri CP, Fevre EM, Finucane MM, Flaxman S, Flood L, Foreman K, Forouzanfar MH, Fowkes FGR, Fransen M, Freeman MK, Gabbe BJ, Gabriel SE, Gakidou E, Ganatra HA, Garcia B, Gaspari F, Gillum RF, Gmel G, Gonzalez-Medina D, Gosselin R, Grainger R, Grant B, Groeger J, Guillemin F, Gunnell D, Gupta R, Haagsma J, Hagan H, Halasa YA, Hall W, Haring D, Haro JM, Harrison JE, Havmoeller R, Hay RJ, Higashi H, Hill C, Hoen B, Hoffman H, Hotez PJ, Hoy D, Huang JJ, Ibeanusi SE, Jacobsen KH, James SL, Jarvis D, Jasrasaria R, Jayaraman S, Johns N, Jonas JB, Karthikeyan G, Kassebaum N, Kawakami N, Keren A, Khoo J-P, King CH, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Laden F, Lalloo R, Laslett LL, Lathlean T, Leasher JL, Lee YY, Leigh J, Levinson D, Lim SS, Limb E, Lin JK, Lipnick M, Lipshultz SE, Liu W, Loane M, Ohno SL, Lyons R, Mabweijano J, MacIntyre MF, Malekzadeh R, Mallinger L, Manivannan S, Marcenes W, March L, Margolis DJ, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGill N, McGrath J, Medina-Mora ME, Meltzer M, Mensah GA, Merriman TR, Meyer A-C, Miglioli V, Miller M, Miller TR, Mitchell PB, Mock C, Mocumbi AO, Moffitt TE, Mokdad AA, Monasta L, Montico M, Moradi-Lakeh M, Moran A, Morawska L, Mori R, Murdoch ME, Mwaniki MK, Naidoo K, Nair MN, Naldi L, Narayan KMV, Nelson PK, Nelson RG, Nevitt MC, Newton CR, Nolte S, Norman P, Norman R, O'Donnell M, O'Hanlon S, Olives C, Omer SB, Ortblad K, Osborne R, Ozgediz D, Page A, Pahari B, Pandian JD, Rivero AP, Patten SB, Pearce N, Padilla RP, Perez-Ruiz F, Perico N, Pesudovs K, Phillips D, Phillips MR, Pierce K, Pion S, Polanczyk GV, Polinder S, Pope CA, 3rd, Popova S, Porrini E, Pourmalek F, Prince M, Pullan RL, Ramaiah KD, Ranganathan D, Razavi H, Regan M, Rehm JT, Rein DB, Remuzzi G, Richardson K, Rivara FP, Roberts T, Robinson

- C, De Leon FR, Ronfani L, Room R, Rosenfeld LC, Rushton L, Sacco RL, Saha S, Sampson U, Sanchez-Riera L, Sanman E, Schwebel DC, Scott JG, Segui-Gomez M, Shahraz S, Shepard DS, Shin H, Shivakoti R, Singh D, Singh GM, Singh JA, Singleton J, Sleet DA, Sliwa K, Smith E, Smith JL, Stapelberg NJC, Steer A, Steiner T, Stolk WA, Stovner LJ, Sudfeld C, Syed S, Tamburlini G, Tavakkoli M, Taylor HR, Taylor JA, Taylor WJ, Thomas B, Thomson WM, Thurston GD, Tleyjeh IM, Tonelli M, Towbin JA, Truelsen T, Tsilimbaris MK, Ubeda C, Undurraga EA, van der Werf MJ, van Os J, Vavilala MS, Venketasubramanian N, Wang M, Wang W, Watt K, Weatherall DJ, Weinstock MA, Weintraub R, Weisskopf MG, Weissman MM, White RA, Whiteford H, Wiebe N, Wiersma ST, Wilkinson JD, Williams HC, Williams SRM, Witt E, Wolfe F, Woolf AD, Wulf S, Yeh P-H, Zaidi AKM, Zheng Z-J, Zonies D, Lopez AD, AlMazroa MA, Memish ZA. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;380(9859):2197-223.
4. Yusuf S, Rangarajan S, Teo K, Islam S, Li W, Liu L, Bo J, Lou Q, Lu F, Liu T, Yu L, Zhang S, Mony P, Swaminathan S, Mohan V, Gupta R, Kumar R, Vijayakumar K, Lear S, Anand S, Wielgosz A, Diaz R, Avezum A, Lopez-Jaramillo P, Lanas F, Yusoff K, Ismail N, Iqbal R, Rahman O, Rosengren A, Yusufali A, Kelishadi R, Kruger A, Puoane T, Szuba A, Chifamba J, Oguz A, McQueen M, McKee M, Dagenais G. Cardiovascular risk and events in 17 low-, middle-, and high-income countries. *New England Journal of Medicine* 2014;371(9):818-27.
 5. Institute for Health Metrics and Evaluation. GBD Database Seattle: Institute for Health Metrics and Evaluation; 2014 [cited 2015 29th April 2015]. Available from: <http://www.healthdata.org/search-gbd-data?s=China>.
 6. Emerging Risk Factors Collaboration. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. *Lancet* 2010;375(9733):2215-22.
 7. Chen Z, Lee L, Chen J, Collins R, Wu F, Guo Y, Linksted P, Peto R. Cohort profile: the Kadoorie Study of Chronic Disease in China (KSCDC). *International Journal of Epidemiology* 2005;34(6):1243-9.
 8. Wong ND. Epidemiological studies of CHD and the evolution of preventive cardiology. *Nature Reviews Cardiology* 2014;11(5):276-89.
 9. Mathers C, Truelsen T, Begg S, Satoh T. Global burden of ischaemic heart disease in the year 2000. Geneva: World Health Organization 2004.
 10. Lefkowitz RJ, Willerson JT. Prospects for cardiovascular research. *JAMA- The Journal of the American Medical Association* 2001;285(5):581-7.
 11. Tegos TJ, Kalodiki E, Sabetai MM, Nicolaides AN. The genesis of atherosclerosis and risk factors: a review. *Angiology* 2001;52(2):89-98.
 12. Dollery C, Libby P. Biology and pathology of atherosclerosis. In: Warrell DA, Cox TM, Firth JD, editors. *Oxford Textbook of Medicine*. 5th ed. Oxford: Oxford University Press; 2010.

13. Libby P, Ridker PM, Hansson GK. Progress and challenges in translating the biology of atherosclerosis. *Nature* 2011;473(7347):317-25.
14. Kullo IJ, Ding K. Mechanisms of Disease: the genetic basis of coronary heart disease. *Nature Clinical Practice Cardiovascular Medicine* 2007;4(10):558-69.
15. Finegold JA, Asaria P, Francis DP. Mortality from ischaemic heart disease by country, region, and age: statistics from World Health Organisation and United Nations. *International Journal of Cardiology* 2013;168(2):934-45.
16. Lewington S, Clarke R, Qizilbash N, Peto R, Collins R. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* 2002;360(9349):1903-13.
17. Prospective Studies Collaboration, Lewington S, Whitlock G, Clarke R, Sherliker P, Emberson J, Halsey J, Qizilbash N, Peto R, Collins R. Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. *Lancet* 2007;370(9602):1829-39.
18. Emerging Risk Factors Collaboration, Di Angelantonio E, Sarwar N, Perry P, Kaptoge S, Ray KK, Thompson A, Wood AM, Lewington S, Sattar N, Packard CJ, Collins R, Thompson SG, Danesh J. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA- The Journal of the American Medical Association* 2009;302(18):1993-2000.
19. Cholesterol Treatment Trialists' Collaboration. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170 000 participants in 26 randomised trials. *Lancet* 2010;376(9753):1670-81.
20. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ* 2004;328(7455):1519.
21. Ronksley PE, Brien SE, Turner BJ, Mukamal KJ, Ghali WA. Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis. *BMJ* 2011;342:d671.
22. Gulsvik AK, Thelle DS, Samuelsen SO, Myrstad M, Mowe M, Wyller TB. Ageing, physical activity and mortality- a 42-year follow-up study. *International Journal of Epidemiology* 2012;41(2):521-30.
23. Berlin JA, Colditz GA. A meta-analysis of physical activity in the prevention of coronary heart disease. *American Journal of Epidemiology* 1990;132(4):612-28.
24. Whitlock G, Lewington S, Sherliker P, Clarke R, Emberson J, Halsey J, Qizilbash N, Collins R, Peto R. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. *Lancet* 2009;373(9669):1083-96.

25. World Health Organization, World Heart Federation, World Stroke Organization. Global Atlas on cardiovascular disease prevention and control. Geneva: World Health Organization, 2011.

26. Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buysschaert I, Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NE, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S, Gudnason V, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJ, Khaw KT, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Muhleisen TW, Muhlestein JB, Munzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nothen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS, Schadt E, Schafer A, Schillert A, Schreiber S, Schrezenmeir J, Schwartz SM, Siscovick DS, Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoop JD, Soranzo N, Spertus JA, Stark K, Stirrups K, Stoll M, Tang WH, Tennstedt S, Thorgeirsson G, Thorleifsson G, Tomaszewski M, Uitterlinden AG, van Rijn AM, Voight BF, Wareham NJ, Wells GA, Wichmann HE, Wild PS, Willenborg C, Witterman JC, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA, Quertermous T, März W, Hengstenberg C, Blankenberg S, Ouwehand WH, Hall AS, Deloukas P, Thompson JR, Stefansson K, Roberts R, Thorsteinsdottir U, O'Donnell CJ, McPherson R, Erdmann J, Samani NJ. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nature Genetics* 2011;43(4):333-8.

27. Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nature Genetics* 2011;43(4):339-44.

28. Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, Ingelsson E, Saleheen D, Erdmann J, Goldstein BA, Stirrups K, König IR, Cazier JB, Johansson A, Hall AS, Lee JY, Willer CJ, Chambers JC, Esko T, Folkersen L, Goel A, Grundberg E, Havulinna AS, Ho WK, Hopewell JC, Eriksson N, Kleber ME, Kristiansson K, Lundmark P, Lyytikäinen LP, Rafelt S, Shungin D, Strawbridge RJ, Thorleifsson G, Tikkanen E, Van Zuydam N, Voight BF, Waite LL, Zhang W, Ziegler A, Absher D, Altshuler D, Balmforth AJ, Barroso I, Braund PS, Burgdorf C, Claudi-Boehm S, Cox D, Dimitriou M, Do R, Doney AS, El Mokhtari N, Eriksson P, Fischer K, Fontanillas P, Franco-Cereceda A, Gigante B, Groop L, Gustafsson S, Hager J, Hallmans G, Han BG, Hunt SE, Kang HM, Illig T, Kessler T, Knowles JW, Kolovou G, Kuusisto J, Langenberg C, Langford C, Leander K, Lokki ML, Lundmark A, McCarthy MI, Meisinger C, Melander O, Mihailov E, Maouche S, Morris AD, Müller-Nurasyid M, Nikus K, Peden JF, Rayner NW, Rasheed A, Rosinger S, Rubin D, Rumpf MP, Schafer A, Sivananthan M, Song C, Stewart AF, Tan ST, Thorgeirsson G, van der Schoot CE, Wagner PJ, Wells GA, Wild PS, Yang TP, Amouyel P, Arveiler D, Basart H, Boehnke M, Boerwinkle E, Brambilla P, Cambien F, Cupples AL, de Faire U, Dehghan A, Diemert P, Epstein SE, Evans A, Ferrario MM, Ferrieres

- J, Gauguier D, Go AS, Goodall AH, Gudnason V, Hazen SL, Holm H, Iribarren C, Jang Y, Kahonen M, Kee F, Kim HS, Klopp N, Koenig W, Kratzer W, Kuulasmaa K, Laakso M, Laaksonen R, Lind L, Ouwehand WH, Parish S, Park JE, Pedersen NL, Peters A, Quertermous T, Rader DJ, Salomaa V, Schadt E, Shah SH, Sinisalo J, Stark K, Stefansson K, Tregouet DA, Virtamo J, Wallentin L, Wareham N, Zimmermann ME, Nieminen MS, Hengstenberg C, Sandhu MS, Pastinen T, Syvanen AC, Hovingh GK, Dedoussis G, Franks PW, Lehtimäki T, Metspalu A, Zalloua PA, Siegbahn A, Schreiber S, Ripatti S, Blankenberg SS, Perola M, Clarke R, Boehm BO, O'Donnell C, Reilly MP, März W, Collins R, Kathiresan S, Hamsten A, Kooner JS, Thorsteinsdóttir U, Danesh J, Palmer CN, Roberts R, Watkins H, Schunkert H, Samani NJ. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nature Genetics* 2013;45(1):25-33.
29. Lieb W, Vasan RS. Genetics of Coronary Artery Disease. *Circulation* 2013;128(10):1131-8.
 30. Timmis AD. Management of stable angina. In: Warrell DA, Cox TM, Firth JD, editors. *Oxford Textbook of Medicine*. 5th ed. Oxford: Oxford University Press; 2010.
 31. Fox K, Mills N. Management of acute coronary syndrome. In: Warrell D, Cox T, Firth J, editors. *Oxford Textbook of Medicine*. 5th ed. Oxford: Oxford University Press; 2010.
 32. Hatano S. Experience from a multicentre stroke register: a preliminary report. *Bulletin of the World Health Organization* 1976;54(5):541-53.
 33. Muir KW. Stroke. *Medicine* 2009;37(2):109-14.
 34. Endres M, Heuschmann PU, Laufs U, Hakim AM. Primary prevention of stroke: blood pressure, lipids, and heart failure. *European Heart Journal* 2011;32(5):545-52.
 35. Donnan GA, Fisher M, Macleod M, Davis SM. Stroke. *Lancet* 2008;371(9624):1612-23.
 36. Bamford J, Sandercock P, Dennis M, Burn J, Warlow C. Classification and natural history of clinically identifiable subtypes of cerebral infarction. *Lancet* 1991;337(8756):1521-6.
 37. Adams HP, Jr., Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, Marsh EE, 3rd. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke* 1993;24(1):35-41.
 38. van Gijn J. Stroke: cerebrovascular disease. In: Warrell DA, Cox TM, Firth JD, editors. *Oxford Textbook of Medicine*. 5th ed. Oxford: Oxford University Press; 2010.
 39. Saenger AK, Christenson RH. Stroke biomarkers: progress and challenges for diagnosis, prognosis, differentiation, and treatment. *Clinical Chemistry* 2010;56(1):21-33.

40. Lee CD, Folsom AR, Blair SN. Physical activity and stroke risk: a meta-analysis. *Stroke* 2003;34(10):2475-81.
41. Kurth T, Kase CS, Berger K, Schaeffner ES, Buring JE, Gaziano JM. Smoking and the risk of hemorrhagic stroke in men. *Stroke* 2003;34(5):1151-5.
42. Reynolds K, Lewis B, Nolen JD, Kinney GL, Sathya B, He J. Alcohol consumption and risk of stroke: a meta-analysis. *JAMA- The Journal of the American Medical Association* 2003;289(5):579-88.
43. Sutherland GR, Auer RN. Primary intracerebral hemorrhage. *Journal of Clinical Neuroscience* 2006;13(5):511-7.
44. Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, Fornage M, Ikram MA, Malik R, Bevan S, Thorsteinsdottir U, Nalls MA, Longstreth W, Wiggins KL, Yadav S, Parati EA, Destefano AL, Worrall BB, Kittner SJ, Khan MS, Reiner AP, Helgadottir A, Achterberg S, Fernandez-Cadenas I, Abboud S, Schmidt R, Walters M, Chen WM, Ringelstein EB, O'Donnell M, Ho WK, Pera J, Lemmens R, Norrving B, Higgins P, Benn M, Sale M, Kuhlenbaumer G, Doney AS, Vicente AM, Delavaran H, Algra A, Davies G, Oliveira SA, Palmer CN, Deary I, Schmidt H, Pandolfo M, Montaner J, Carty C, de Bakker PI, Kostulas K, Ferro JM, van Zuydam NR, Valdimarsson E, Nordestgaard BG, Lindgren A, Thijs V, Slowik A, Saleheen D, Pare G, Berger K, Thorleifsson G, Hofman A, Mosley TH, Mitchell BD, Furie K, Clarke R, Levi C, Seshadri S, Gschwendtner A, Boncoraglio GB, Sharma P, Bis JC, Gretarsdottir S, Psaty BM, Rothwell PM, Rosand J, Meschia JF, Stefansson K, Dichgans M, Markus HS. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurology* 2012;11(11):951-62.
45. Bellenguez C, Bevan S, Gschwendtner A, Spencer CCA, Burgess AI, Pirinen M, Jackson CA, Traylor M, Strange A, Su Z, Band G, Syme PD, Malik R, Pera J, Norrving B, Lemmens R, Freeman C, Schanz R, James T, Poole D, Murphy L, Segal H, Cortellini L, Cheng Y-C, Woo D, Nalls MA, Muller-Myhsok B, Meisinger C, Seedorf U, Ross-Adams H, Boonen S, Wloch-Kopec D, Valant V, Slark J, Furie K, Delavaran H, Langford C, Deloukas P, Edkins S, Hunt S, Gray E, Dronov S, Peltonen L, Gretarsdottir S, Thorleifsson G, Thorsteinsdottir U, Stefansson K, Boncoraglio GB, Parati EA, Attia J, Holliday E, Levi C, Franzosi M-G, Goel A, Helgadottir A, Blackwell JM, Bramon E, Brown MA, Casas JP, Corvin A, Duncanson A, Jankowski J, Mathew CG, Palmer CNA, Plomin R, Rautanen A, Sawcer SJ, Trembath RC, Viswanathan AC, Wood NW, Worrall BB, Kittner SJ, Mitchell BD, Kissela B, Meschia JF, Thijs V, Lindgren A, Macleod MJ, Slowik A, Walters M, Rosand J, Sharma P, Farrall M, Sudlow CLM, Rothwell PM, Dichgans M, Donnelly P, Markus HS. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nature Genetics* 2012;44(3):328-33.
46. Woo D, Falcone GJ, Devan WJ, Brown WM, Biffi A, Howard TD, Anderson CD, Brouwers HB, Valant V, Battey TW, Radmanesh F, Raffeld MR, Baedorf-Kassis S, Deka R, Woo JG, Martin LJ, Haverbusch M, Moomaw CJ, Sun G, Broderick JP, Flaherty ML, Martini SR, Kleindorfer DO, Kissela B, Comeau ME, Jagiella JM, Schmidt H, Freudenberger P, Pichler A, Enzinger C, Hansen BM, Norrving B, Jimenez-Conde J, Giralt-Steinhauer E, Elosua R, Cuadrado-Godia E, Soriano C, Roquer J, Kraft P, Ayres AM, Schwab K, McCauley JL, Pera J, Urbanik A, Rost NS,

- Goldstein JN, Viswanathan A, Stoger EM, Tirschwell DL, Selim M, Brown DL, Silliman SL, Worrall BB, Meschia JF, Kidwell CS, Montaner J, Fernandez-Cadenas I, Delgado P, Malik R, Dichgans M, Greenberg SM, Rothwell PM, Lindgren A, Slowik A, Schmidt R, Langefeld CD, Rosand J. Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. *American Journal of Human Genetics* 2014;94(4):511-21.
47. Keir SL, Wardlaw JM, Warlow CP. Stroke epidemiology studies have underestimated the frequency of intracerebral haemorrhage. A systematic review of imaging in epidemiological studies. *Journal of Neurology* 2002;249(9):1226-31.
 48. Krishnamurthi RV, Feigin VL, Forouzanfar MH, Mensah GA, Connor M, Bennett DA, Moran AE, Sacco RL, Anderson LM, Truelsen T, O'Donnell M, Venketasubramanian N, Barker-Collo S, Lawes CM, Wang W, Shinohara Y, Witt E, Ezzati M, Naghavi M, Murray C. Global and regional burden of first-ever ischaemic and haemorrhagic stroke during 1990-2010: findings from the Global Burden of Disease Study 2010. *Lancet Global Health* 2013;1(5):e259-81.
 49. Moran AE, Forouzanfar MH, Roth GA, Mensah GA, Ezzati M, Flaxman A, Murray CJL, Naghavi M. The Global Burden of Ischemic Heart Disease in 1990 and 2010: The Global Burden of Disease 2010 Study. *Circulation* 2014;129(14):1493-501.
 50. Moran AE, Forouzanfar MH, Roth GA, Mensah GA, Ezzati M, Murray CJ, Naghavi M. Temporal trends in ischemic heart disease mortality in 21 world regions, 1980 to 2010: the Global Burden of Disease 2010 study. *Circulation* 2014;129(14):1483-92.
 51. Moran AE, Tzong KY, Forouzanfar MH, Roth G, Mensah GA, Ezzati M, Murray CJL, Naghavi M. Variations in Ischemic Heart Disease Burden by Age, Country, and Income: The Global Burden of Diseases, Injuries and Risk Factors 2010 Study. *Global Heart* 2014;9(1):91-9.
 52. Omran AR. The Epidemiologic Transition: A Theory of the Epidemiology of Population Change. *The Milbank Quarterly* 2005;83(4):731-57.
 53. Yusuf S, Reddy S, Ounpuu S, Anand S. Global burden of cardiovascular diseases: part I: general considerations, the epidemiologic transition, risk factors, and impact of urbanization. *Circulation* 2001;104(22):2746-53.
 54. Zhang XH, Lu ZL, Liu L. Coronary heart disease in China. *Heart* 2008;94(9):1126-31.
 55. Liu M, Wu B, Wang W-Z, Lee L-M, Zhang S-H, Kong L-Z. Stroke in China: epidemiology, prevention, and management strategies. *Lancet Neurology* 2007;6(5):456-64.
 56. Huang CY, Chan FL, Yu YL, Woo E, Chin D. Cerebrovascular disease in Hong Kong Chinese. *Stroke* 1990;21(2):230-5.
 57. Chen Z, Whitlock G. Unveiling the causes of heart disease in China. *Heart* 2009;95(22):1818-9.

58. Wu Z, Yao C, Zhao D, Wu G, Wang W, Liu J, Zeng Z, Wu Y. Sino-MONICA project: a collaborative study on trends and determinants in cardiovascular diseases in China, Part i: morbidity and mortality monitoring. *Circulation* 2001;103(3):462-8.
59. Yusuf S, Reddy S, Ôunpuu S, Anand S. Global Burden of Cardiovascular Diseases: Part II: Variations in Cardiovascular Disease by Specific Ethnic Groups and Geographic Regions and Prevention Strategies. *Circulation* 2001;104(23):2855-64.
60. Moran A, Gu D, Zhao D, Coxson P, Wang YC, Chen C-S, Liu J, Cheng J, Bibbins-Domingo K, Shen Y-M, He J, Goldman L. Future Cardiovascular Disease in China: Markov Model and Risk Factor Scenario Projections From the Coronary Heart Disease Policy Model–China. *Circulation: Cardiovascular Quality and Outcomes* 2010;3(3):243-52.
61. Dayan C, Williams G. Diabetes. In: Warrell DA, Cox TM, Firth JD, editors. *Oxford Textbook of Medicine*. 5th ed. Oxford: Oxford University Press; 2010.
62. Hall J. Insulin, Glucagon and Diabetes Mellitus. *Guyton and Hall Textbook of Medical Physiology*. 13th ed. Philadelphia: Elsevier; 2015.
63. Pocock G, Richards C, Richards D. The endocrine pancreas and the regulation of plasma glucose. *Human Physiology*. 4th ed. Oxford: Oxford University Press; 2013.
64. American Diabetes Association. Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* 2014;37(Supplement 1):S81-S90.
65. International Diabetes Federation. *IDF Diabetes Atlas*. Brussels: International Diabetes Federation, 2013.
66. Chen L, Magliano DJ, Zimmet PZ. The worldwide epidemiology of type 2 diabetes mellitus-present and future perspectives. *Nature Reviews Endocrinology* 2012;8(4):228-36.
67. Zimmet PZ, Magliano DJ, Herman WH, Shaw JE. Diabetes: a 21st century challenge. *Lancet Diabetes and Endocrinology* 2014;2(1):56-64.
68. Hu FB. Globalization of diabetes: the role of diet, lifestyle, and genes. *Diabetes Care* 2011;34(6):1249-57.
69. Kharroubi AT, Darwish HM. Diabetes mellitus: The epidemic of the century. *World Journal of Diabetes* 2015;6(6):850-67.
70. The DIAMOND Project Group. Incidence and trends of childhood Type 1 diabetes worldwide 1990-1999. *Diabetic Medicine* 2006;23(8):857-66.
71. Mohlke KL, Boehnke M. Recent advances in understanding the genetic architecture of type 2 diabetes. *Human Molecular Genetics* 2015:doi:10.1093/hmg/ddv264.

72. Basile KJ, Guy VC, Schwartz S, Grant SF. Overlap of genetic susceptibility to type 1 diabetes, type 2 diabetes, and latent autoimmune diabetes in adults. *Current Diabetes Reports* 2014;14(11):550.
73. Hu FB, Manson JE, Stampfer MJ, Colditz G, Liu S, Solomon CG, Willett WC. Diet, lifestyle, and the risk of type 2 diabetes mellitus in women. *New England Journal of Medicine* 2001;345(11):790-7.
74. Langenberg C, Sharp SJ, Schulze MB, Rolandsson O, Overvad K, Forouhi NG, Spranger J, Drogan D, Huerta JM, Arriola L, de Lauzon-Guillan B, Tormo MJ, Ardanaz E, Balkau B, Beulens JW, Boeing H, Bueno-de-Mesquita HB, Clavel-Chapelon F, Crowe FL, Franks PW, Gonzalez CA, Grioni S, Halkjaer J, Hallmans G, Kaaks R, Kerrison ND, Key TJ, Khaw KT, Mattiello A, Nilsson P, Norat T, Palla L, Palli D, Panico S, Quiros JR, Romaguera D, Romieu I, Sacerdote C, Sanchez MJ, Slimani N, Sluijs I, Spijkerman AM, Teucher B, Tjonneland A, Tumino R, van der AD, van der Schouw YT, Feskens EJ, Riboli E, Wareham NJ. Long-term risk of incident type 2 diabetes and measures of overall and regional obesity: the EPIC-InterAct case-cohort study. *PLoS Medicine* 2012;9(6):e1001230.
75. Jeon CY, Lokken RP, Hu FB, van Dam RM. Physical activity of moderate intensity and risk of type 2 diabetes: a systematic review. *Diabetes Care* 2007;30(3):744-52.
76. Willi C, Bodenmann P, Ghali WA, Faris PD, Cornuz J. Active smoking and the risk of type 2 diabetes: A systematic review and meta-analysis. *JAMA- The Journal of the American Medical Association* 2007;298(22):2654-64.
77. National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes* 1979;28(12):1039-57.
78. World Health Organization, International Diabetes Federation. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation. Geneva: World Health Organization, 2006.
79. Danaei G, Finucane MM, Lu Y, Singh GM, Cowan MJ, Paciorek CJ, Lin JK, Farzadfar F, Khang Y-H, Stevens GA, Rao M, Ali MK, Riley LM, Robinson CA, Ezzati M, Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. *Lancet* 2011;378(9785):31-40.
80. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL, Lindgren CM, Magi R, Morris AP, Randall J, Johnson T, Elliott P, Rybin D, Thorleifsson G, Steinthorsdottir V, Henneman P, Grallert H, Dehghan A, Hottenga JJ, Franklin CS, Navarro P, Song K, Goel A, Perry JRB, Egan JM, Lajunen T, Grarup N, Sparso T, Doney A, Voight BF, Stringham HM, Li M, Kanoni S, Shrader P, Cavalcanti-Proenca C, Kumari M, Qi L, Timpson NJ, Gieger C, Zabena C, Rocheleau G, Ingelsson E, An P, O'Connell J, Luan Ja, Elliott A, McCarroll SA, Payne F, Roccascaccia RM, Pattou F, Sethupathy P, Ardlie K, Ariyurek Y, Balkau B, Barter P, Beilby JP, Ben-Shlomo Y, Benediktsson R, Bennett AJ, Bergmann S, Bochud M, Boerwinkle E, Bonnefond A, Bonnycastle

LL, Borch-Johnsen K, Bottcher Y, Brunner E, Bumpstead SJ, Charpentier G, Chen Y-DI, Chines P, Clarke R, Coin LJM, Cooper MN, Cornelis M, Crawford G, Crisponi L, Day INM, de Geus EJC, Delplanque J, Dina C, Erdos MR, Fedson AC, Fischer-Rosinsky A, Forouhi NG, Fox CS, Frants R, Franzosi MG, Galan P, Goodarzi MO, Graessler J, Groves CJ, Grundy S, Gwilliam R, Gyllensten U, Hadjadj S, Hallmans G, Hammond N, Han X, Hartikainen A-L, Hassanali N, Hayward C, Heath SC, Hercberg S, Herder C, Hicks AA, Hillman DR, Hingorani AD, Hofman A, Hui J, Hung J, Isomaa B, Johnson PRV, Jorgensen T, Julia A, Kaakinen M, Kaprio J, Kesaniemi YA, Kivimaki M, Knight B, Koskinen S, Kovacs P, Kyvik KO, Lathrop GM, Lawlor DA, Le Bacquer O, Lecoeur C, Li Y, Lyssenko V, Mahley R, Mangino M, Manning AK, Martinez-Larrad MT, McAteer JB, McCulloch LJ, McPherson R, Meisinger C, Melzer D, Meyre D, Mitchell BD, Morken MA, Mukherjee S, Naitza S, Narisu N, Neville MJ, Oostra BA, Orru M, Pakyz R, Palmer CNA, Paolisso G, Pattaro C, Pearson D, Peden JF, Pedersen NL, Perola M, Pfeiffer AFH, Pichler I, Polasek O, Posthuma D, Potter SC, Pouta A, Province MA, Psaty BM, Rathmann W, Rayner NW, Rice K, Ripatti S, Rivadeneira F, Roden M, Rolandsson O, Sandbaek A, Sandhu M, Sanna S, Sayer AA, Scheet P, Scott LJ, Seedorf U, Sharp SJ, Shields B, Sigurdsson G, Sijbrands EJG, Silveira A, Simpson L, Singleton A, Smith NL, Sovio U, Swift A, Syddall H, Syvanen A-C, Tanaka T, Thorand B, Tichet J, Tonjes A, Tuomi T, Uitterlinden AG, van Dijk KW, van Hoek M, Varma D, Visvikis-Siest S, Vitart V, Vogelzangs N, Waeber G, Wagner PJ, Walley A, Walters GB, Ward KL, Watkins H, Weedon MN, Wild SH, Willemssen G, Witteman JCM, Yarnell JWG, Zeggini E, Zelenika D, Zethelius B, Zhai G, Zhao JH, Zillikens MC, Diagram Consortium, Giant Consortium, Global BPgen Consortium, Borecki IB, Loos RJF, Meneton P, Magnusson PKE, Nathan DM, Williams GH, Hattersley AT, Silander K, Salomaa V, Smith GD, Bornstein SR, Schwarz P, Spranger J, Karpe F, Shuldiner AR, Cooper C, Dedoussis GV, Serrano-Rios M, Morris AD, Lind L, Palmer LJ, Hu FB, Franks PW, Ebrahim S, Marmot M, Kao WHL, Pankow JS, Sampson MJ, Kuusisto J, Laakso M, Hansen T, Pedersen O, Pramstaller PP, Wichmann HE, Illig T, Rudan I, Wright AF, Stumvoll M, Campbell H, Wilson JF, Hamsten A on behalf of Procardis Consortium, Magic Investigators, Bergman RN, Buchanan TA, Collins FS, Mohlke KL, Tuomilehto J, Valle TT, Altshuler D, Rotter JI, Siscovick DS, Penninx BWJH, Boomsma DI, Deloukas P, Spector TD, Frayling TM, Ferrucci L, Kong A, Thorsteinsdottir U, Stefansson K, van Duijn CM, Aulchenko YS, Cao A, Scuteri A, Schlessinger D, Uda M, Ruukonen A, Jarvelin M-R, Waterworth DM, Vollenweider P, Peltonen L, Mooser V, Abecasis GR, Wareham NJ, Sladek R, Froguel P, Watanabe RM, Meigs JB, Groop L, Boehnke M, McCarthy MI, Florez JC, Barroso I. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nature Genetics* 2010;42(2):105-16.

81. Scott RA, Lagou V, Welch RP, Wheeler E, Montasser ME, Luan Ja, Magi R, Strawbridge RJ, Rehnberg E, Gustafsson S, Kanoni S, Rasmussen-Torvik LJ, Yengo L, Lecoeur C, Shungin D, Sanna S, Sidore C, Johnson PCD, Jukema JW, Johnson T, Mahajan A, Verweij N, Thorleifsson G, Hottenga J-J, Shah S, Smith AV, Sennblad B, Gieger C, Salo P, Perola M, Timpson NJ, Evans DM, Pourcain BS, Wu Y, Andrews JS, Hui J, Bielak LF, Zhao W, Horikoshi M, Navarro P, Isaacs A, O'Connell JR, Stirrups K, Vitart V, Hayward C, Esko T, Mihailov E, Fraser RM, Fall T, Voight BF, Raychaudhuri S, Chen H, Lindgren CM, Morris AP, Rayner NW, Robertson N, Rybin D, Liu C-T, Beckmann JS, Willems SM, Chines PS, Jackson AU, Kang HM, Stringham HM, Song K, Tanaka T, Peden JF, Goel A, Hicks AA, An P, Muller-Nurasyid M, Franco-Cereceda A, Folkersen L, Marullo L, Jansen H, Oldehinkel AJ, Bruinenberg M, Pankow JS, North KE, Forouhi NG, Loos RJF, Edkins S, Varga TV, Hallmans G, Oksa H, Antonella M, Nagaraja R, Trompet S, Ford I, Bakker SJL, Kong A, Kumari M, Gigante B, Herder C, Munroe PB, Caulfield

- M, Antti J, Mangino M, Small K, Miljkovic I, Liu Y, Atalay M, Kiess W, James AL, Rivadeneira F, Uitterlinden AG, Palmer CNA, Doney ASF, Willemssen G, Smit JH, Campbell S, Polasek O, Bonnycastle LL, Hercberg S, Dimitriou M, Bolton JL, Fowkes GR, Kovacs P, Lindstrom J, Zemunik T, Bandinelli S, Wild SH, Basart HV, Rathmann W, Grallert H, DIABetes Genetics Replication Meta-analysis Consortium, Maerz W, Kleber ME, Boehm BO, Peters A, Pramstaller PP, Province MA, Borecki IB, Hastie ND, Rudan I, Campbell H, Watkins H, Farrall M, Stumvoll M, Ferrucci L, Waterworth DM, Bergman RN, Collins FS, Tuomilehto J, Watanabe RM, de Geus EJC, Penninx BW, Hofman A, Oostra BA, Psaty BM, Vollenweider P, Wilson JF, Wright AF, Hovingh GK, Metspalu A, Uusitupa M, Magnusson PKE, Kyvik KO, Kaprio J, Price JF, Dedoussis GV, Deloukas P, Meneton P, Lind L, Boehnke M, Shuldiner AR, van Duijn CM, Morris AD, Toenjes A, Peyser PA, Beilby JP, Korner A, Kuusisto J, Laakso M, Bornstein SR, Schwarz PEH, Lakka TA, Rauramaa R, Adair LS, Smith GD, Spector TD, Illig T, de Faire U, Hamsten A, Gudnason V, Kivimaki M, Hingorani A, Keinanen-Kiukkaanniemi SM, Saaristo TE, Boomsma DI, Stefansson K, van der Harst P, Dupuis J, Pedersen NL, Sattar N, Harris TB, Cucca F, Ripatti S, Salomaa V, Mohlke KL, Balkau B, Froguel P, Pouta A, Jarvelin M-R, Wareham NJ, Bouatia-Naji N, McCarthy MI, Franks PW, Meigs JB, Teslovich TM, Florez JC, Langenberg C, Ingelsson E, Prokopenko I, Barroso I. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nature Genetics* 2012;44(9):991-1005.
82. Sung J, Song Y-M, Ebrahim S, Lawlor DA. Fasting blood glucose and the risk of stroke and myocardial infarction. *Circulation* 2009;119(6):812-9.
 83. Hyvarinen M, Qiao Q, Tuomilehto J, Laatikainen T, Heine RJ, Stehouwer CDA, Alberti KGMM, Pyorala K, Zethelius B, Stegmayr B, DECODE Study Group. Hyperglycemia and stroke mortality: comparison between fasting and 2-h glucose criteria. *Diabetes Care* 2009;32(2):348-54.
 84. Orenchia AJ, Daviglus ML, Dyer AR, Walsh M, Greenland P, Stamler J. One-hour postload plasma glucose and risks of fatal coronary heart disease and stroke among nondiabetic men and women: the Chicago Heart Association Detection Project in Industry (CHA) Study. *Journal of Clinical Epidemiology* 1997;50(12):1369-76.
 85. National Collaborative Group of Diabetes Study. [A mass survey of diabetes mellitus in a population of 300,000 in 14 provinces and municipalities in China]. *Zhonghua Nei Ke Za Zhi* 1981;20(11):678-83.
 86. Chan JC, Zhang Y, Ning G. Diabetes in China: a societal solution for a personal challenge. *Lancet Diabetes and Endocrinology* 2014;2(12):969-79.
 87. Xu Y, Wang L, He J, Bi Y, Li M, Wang T, Wang L, Jiang Y, Dai M, Lu J, Xu M, Li Y, Hu N, Li J, Mi S, Chen C-S, Li G, Mu Y, Zhao J, Kong L, Chen J, Lai S, Wang W, Zhao W, Ning G, China Noncommunicable Disease Surveillance Group. Prevalence and control of diabetes in Chinese adults. *JAMA- The Journal of the American Medical Association* 2013;310(9):948-59.
 88. Kong AP, Xu G, Brown N, So WY, Ma RC, Chan JC. Diabetes and its comorbidities-where East meets West. *Nature Reviews Endocrinology* 2013;9(9):537-47.

89. Ma RC, Lin X, Jia W. Causes of type 2 diabetes in China. *Lancet Diabetes and Endocrinology* 2014;2(12):980-91.
90. Shah AD, Langenberg C, Rapsomaniki E, Denaxas S, Pujades-Rodriguez M, Gale CP, Deanfield J, Smeeth L, Timmis A, Hemingway H. Type 2 diabetes and incidence of cardiovascular diseases: a cohort study in 1·9 million people. *Lancet Diabetes and Endocrinology* 2015;3(2):105-13.
91. Asia Pacific Cohort Studies Collaboration. The effects of diabetes on the risks of major cardiovascular diseases and death in the Asia-Pacific region. *Diabetes Care* 2003;26(2):360-6.
92. Reusch JEB, Wang CCL. Cardiovascular disease in diabetes: where does glucose fit in? *Journal of Clinical Endocrinology and Metabolism* 2011;96(8):2367-76.
93. Bianchi C, Miccoli R, Penno G, Del Prato S. Primary prevention of cardiovascular disease in people with dysglycemia. *Diabetes Care* 2008;31 Suppl 2:S208-14.
94. Duckworth WC. Hyperglycemia and cardiovascular disease. *Current Atherosclerosis Reports* 2001;3(5):383-91.
95. Szuszkiewicz-Garcia MM, Davidson JA. Cardiovascular Disease in Diabetes Mellitus: Risk Factors and Medical Therapy. *Endocrinology and Metabolism Clinics of North America* 2014;43(1):25-40.
96. Kovacic JC, Castellano JM, Farkouh ME, Fuster V. The relationships between cardiovascular disease and diabetes: focus on pathogenesis. *Endocrinology and Metabolism Clinics of North America* 2014;43(1):41-57.
97. Punthakee Z, Werstuck GH, Gerstein HC. Diabetes and cardiovascular disease: explaining the relationship. *Reviews in Cardiovascular Medicine* 2007;8(3):145-53.
98. Gerstein HC, Werstuck GH. Dysglycaemia, vasculopenia, and the chronic consequences of diabetes. *Lancet Diabetes and Endocrinology* 2013;1(1):71-8.
99. Quinn TJ, Dawson J, Walters MR. Sugar and stroke: cerebrovascular disease and blood glucose control. *Cardiovascular therapeutics* 2011;29(6):e31-42.
100. Gerstein HC, Yusuf S. Dysglycaemia and risk of cardiovascular disease. *Lancet* 1996;347(9006):949-50.
101. Gerstein HC. Glucose: a continuous risk factor for cardiovascular disease. *Diabetic Medicine* 1997;14 Suppl 3:S25-31.
102. Park C, Guallar E, Linton JA, Lee D-C, Jang Y, Son DK, Han E-J, Baek SJ, Yun YD, Jee SH, Samet JM. Fasting glucose level and the risk of incident atherosclerotic cardiovascular diseases. *Diabetes Care* 2013;36(7):1988-93.

103. Lawes CMM, Parag V, Bennett DA, Suh I, Lam TH, Whitlock G, Barzi F, Woodward M, Asia Pacific Cohort Studies Collaboration. Blood glucose and risk of cardiovascular disease in the Asia Pacific region. *Diabetes Care* 2004;27(12):2836-42.
104. Einarson TR, Machado M, Henk Hemels ME. Blood glucose and subsequent cardiovascular disease: update of a meta-analysis. *Current Medical Research and Opinion* 2011;27(11):2155-63.
105. Sarwar N, Aspelund T, Eiriksdottir G, Gobin R, Seshasai SRK, Forouhi NG, Sigurdsson G, Danesh J, Gudnason V. Markers of dysglycaemia and risk of coronary heart disease in people without diabetes: Reykjavik prospective study and systematic review. *PLoS Medicine* 2010;7(5):e1000278.
106. Sattar N. Revisiting the links between glycaemia, diabetes and cardiovascular disease. *Diabetologia* 2013;56(4):686-95.
107. World Health Organization. Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus. Abbreviated Report of a WHO Consultation. Geneva: World Health Organization, 2011.
108. Tang X, He L, Cao Y, Wang J, Li N, Tian J, Liu J, Yu L, Xu H, Liu W, Wen S, Dou H, Hu Y. [Gender-specific differences in relative effects of cardiovascular risk factors among rural population]. *Beijing da xue xue bao Yi xue ban = Journal of Peking University Health sciences* 2011;43(3):379-85.
109. Liu LH, Chia LG. [The effects of hypertension, diabetes mellitus, atrial fibrillation, transient ischemic attack and smoking on stroke in Chinese people]. *Zhonghua Yi Xue Za Zhi (Taipei)* 1991;47(2):110-5.
110. He L, Tang X, Song Y, Li N, Li J, Zhang Z, Liu J, Yu L, Xu H, Zhang J, Hu Y. Prevalence of cardiovascular disease and risk factors in a rural district of Beijing, China: a population-based survey of 58,308 residents. *BMC Public Health* 2012;12:34.
111. Wirehn A-BE, Ostgren CJ, Carstensen JM. Age and gender differences in the impact of diabetes on the prevalence of ischemic heart disease: a population-based register study. *Diabetes Research and Clinical Practice* 2008;79(3):497-502.
112. Hammoudeh AJ, Al-Tarawneh H, Elharassis A, Haddad J, Mahadeen Z, Badran N, Izraiq M, Al-Mousa E. Prevalence of conventional risk factors in Jordanians with coronary heart disease: the Jordan Hyperlipidemia and Related Targets Study (JoHARTS). *International Journal of Cardiology* 2006;110(2):179-83.
113. Yang ZJ, Liu J, Ge JP, Chen L, Zhao ZG, Yang WY. Prevalence of cardiovascular disease risk factor in the Chinese population: The 2007-2008 China National Diabetes and Metabolic Disorders Study. *European Heart Journal* 2012;33(2):213-20.

114. Kawano H, Soejima H, Kojima S, Kitagawa A, Ogawa H, Japanese Acute Coronary Syndrome Study Investigators. Sex differences of risk factors for acute myocardial infarction in Japanese patients. *Circulation Journal* 2006;70(5):513-7.
115. Teo KK, Liu L, Chow CK, Wang X, Islam S, Jiang L, Sanderson JE, Rangarajan S, Yusuf S, INTERHEART Investigators in China. Potentially modifiable risk factors associated with myocardial infarction in China: the INTERHEART China study. *Heart* 2009;95(22):1857-64.
116. Reuterwall C, Hallqvist J, Ahlbom A, De Faire U, Diderichsen F, Hogstedt C, Pershagen G, Theorell T, Wiman B, Wolk A. Higher relative, but lower absolute risks of myocardial infarction in women than in men: analysis of some major risk factors in the SHEEP study. The SHEEP Study Group. *Journal of Internal Medicine* 1999;246(2):161-74.
117. Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, McQueen M, Budaj A, Pais P, Varigos J, Lisheng L, INTERHEART Study Investigators. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* 2004;364(9438):937-52.
118. Mushlin AI, Christos PJ, Abu-Raddad L, Chemaitelly H, Deleu D, Gehani AR. The importance of diabetes mellitus in the global epidemic of cardiovascular disease: the case of the state of Qatar. *Transactions of the American Clinical and Climatological Association* 2012;123:193-207.
119. Tavani A, Bertuzzi M, Gallus S, Negri E, La Vecchia C. Diabetes mellitus as a contributor to the risk of acute myocardial infarction. *Journal of Clinical Epidemiology* 2002;55(11):1082-7.
120. de Jesus Llibre J, Valhuerdi A, Fernandez O, Llibre JC, Porto R, Lopez AM, Marcheco B, Moreno C. Prevalence of stroke and associated risk factors in older adults in Havana City and Matanzas Provinces, Cuba (10/66 population-based study). *MEDICC Review* 2010;12(3):20-6.
121. Thompson SG, Greenberg G, Meade TW. Risk factors for stroke and myocardial infarction in women in the United Kingdom as assessed in general practice: a case-control study. *British Heart Journal* 1989;61(5):403-9.
122. O'Donnell MJ, Xavier D, Liu L, Zhang H, Chin SL, Rao-Melacini P, Rangarajan S, Islam S, Pais P, McQueen MJ, Mondo C, Damasceno A, Lopez-Jaramillo P, Hankey GJ, Dans AL, Yusuf K, Truelsen T, Diener H-C, Sacco RL, Ryglewicz D, Czlankowska A, Weimar C, Wang X, Yusuf S, INTERSTROKE investigators. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet* 2010;376(9735):112-23.
123. Giroud M, Creisson E, Fayolle H, Andre N, Becker F, Martin D, Dumas R. Risk factors for primary cerebral hemorrhage: a population-based study--the Stroke Registry of Dijon. *Neuroepidemiology* 1995;14(1):20-6.

124. Wu Y, Zhang L, Yuan X, Wu Y, Yi D. Quantifying links between stroke and risk factors: a study on individual health risk appraisal of stroke in a community of Chongqing. *Neurological Sciences* 2011;32(2):211-9.
125. Zodpey SP, Tiwari RR, Kulkarni HR. Risk factors for haemorrhagic stroke: a case-control study. *Public Health* 2000;114(3):177-82.
126. Abu-Zeid HA, Choi NW, Maini KK, Hsu PH, Nelson NA. Relative role of factors associated with cerebral infarction and cerebral hemorrhage. A matched pair case-control study. *Stroke* 1977;8(1):106-12.
127. Campbell PT, Newton CC, Patel AV, Jacobs EJ, Gapstur SM. Diabetes and cause-specific mortality in a prospective cohort of one million U.S. adults. *Diabetes Care* 2012;35(9):1835-44.
128. Mulnier HE, Seaman HE, Raleigh VS, Soedamah-Muthu SS, Colhoun HM, Lawrenson RA, de Vries CS. Risk of myocardial infarction in men and women with type 2 diabetes in the UK: a cohort study using the General Practice Research Database. *Diabetologia* 2008;51(9):1639-45.
129. Spencer EA, Pirie KL, Stevens RJ, Beral V, Brown A, Liu B, Green J, Reeves GK, Million Women Study Collaborators. Diabetes and modifiable risk factors for cardiovascular disease: the prospective Million Women Study. *European Journal of Epidemiology* 2008;23(12):793-9.
130. Shen C, Schooling CM, Chan WM, Lee SY, Leung GM, Lam TH. Self-reported diabetes and mortality in a prospective Chinese elderly cohort study in Hong Kong. *Preventive Medicine* 2014;64:20-6.
131. Liu J, Grundy SM, Wang W, Smith SC, Jr., Vega GL, Wu Z, Zeng Z, Wang W, Zhao D. Ten-year risk of cardiovascular incidence related to diabetes, prediabetes, and the metabolic syndrome. *American Heart Journal* 2007;153(4):552-8.
132. Wu Y, Liu X, Li X, Li Y, Zhao L, Chen Z, Li Y, Rao X, Zhou B, Detrano R, Liu K, USA-PRC Collaborative Study of Cardiovascular, Cardiopulmonary Epidemiology Research Group, China Multicenter Collaborative Study of Cardiovascular Epidemiology Research Group. Estimation of 10-year risk of fatal and nonfatal ischemic cardiovascular diseases in Chinese adults. *Circulation* 2006;114(21):2217-25.
133. 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;57(8):1542-51.
134. Woo J, Lau EMC. Risk factors predisposing to stroke in an elderly Chinese population - a longitudinal study. *Neuroepidemiology* 1990;9(3):131-4.

135. Hamer M, Batty GD, Stamatakis E, Kivimaki M. Comparison of risk factors for fatal stroke and ischemic heart disease: a prospective follow up of the health survey for England. *Atherosclerosis* 2011;219(2):807-10.
136. Chen H-F, Lee S-P, Li C-Y. Sex differences in the incidence of hemorrhagic and ischemic stroke among diabetics in Taiwan. *Journal of Women's Health* 2009;18(5):647-54.
137. Roche MM, Wang PP. Sex differences in all-cause and cardiovascular mortality, hospitalization for individuals with and without diabetes, and patients with diabetes diagnosed early and late. *Diabetes Care* 2013;36(9):2582-90.
138. Peters SAE, Huxley RR, Woodward M. Diabetes as a risk factor for stroke in women compared with men: a systematic review and meta-analysis of 64 cohorts, including 775,385 individuals and 12,539 strokes. *Lancet* 2014;383(9933):1973-80.
139. Hoogwerf BJ, Sprecher DL, Pearce GL, Acevedo M, Frolkis JP, Foody JM, Cross JA, Pashkow FJ, Robinson K, Vidt DG. Blood glucose concentrations ≤ 125 mg/dl and coronary heart disease risk. *American Journal of Cardiology* 2002;89(5):596-9.
140. Reunanen A, Pyorala K, Aromaa A, Maatela J, Knekt P. Glucose tolerance and coronary heart disease in middle-aged Finnish men: Social Insurance Institution's Coronary Heart Disease Study. *Journal of Chronic Diseases* 1979;32(11-12):747-58.
141. Gerstein HC, Pais P, Pogue J, Yusuf S. Relationship of glucose and insulin levels to the risk of myocardial infarction: a case-control study. *Journal of the American College of Cardiology* 1999;33(3):612-9.
142. Singh GM, Danaei G, Farzadfar F, Stevens GA, Woodward M, Wormser D, Kaptoge S, Whitlock G, Qiao Q, Lewington S, Di Angelantonio E, Vander Hoorn S, Lawes CM, Ali MK, Mozaffarian D, Ezzati M. The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. *PLoS ONE* 2013;8(7):e65174.
143. Lewington S, Bragg F, Clarke R. A review on metaanalysis of biomarkers: promises and pitfalls. *Clinical Chemistry* 2012;58(8):1192-204.
144. Sui X, Lavie CJ, Hooker SP, Lee DC, Colabianchi N, Lee CD, Blair SN. A prospective study of fasting plasma glucose and risk of stroke in asymptomatic men. *Mayo Clinic Proceedings* 2011;86(11):1042-9.
145. Burgess S, Thompson SG. Mendelian Randomization. Methods for Using Genetic Variants in Causal Estimation. 1st ed. Boca Raton, Florida: CRC Press; 2015.
146. Lewis SJ. Mendelian randomization as applied to coronary heart disease, including recent advances incorporating new technology. *Circulation Cardiovascular Genetics* 2010;3(1):109-17.

147. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *International Journal of Epidemiology* 2003;32(1):1-22.
148. Benn M, Tybjaerg-Hansen A, McCarthy MI, Jensen GB, Grande P, Nordestgaard BG. Nonfasting glucose, ischemic heart disease, and myocardial infarction: a mendelian randomization study. *Journal of the American College of Cardiology* 2012;59(25):2356-65.
149. Rasmussen-Torvik LJ, Li M, Kao WH, Couper D, Boerwinkle E, Bielinski SJ, Folsom AR, Pankow JS. Association of a fasting glucose genetic risk score with subclinical atherosclerosis: The Atherosclerosis Risk in Communities (ARIC) study. *Diabetes* 2011;60(1):331-5.
150. Ross S., Gerstein H.C., Eikelboom J., Anand S.S., Yusuf S., Pare G. Mendelian randomization analysis supports the causal role of dysglycaemia and diabetes in the risk of coronary artery disease. *European Heart Journal* 2015;36(23):1454-62.
151. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. *International Journal of Epidemiology* 2013;42(4):1134-44.
152. Burgess S, Thompson SG. Avoiding bias from weak instruments in Mendelian randomization studies. *International Journal of Epidemiology* 2011;40(3):755-64.
153. Xiao R, Boehnke M. Quantifying and correcting for the winner's curse in quantitative-trait association studies. *Genetic Epidemiology* 2011;35(3):133-8.
154. Chen J, Campbell T, Li J, Peto R, Boreham J, Feng Z, Youngman L. Diet, life-style, and mortality in China: a study of the characteristics of 65 Chinese counties. Oxford: Oxford University Press; 1990.
155. Chen J, Peto R, Pan W, Liu B, Campbell T, Boreham J, Parpia B, Cassano P, Chen Z, Feng Z, Gelband H, Li J, Pan H, Root M, Wu Y, Youngman L. Geographic Study of Mortality, Biochemistry, Diet and Lifestyle in Rural China. Oxford: Oxford University Press; 2006.
156. Li JY, Liu BQ, Li GY, Chen ZJ, Sun XI, Rong SD. Atlas of cancer mortality in the People's Republic of China. An aid for cancer control and research. *International Journal of Epidemiology* 1981;10(2):127-33.
157. Jiang C, Thomas GN, Lam TH, Schooling CM, Zhang W, Lao X, Adab P, Liu B, Leung GM, Cheng KK. Cohort profile: The Guangzhou Biobank Cohort Study, a Guangzhou-Hong Kong-Birmingham collaboration. *International Journal of Epidemiology* 2006;35(4):844-52.
158. Zheng W, Chow WH, Yang G, Jin F, Rothman N, Blair A, Li HL, Wen W, Ji BT, Li Q, Shu XO, Gao YT. The Shanghai Women's Health Study: rationale, study design, and baseline characteristics. *American Journal of Epidemiology* 2005;162(11):1123-31.

159. Niu SR, Yang GH, Chen ZM, Wang JL, Wang GH, He XZ, Schoepff H, Boreham J, Pan HC, Peto R. Emerging tobacco hazards in China: 2. Early mortality results from a prospective study. *BMJ* 1998;317(7170):1423-4.
160. Chen Z, Peto R, Collins R, MacMahon S, Lu J, Li W. Serum cholesterol concentration and coronary heart disease in population with low cholesterol concentrations. *BMJ* 1991;303(6797):276-82.
161. He J, Gu D, Wu X, Reynolds K, Duan X, Yao C, Wang J, Chen C-S, Chen J, Wildman RP, Klag MJ, Whelton PK. Major Causes of Death among Men and Women in China. *New England Journal of Medicine* 2005;353(11):1124-34.
162. Yuan JM, Ross RK, Gao YT, Yu MC. Body weight and mortality: a prospective evaluation in a cohort of middle-aged men in Shanghai, China. *International Journal of Epidemiology* 1998;27(5):824-32.
163. Sai XY, He Y, Men K, Wang B, Huang JY, Shi QL, Zhang L, Li LS, Choi BC, Yan YP. All-cause mortality and risk factors in a cohort of retired military male veterans, Xi'an, China: an 18-year follow up study. *BMC Public Health* 2007;7:290.
164. Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, Li L, China Kadoorie Biobank Collaborative Group. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *International Journal of Epidemiology* 2011;40(6):1652-66.
165. Doll R. Weak Associations in Epidemiology : Importance, Detection, and Interpretation. *Journal of Epidemiology* 1996;6(4):S11-20.
166. Collins R. UK biobank: the need for large prospective epidemiological studies. *Journal of Epidemiology and Community Health* 2011;65(Suppl 1):A37.
167. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *International Journal of Epidemiology* 2013;42(4):1012-4.
168. Kadoorie Study of Chronic Disease in China. [Manual for the survey in the field]. University of Oxford, Chinese Center for Disease Control and Prevention, 2004.
169. Kadoorie Study of Chronic Disease in China. [Instructions for use of the Johnson and Johnson SureStep Plus meter]. University of Oxford, Chinese Center for Disease Control and Prevention, 2004.
170. Kadoorie Study of Chronic Disease in China. [Hepatitis B virus surface antigen blood test strips instructions]. University of Oxford, Chinese Center for Disease Control and Prevention, 2004.
171. Kadoorie Study of Chronic Disease in China. [Blood tests in the field]. University of Oxford, Chinese Center for Disease Control and Prevention, 2004.

172. China Kadoorie Biobank. [Second resurvey blood sampling and on-site blood glucose and lipid testing instruction manual]. University of Oxford, Chinese Center for Disease Control and Prevention, 2013.
173. Clarke R, Shipley M, Lewington S, Youngman L, Collins R, Marmot M, Peto R. Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. *American Journal of Epidemiology* 1999;150(4):341-53.
174. MacMahon S, Peto R, Cutler J, Collins R, Sorlie P, Neaton J, Abbott R, Godwin J, Dyer A, Stamler J. Blood pressure, stroke, and coronary heart disease. Part 1, Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet* 1990;335(8692):765-74.
175. Berglund L. Regression dilution bias: tools for correction methods and sample size calculation. *Upsala Journal of Medical Sciences* 2012;117(3):279-83.
176. World Health Organization. International Classification of Diseases (ICD) Version 10 2010 [15th January 2014]. Available from: <http://apps.who.int/classifications/icd10/browse/2010/en>.
177. Yang G, Murray C, Zhang Z. Exploring adult mortality in China: levels, patterns and causes. Beijing: Hua Xia Press; 1991.
178. Yang G, Rao C, Ma J, Wang L, Wan X, Dubrovsky G, Lopez AD. Validation of verbal autopsy procedures for adult deaths in China. *International Journal of Epidemiology* 2006;35(3):741-8.
179. Yang W, Lu J, Weng J, Jia W, Ji L, Xiao J, Shan Z, Liu J, Tian H, Ji Q, Zhu D, Ge J, Lin L, Chen L, Guo X, Zhao Z, Li Q, Zhou Z, Shan G, He J, China National Diabetes and Metabolic Disorders Study Group. Prevalence of diabetes among men and women in China. *New England Journal of Medicine* 2010;362(12):1090-101.
180. Hu D, Sun L, Fu P, Xie J, Lu J, Zhou J, Yu D, Whelton PK, He J, Gu D. Prevalence and risk factors for type 2 diabetes mellitus in the Chinese adult population: the InterASIA Study. *Diabetes Research and Clinical Practice* 2009;84(3):288-95.
181. Liu S, Wang W, Zhang J, He Y, Yao C, Zeng Z, Piao J, Howard BV, Fabsitz RR, Best L, Yang X, Lee ET. Prevalence of diabetes and impaired fasting glucose in Chinese adults, China National Nutrition and Health Survey, 2002. *Preventing Chronic Disease* 2011;8(1):A13.
182. The Task Force on Diabetes and Cardiovascular Diseases of the European Society of Cardiology (ESC) and of the European Association for the Study of Diabetes (EASD). Guidelines on diabetes, pre-diabetes, and cardiovascular diseases: executive summary. *European Heart Journal* 2007;28(1):88-136.
183. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2012;35(S1):S64-S71.

184. Soedamah-Muthu SS, Fuller JH, Mulnier HE, Raleigh VS, Lawrenson RA, Colhoun HM. All-cause mortality rates in patients with type 1 diabetes mellitus compared with a non-diabetic population from the UK general practice research database, 1992-1999. *Diabetologia* 2006;49(4):660-6.
185. OneTouch SureStep [14th September 2012]. Available from: <http://www.onetouch.com/support/products/surestep>.
186. Rosner B, Willett WC, Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Statistics in Medicine* 1989;8(9):1051-69.
187. Knuiman MW, Divitini ML, Buzas JS, Fitzgerald PE. Adjustment for regression dilution in epidemiological regression analyses. *Annals of Epidemiology* 1998;8(1):56-63.
188. Thompson S, Kaptoge S, White I, Wood A, Perry P, Danesh J, Emerging Risk Factors Collaboration. Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies. *International Journal of Epidemiology* 2010;39(5):1345-59.
189. Fibrinogen Studies Collaboration. Correcting for multivariate measurement error by regression calibration in meta-analyses of epidemiological studies. *Statistics in Medicine* 2009;28(7):1067-92.
190. Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. *American Journal of Epidemiology* 1990;132(4):734-45.
191. Chen M, Bergman RN, Pacini G, Porte D, Jr. Pathogenesis of age-related glucose intolerance in man: insulin resistance and decreased beta-cell function. *Journal of Clinical Endocrinology and Metabolism* 1985;60(1):13-20.
192. Wang H, Qiu Q, Tan LL, Liu T, Deng XQ, Chen YM, Chen W, Yu XQ, Hu BJ, Chen WQ. Prevalence and determinants of diabetes and impaired fasting glucose among urban community-dwelling adults in Guangzhou, China. *Diabetes and Metabolism* 2009;35(5):378-84.
193. Li H, Oldenburg B, Chamberlain C, O'Neil A, Xue B, Jolley D, Hall R, Dong Z, Guo Y. Diabetes prevalence and determinants in adults in China mainland from 2000 to 2010: a systematic review. *Diabetes Research and Clinical Practice* 2012;98(2):226-35.
194. Villegas R, Shu X-O, Li H, Yang G, Matthews CE, Leitzmann M, Li Q, Cai H, Gao Y-T, Zheng W. Physical activity and the incidence of type 2 diabetes in the Shanghai women's health study. *International Journal of Epidemiology* 2006;35(6):1553-62.
195. Qin L, Corpeleijn E, Jiang C, Thomas GN, Schooling CM, Zhang W, Cheng KK, Leung GM, Stolk RP, Lam TH. Physical activity, adiposity, and diabetes risk in

middle-aged and older Chinese population: the Guangzhou Biobank Cohort Study. *Diabetes Care* 2010;33(11):2342-8.

196. Sattar N, Gill JM. Type 2 diabetes as a disease of ectopic fat? *BMC Medicine* 2014;12:123.
197. Hamdy O, Porramatikul S, Al-Ozairi E. Metabolic obesity: the paradox between visceral and subcutaneous fat. *Current Diabetes Reviews* 2006;2(4):367-73.
198. Hu D, Xie J, Fu P, Zhou J, Yu D, Whelton PK, He J, Gu D. Central rather than overall obesity is related to diabetes in the Chinese population: the InterASIA study. *Obesity* 2007;15(11):2809-16.
199. Wildman RP, Gu D, Reynolds K, Duan X, Wu X, He J. Are waist circumference and body mass index independently associated with cardiovascular disease risk in Chinese adults? *American Journal of Clinical Nutrition* 2005;82(6):1195-202.
200. He Y, Zhai F, Ma G, Feskens EJ, Zhang J, Fu P, Van't Veer P, Yang X. Abdominal obesity and the prevalence of diabetes and intermediate hyperglycaemia in Chinese adults. *Public Health Nutrition* 2009;12(8):1078-84.
201. Klein S. The case of visceral fat: argument for the defense. *Journal of Clinical Investigation* 2004;113(11):1530-2.
202. Nazare JA, Smith JD, Borel AL, Haffner SM, Balkau B, Ross R, Massien C, Almeras N, Despres JP. Ethnic influences on the relations between abdominal subcutaneous and visceral adiposity, liver fat, and cardiometabolic risk profile: the International Study of Prediction of Intra-Abdominal Adiposity and Its Relationship With Cardiometabolic Risk/Intra-Abdominal Adiposity. *American Journal of Clinical Nutrition* 2012;96(4):714-26.
203. World Health Organization expert consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet* 2004;363(9403):157-63.
204. Manolopoulos KN, Karpe F, Frayn KN. Gluteofemoral body fat as a determinant of metabolic health. *International Journal of Obesity* 2010;34(6):949-59.
205. Lago RM, Singh PP, Nesto RW. Diabetes and hypertension. *Nature Clinical Practice Endocrinology and Metabolism* 2007;3(10):667.
206. Huxley R, Barzi F, Stolk R, Caterson I, Gill T, Lam TH, Omari A, Woodward M, Obesity in Asia Collaboration. Ethnic comparisons of obesity in the Asia-Pacific region: protocol for a collaborative overview of cross-sectional studies. *Obesity Reviews* 2005;6(3):193-8.
207. Midthjell K, Holmen J, Bjorndal A, Lund-Larsen PG. Is questionnaire information valid in the study of a chronic disease such as diabetes? The Nord-Trondelag Diabetes Study. *Journal of Epidemiology and Community Health* 1992;46(5):537-42.

208. Wada K, Yatsuya H, Ouyang P, Otsuka R, Mitsuhashi H, Takefuji S, Matsushita K, Sugiura K, Hotta Y, Toyoshima H, Tamakoshi K. Self-reported medical history was generally accurate among Japanese workplace population. *Journal of Clinical Epidemiology* 2009;62(3):306-13.
209. Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *Journal of Clinical Epidemiology* 2004;57(10):1096-103.
210. Haapanen N, Miilunpalo S, Pasanen M, Oja P, Vuori I. Agreement between questionnaire data and medical records of chronic diseases in middle-aged and elderly Finnish men and women. *American Journal of Epidemiology* 1997;145(8):762-9.
211. Kehoe R, Wu SY, Leske MC, Chylack LT, Jr. Comparing self-reported and physician-reported medical history. *American Journal of Epidemiology* 1994;139(8):813-8.
212. Qiao Q, Pyorala K, Pyorala M, Nissinen A, Lindstrom J, Tilvis R, Tuomilehto J. Two-hour glucose is a better risk predictor for incident coronary heart disease and cardiovascular mortality than fasting glucose. *European Heart Journal* 2002;23(16):1267-75.
213. Youngman LD, Clark S, Manley S, Peto R, Collins R. Reliable Measurement of Glycated Hemoglobin in Frozen Blood Samples: Implications for Epidemiologic Studies. *Clinical Chemistry* 2002;48(9):1627-9.
214. Rolka DB, Narayan KM, Thompson TJ, Goldman D, Lindenmayer J, Alich K, Bacall D, Benjamin EM, Lamb B, Stuart DO, Engelgau MM. Performance of recommended screening tests for undiagnosed diabetes and dysglycemia. *Diabetes Care* 2001;24(11):1899-903.
215. Engelgau MM, Thompson TJ, Smith PJ, Herman WH, Aubert RE, Gunter EW, Wetterhall SF, Sous ES, Ali MA. Screening for diabetes mellitus in adults. The utility of random capillary blood glucose measurements. *Diabetes Care* 1995;18(4):463-6.
216. Qiao Q, Keinänen-Kiukaanniemi S, Rajala U, Uusimaki A, Kivela SL. Random capillary whole blood glucose test as a screening test for diabetes mellitus in a middle-aged population. *Scandinavian Journal of Clinical and Laboratory Investigation* 1995;55(1):3-8.
217. Ziemer DC, Kolm P, Foster JK, Weintraub WS, Vaccarino V, Rhee MK, Varughese RM, Tsui CW, Koch DD, Twombly JG, Narayan KM, Phillips LS. Random plasma glucose in serendipitous screening for glucose intolerance: screening for impaired glucose tolerance study 2. *Journal of General Internal Medicine* 2008;23(5):528-35.
218. Tchobroutsky G. Blood glucose levels in diabetic and non-diabetic subjects. *Diabetologia* 1991;34(2):67-73.

219. Barrett-Connor E. Factors associated with the distribution of fasting plasma glucose in an adult community. *American Journal of Epidemiology* 1980;112(4):518-23.
220. Lim TO, Ding LM, Zaki M, Suleiman AB, Kew ST, Maimunah AH, Rugayah B. Distribution of blood glucose in a national sample of Malaysian adults. *Medical Journal of Malaysia* 2000;55(1):65-77.
221. The DECODE Study Group. Age- and sex-specific prevalences of diabetes and impaired glucose regulation in 13 European cohorts. *Diabetes Care* 2003;26(1):61-9.
222. The DECODA Study Group. Age- and sex-specific prevalence of diabetes and impaired glucose regulation in 11 Asian cohorts. *Diabetes Care* 2003;26(6):1770-80.
223. Ying X, Song Z-y, Zhao C-j, Jiang Y. Body mass index, waist circumference, and cardiometabolic risk factors in young and middle-aged Chinese women. *Journal of Zhejiang University SCIENCE B* 2010;11(9):639-46.
224. Emerging Risk Factors Collaboration. Glycated hemoglobin measurement and prediction of cardiovascular disease. *JAMA- The Journal of the American Medical Association* 2014;311(12):1225-33.
225. Fibrinogen Studies Collaboration. Regression dilution methods for meta-analysis: assessing long-term variability in plasma fibrinogen among 27 247 adults in 15 prospective studies. *International Journal of Epidemiology* 2006;35(6):1570-8.
226. Moebus S, Gores L, Losch C, Jockel KH. Impact of time since last caloric intake on blood glucose levels. *European Journal of Epidemiology* 2011;26(9):719-28.
227. Kawada T. Random blood glucose measurement for epidemiological studies: its significance and limitations. *Current Medical Research and Opinion* 2012;28(3):447-8.
228. Savoca R, Jaworek B, Huber AR. New "plasma referenced" POCT glucose monitoring systems--are they suitable for glucose monitoring and diagnosis of diabetes? *Clinica Chimica Acta* 2006;372(1-2):199-201.
229. Boyd R, Leigh B, Stuart P. Capillary versus venous bedside blood glucose estimations. *Emergency Medicine Journal* 2005;22(3):177-9.
230. Kumar G, Sng BL, Kumar S. Correlation of capillary and venous blood glucometry with laboratory determination. *Prehospital Emergency Care* 2004;8(4):378-83.
231. Sacks DB, Bruns DE, Goldstein DE, Maclaren NK, McDonald JM, Parrott M. Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. *Clinical Chemistry* 2002;48(3):436-72.

232. Olansky L, Kennedy L. Finger-stick glucose monitoring: issues of accuracy and specificity. *Diabetes Care* 2010;33(4):948-9.
233. Tonyushkina K, Nichols JH. Glucose meters: a review of technical challenges to obtaining accurate results. *Journal of Diabetes Science and Technology* 2009;3(4):971-80.
234. Acon Laboratories Inc. On Call Advanced Blood Glucose Monitoring System User's Manual. San Diego Acon Laboratories Inc; 2010.
235. Evaluation Report for the Kadoorie Biobank Study POC devices. Oxford: Clinical Trial Service Unit and Epidemiological Studies Unit, 2013.
236. Global Burden of Disease Study 2010. China Global Burden of Disease Study 2010 (GBD 2010) Results 1990-2010 Seattle, United States: Institute for Health Metrics and Evaluation (IHME); 2013 [cited 2014 21st May]. Available from: <http://ghdx.healthdata.org/record/china-global-burden-disease-study-2010-gbd-2010-results-1990-2010>.
237. World Health Organization. China health databank 2011. Geneva: 2011.
238. Guha S, Ghosh A, Chatterjee N, Pal S, Guha S, Chatterjee S, Deb PK, Dawn B. Risk factors for coronary heart disease in Indians: a case-control study from Eastern India. *Indian Heart Journal* 2005;57(6):738-40.
239. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10(1):37-48.
240. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Medical Research Methodology* 2008;8:70.
241. 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. *Statistics in Medicine* 1991;10(7):1025-35.
242. Early Breast Cancer Trialists' Collaborative Group. Arithmetic details of tests for trend, for heterogeneity and for interaction. *Treatment of early breast cancer Worldwide evidence 1985-1990 A systematic overview of all available randomized trials of adjuvant endocrine and cytotoxic therapy*. 1. Oxford: Oxford University Press; 1990.
243. Glynn RJ, Rosner B. Comparison of risk factors for the competing risks of coronary heart disease, stroke, and venous thromboembolism. *American Journal of Epidemiology* 2005;162(10):975-82.
244. Hart CL, Hole DJ, Smith GD. Risk factors and 20-year stroke mortality in men and women in the Renfrew/Paisley study in Scotland. *Stroke* 1999;30(10):1999-2007.

245. Huxley R, Barzi F, Woodward M. Excess risk of fatal coronary heart disease associated with diabetes in men and women: meta-analysis of 37 prospective cohort studies. *BMJ* 2006;332(7533):73-8.
246. Rodriguez BL, Lau N, Burchfiel CM, Abbott RD, Sharp DS, Yano K, Curb JD. Glucose intolerance and 23-year risk of coronary heart disease and total mortality: the Honolulu Heart Program. *Diabetes Care* 1999;22(8):1262-5.
247. Burchfiel CM, Curb JD, Rodriguez BL, Abbott RD, Chiu D, Yano K. Glucose intolerance and 22-year stroke incidence. The Honolulu Heart Program. *Stroke* 1994;25(5):951-7.
248. Gerstein H, Punthakee Z. Dysglycemia and the risk of cardiovascular events. In: Yusuf S, Cairns JA, Camm AJ, Fallen EL, Gersh BJ, editors. *Evidence-Based Cardiology*. Third ed. Chichester: Wiley-Blackwell; 2010. p. 179-89.
249. Chien K-L, Lee B-C, Lin H-J, Hsu H-C, Chen M-F. Association of fasting and post-prandial hyperglycemia on the risk of cardiovascular and all-cause death among non-diabetic Chinese. *Diabetes Research and Clinical Practice* 2009;83(2):e47-50.
250. Hart CL, Hole DJ, Smith GD. Comparison of risk factors for stroke incidence and stroke mortality in 20 years of follow-up in men and women in the Renfrew/Paisley Study in Scotland. *Stroke* 2000;31(8):1893-6.
251. Balkau B, Shipley M, Jarrett RJ, Pyorala K, Pyorala M, Forhan A, Eschwege E. High blood glucose concentration is a risk factor for mortality in middle-aged nondiabetic men. 20-year follow-up in the Whitehall Study, the Paris Prospective Study, and the Helsinki Policemen Study. *Diabetes Care* 1998;21(3):360-7.
252. Allen N, Sudlow C, Downey P, Peakman T, Danesh J, Elliott P, Gallacher J, Green J, Matthews P, Pell J, Sprosen T, Collins R. UK Biobank: Current status and what it means for epidemiology. *Health Policy and Technology* 2012;1(3):123-6.
253. Liu X, Xu G, Wu W, Zhang R, Yin Q, Zhu W. Subtypes and one-year survival of first-ever stroke in Chinese patients: The Nanjing Stroke Registry. *Cerebrovascular Diseases* 2006;22(2-3):130-6.
254. Barr EL, Tonkin AM, Welborn TA, Shaw JE. Validity of self-reported cardiovascular disease events in comparison to medical record adjudication and a statewide hospital morbidity database: the AusDiab study. *Internal Medicine Journal* 2009;39(1):49-53.
255. Machon M, Arriola L, Larranaga N, Amiano P, Moreno-Iribas C, Agudo A, Ardanaz E, Barricarte A, Buckland G, Chirlaque MD, Gavrila D, Huerta JM, Martinez C, Molina E, Navarro C, Quiros JR, Rodriguez L, Sanchez MJ, Gonzalez CA, Dorronsoro M. Validity of self-reported prevalent cases of stroke and acute myocardial infarction in the Spanish cohort of the EPIC study. *Journal of Epidemiology and Community Health* 2013;67(1):71-5.

256. Lampe FC, Walker M, Lennon LT, Whincup PH, Ebrahim S. Validity of a self-reported history of doctor-diagnosed angina. *Journal of Clinical Epidemiology* 1999;52(1):73-81.
257. Engstad T, Bonna KH, Viitanen M. Validity of self-reported stroke : The Tromso Study. *Stroke* 2000;31(7):1602-7.
258. Yamagishi K, Ikeda A, Iso H, Inoue M, Tsugane S, JPHC Study Group. Self-reported stroke and myocardial infarction had adequate sensitivity in a population-based prospective study JPHC (Japan Public Health Center)-based Prospective Study. *Journal of Clinical Epidemiology* 2009;62(6):667-73.
259. Jackson CA, Mishra GD, Tooth L, Byles J, Dobson A. Moderate agreement between self-reported stroke and hospital-recorded stroke in two cohorts of Australian women: a validation study. *BMC Medical Research Methodology* 2015;15:7.
260. American Diabetes Association. 6. Glycemic Targets. *Diabetes Care* 2015;38(S1):S33-S40.
261. Tunstall-Pedoe H, Kuulasmaa K, Amouyel P, Arveiler D, Rajakangas AM, Pajak A. Myocardial infarction and coronary deaths in the World Health Organization MONICA Project. Registration procedures, event rates, and case-fatality rates in 38 populations from 21 countries in four continents. *Circulation* 1994;90(1):583-612.
262. Zhao D, Liu J, Wang W, Zeng Z, Cheng J, Liu J, Sun J, Wu Z. Epidemiological transition of stroke in China: twenty-one-year observational study from the Sino-MONICA-Beijing Project. *Stroke* 2008;39(6):1668-74.
263. Jiang B, Wang W-z, Chen H, Hong Z, Yang Q-d, Wu S-p, Du X-l, Bao Q-j. Incidence and trends of stroke and its subtypes in China: results from three large cities. *Stroke* 2006;37(1):63-8.
264. Institute for Health Metrics and Evaluation. GBD Profile: China. Seattle: Institute for Health Metrics and Evaluation, 2013.
265. Lacey B. Blood pressure and stroke pathological types in China: an analysis of 500,000 men and women in the China Kadoorie Biobank study. Oxford: University of Oxford; 2013.
266. Wu B, Lin S, Hao Z, Yang J, Xu Y, Wu L, Zhang S, Liu M. Proportion, risk factors and outcome of lacunar infarction: a hospital-based study in a Chinese population. *Cerebrovascular Diseases* 2010;29(2):181-7.
267. Wannamethee G, Whincup PH, Shaper AG, Walker M, MacFarlane PW. Factors determining case fatality in myocardial infarction "who dies in a heart attack"? *British Heart Journal* 1995;74(3):324-31.
268. He J, Klag MJ, Wu Z, Whelton PK. Stroke in the People's Republic of China: I. Geographic Variations in Incidence and Risk Factors. *Stroke* 1995;26(12):2222-7.

269. Wang L, Kong L, Wu F, Bai Y, Burton R. Preventing chronic diseases in China. *Lancet* 2005;366(9499):1821-4.
270. UK Prospective Diabetes Study Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet* 1998;352(9131):837-53.
271. The Diabetes Control and Complications Trial Research Group. The Effect of Intensive Treatment of Diabetes on the Development and Progression of Long-Term Complications in Insulin-Dependent Diabetes Mellitus. *New England Journal of Medicine* 1993;329(14):977-86.
272. Stratton IM, Adler AI, Neil HA, Matthews DR, Manley SE, Cull CA, Hadden D, Turner RC, Holman RR. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. *BMJ* 2000;321(7258):405-12.
273. Khaw K-T, Wareham N. Glycated hemoglobin as a marker of cardiovascular risk. *Current Opinion in Lipidology* 2006;17(6):637-43.
274. Duckworth W, Abraira C, Moritz T, Reda D, Emanuele N, Reaven PD, Zieve FJ, Marks J, Davis SN, Hayward R, Warren SR, Goldman S, McCarren M, Vitek ME, Henderson WG, Huang GD, VADT Investigators. Glucose control and vascular complications in veterans with type 2 diabetes. *New England Journal of Medicine* 2009;360(2):129-39.
275. Advance Collaborative Group, Patel A, MacMahon S, Chalmers J, Neal B, Billot L, Woodward M, Marre M, Cooper M, Glasziou P, Grobbee D, Hamet P, Harrap S, Heller S, Liu L, Mancia G, Mogensen CE, Pan C, Poulter N, Rodgers A, Williams B, Bompoint S, de Galan BE, Joshi R, Travert F. Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. *New England Journal of Medicine* 2008;358(24):2560-72.
276. Accord Study Group, Gerstein HC, Miller ME, Genuth S, Ismail-Beigi F, Buse JB, Goff DC, Jr., Probstfield JL, Cushman WC, Ginsberg HN, Bigger JT, Grimm RH, Jr., Byington RP, Rosenberg YD, Friedewald WT. Long-term effects of intensive glucose lowering on cardiovascular outcomes. *New England Journal of Medicine* 2011;364(9):818-28.
277. Hemmingsen B, Lund SS, Gluud C, Vaag A, Almdal T, Hemmingsen C, Wetterslev J. Intensive glycaemic control for patients with type 2 diabetes: systematic review with meta-analysis and trial sequential analysis of randomised clinical trials. *BMJ* 2011;343:d6898.
278. Holman RR, Paul SK, Bethel MA, Matthews DR, Neil HAW. 10-year follow-up of intensive glucose control in type 2 diabetes. *New England Journal of Medicine* 2008;359(15):1577-89.

279. Nathan DM, Cleary PA, Backlund J-YC, Genuth SM, Lachin JM, Orchard TJ, Raskin P, Zinman B, The Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) Study Research Group. Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. *New England Journal of Medicine* 2005;353(25):2643-53.
280. Hayward RA, Reaven PD, Wiitala WL, Bahn GD, Reda DJ, Ge L, McCarren M, Duckworth WC, Emanuele NV. Follow-up of Glycemic Control and Cardiovascular Outcomes in Type 2 Diabetes. *New England Journal of Medicine* 2015;372(23):2197-206.
281. UK Prospective Diabetes Study Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. *BMJ* 1998;317(7160):703-13.
282. Cholesterol Treatment Trialists' Collaborators. Efficacy of cholesterol-lowering therapy in 18 686 people with diabetes in 14 randomised trials of statins: a meta-analysis. *Lancet* 2008;371(9607):117-25.
283. American Diabetes Association. 8. Cardiovascular Disease and Risk Management. *Diabetes Care* 2015;38(S1):S49-S57.
284. Emberson JR. Personal communication. 2015.
285. UK Prospective Diabetes Study Group. U.K. Prospective Diabetes Study 27. Plasma lipids and lipoproteins at diagnosis of NIDDM by age and sex. *Diabetes Care* 1997;20(11):1683-7.
286. Kanaya AM, Grady D, Barrett-Connor E. Explaining the sex difference in coronary heart disease mortality among patients with type 2 diabetes mellitus: a meta-analysis. *Archives of Internal Medicine* 2002;162(15):1737-45.
287. Barrett-Connor EL, Cohn BA, Wingard DL, Edelstein SL. Why is diabetes mellitus a stronger risk factor for fatal ischemic heart disease in women than in men?: The Rancho Bernardo Study. *JAMA- The Journal of the American Medical Association* 1991;265(5):627-31.
288. Regensteiner JG, Golden S, Huebschmann AG, Barrett-Connor E, Chang AY, Chyun D, Fox CS, Kim C, Mehta N, Reckelhoff JF, Reusch JEB, Rexrode KM, Sumner AE, Welty FK, Wenger NK, Anton B. Sex Differences in the Cardiovascular Consequences of Diabetes Mellitus: A Scientific Statement From the American Heart Association. *Circulation* 2015;doi:10.1161/cir.0000000000000343.
289. Dale AC, Nilsen TI, Vatten L, Midthjell K, Wiseth R. Diabetes mellitus and risk of fatal ischaemic heart disease by gender: 18 years follow-up of 74,914 individuals in the HUNT 1 Study. *European Heart Journal* 2007;28(23):2924-9.
290. Wexler DJ, Grant RW, Meigs JB, Nathan DM, Cagliero E. Sex disparities in treatment of cardiac risk factors in patients with type 2 diabetes. *Diabetes Care* 2005;28(3):514-20.

291. Arnetz L, Ekberg NR, Alvarsson M. Sex differences in type 2 diabetes: focus on disease course and outcomes. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy* 2014;7:409-20.
292. Lee WL, Cheung AM, Cape D, Zinman B. Impact of diabetes on coronary artery disease in women and men: a meta-analysis of prospective studies. *Diabetes Care* 2000;23(7):962-8.
293. Witt DR, Boucher JL, Hayes SN. The Problem of Diabetes and Cardiovascular Disease: Include Women in the Solution. *Diabetes Spectrum* 2013;26(3):139-41.
294. Rivellese AA, Riccardi G, Vaccaro O. Cardiovascular risk in women with diabetes. *Nutrition, Metabolism, and Cardiovascular Diseases* 2010;20(6):474-80.
295. Cui R, Iso H, Yamagishi K, Saito I, Kokubo Y, Inoue M, Tsugane S. Diabetes mellitus and risk of stroke and its subtypes among Japanese: the Japan public health center study. *Stroke* 2011;42(9):2611-4.
296. Wei WQ, Denny JC. Extracting research-quality phenotypes from electronic health records to support precision medicine. *Genome Medicine* 2015;7(1):41.
297. McCormick N, Lacaille D, Bhole V, Avina-Zubieta JA. Validity of myocardial infarction diagnoses in administrative databases: a systematic review. *PLoS ONE* 2014;9(3):e92286.
298. Herrett E, Shah AD, Boggon R, Denaxas S, Smeeth L, van Staa T, Timmis A, Hemingway H. Completeness and diagnostic validity of recording acute myocardial infarction events in primary care, hospital care, disease registry, and national mortality records: cohort study. *BMJ* 2013;346:f2350.
299. Wang L, Yang G, Jiemin M, Rao C, Wan X, Dubrovsky G, Lopez AD. Evaluation of the quality of cause of death statistics in rural China using verbal autopsies. *Journal of Epidemiology and Community Health* 2007;61(6):519-26.
300. Rao C, Yang G, Hu J, Ma J, Xia W, Lopez AD. Validation of cause-of-death statistics in urban China. *International Journal of Epidemiology* 2007;36(3):642-51.
301. Wildenschild C, Mehnert F, Thomsen RW, Iversen HK, Vestergaard K, Ingeman A, Johnsen SP. Registration of acute stroke: validity in the Danish Stroke Registry and the Danish National Registry of Patients. *Clinical Epidemiology* 2014;6:27-36.
302. Aboa-Eboule C, Mengue D, Benzenine E, Hommel M, Giroud M, Bejot Y, Quantin C. How accurate is the reporting of stroke in hospital discharge data? A pilot validation study using a population-based stroke registry as control. *Journal of Neurology* 2013;260(2):605-13.
303. Cheng CL, Kao YH, Lin SJ, Lee CH, Lai ML. Validation of the National Health Insurance Research Database with ischemic stroke cases in Taiwan. *Pharmacoepidemiology and Drug Safety* 2011;20(3):236-42.

304. Leibson CL, Naessens JM, Brown RD, Whisnant JP. Accuracy of hospital discharge abstracts for identifying stroke. *Stroke* 1994;25(12):2348-55.
305. Lindsted KD, Fraser GE, Steinkohl M, Beeson WL. Healthy volunteer effect in a cohort study: Temporal resolution in the adventist health study. *Journal of Clinical Epidemiology* 1996;49(7):783-90.
306. Ford ES, Zhao G, Li C. Pre-diabetes and the risk for cardiovascular disease: a systematic review of the evidence. *Journal of the American College of Cardiology* 2010;55(13):1310-7.
307. Goto A, Arah OA, Goto M, Terauchi Y, Noda M. Severe hypoglycaemia and cardiovascular disease: systematic review and meta-analysis with bias analysis. *BMJ* 2013;347:f4533.
308. Chien K-L, Hsu H-C, Su T-C, Chen M-F, Lee Y-T, Hu FB. Fasting and postchallenge hyperglycemia and risk of cardiovascular disease in Chinese: the Chin-Shan Community Cardiovascular Cohort study. *American Heart Journal* 2008;156(5):996-1002.
309. Wannamethee SG, Perry IJ, Shaper AG. Nonfasting serum glucose and insulin concentrations and the risk of stroke. *Stroke* 1999;30(9):1780-6.
310. Chiasson J-L, Josse RG, Gomis R, Hanefeld M, Karasik A, Laakso M, Stop-NIDDM Trial Research Group. Acarbose treatment and the risk of cardiovascular disease and hypertension in patients with impaired glucose tolerance: the STOP-NIDDM trial. *JAMA- The Journal of the American Medical Association* 2003;290(4):486-94.
311. Dream Trial Investigators, Gerstein HC, Yusuf S, Bosch J, Pogue J, Sheridan P, Dinccag N, Hanefeld M, Hoogwerf B, Laakso M, Mohan V, Shaw J, Zinman B, Holman RR. Effect of rosiglitazone on the frequency of diabetes in patients with impaired glucose tolerance or impaired fasting glucose: a randomised controlled trial. *Lancet* 2006;368(9541):1096-105.
312. Origin Trial Investigators, Gerstein HC, Bosch J, Dagenais GR, Diaz R, Jung H, Maggioni AP, Pogue J, Probstfield J, Ramachandran A, Riddle MC, Ryden LE, Yusuf S. Basal insulin and cardiovascular and other outcomes in dysglycemia. *New England Journal of Medicine* 2012;367(4):319-28.
313. NAVIGATOR Study Group. Effect of nateglinide on the incidence of diabetes and cardiovascular events. *New England Journal of Medicine* 2010;362(16):1463-76.
314. Davey Smith G, Phillips AN. Inflation in epidemiology: "The proof and measurement of association between two things" revisited. *BMJ* 1996;312(7047):1659-61.
315. Carroll R. Measurement error in epidemiologic studies Texas [cited 2015 11th August 2015]. Available from: <http://www.biostat.washington.edu/~yanez/b578D/materials/gail03.pdf>.
316. Lindsberg PJ, Roine RO. Hyperglycemia in Acute Stroke. *Stroke* 2004;35(2):363-4.

317. Ceriello A. Acute hyperglycaemia: a 'new' risk factor during myocardial infarction. *European Heart Journal* 2005;26(4):328-31.
318. Stears AJ, Woods SH, Watts MM, Burton TJ, Graggaber J, Mir FA, Brown MJ. A double-blind, placebo-controlled, crossover trial comparing the effects of amiloride and hydrochlorothiazide on glucose tolerance in patients with essential hypertension. *Hypertension* 2012;59(5):934-42.
319. Wilson PW, Odell PM, Kannel WB. An updated coronary risk profile. A statement for health professionals. *Circulation* 1991;83(1):356-62.
320. Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, Brindle P. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ* 2008;336(7659):1475-82.
321. Zoungas S, de Galan BE, Ninomiya T, Grobbee D, Hamet P, Heller S, MacMahon S, Marre M, Neal B, Patel A, Woodward M, Chalmers J, on behalf of the ADVANCE Collaborative Group. Combined Effects of Routine Blood Pressure Lowering and Intensive Glucose Control on Macrovascular and Microvascular Outcomes in Patients With Type 2 Diabetes: New results from the ADVANCE trial. *Diabetes Care* 2009;32(11):2068-74.
322. Davey Smith G, Ebrahim S, Lewis S, Palmer LJ. Population health aspects of genetic epidemiology: genomic profiling, personalised medicine and Mendelian randomisation. In: Palmer LJ, Burton PR, Davey Smith G, editors. *An introduction to genetic epidemiology*. Bristol: The Policy Press; 2011. p. 175-215.
323. Cauchi S, Proenca C, Choquet H, Gaget S, De Graeve F, Marre M, Balkau B, Tichet J, Meyre D, Vaxillaire M, Froguel P, D.E.S.I.R. Study Group. Analysis of novel risk loci for type 2 diabetes in a general French population: the D.E.S.I.R. study. *Journal of Molecular Medicine* 2008;86(3):341-8.
324. Takeuchi F, Katsuya T, Chakrewarthy S, Yamamoto K, Fujioka A, Serizawa M, Fujisawa T, Nakashima E, Ohnaka K, Ikegami H, Sugiyama T, Nabika T, Kasturiratne A, Yamaguchi S, Kono S, Takayanagi R, Yamori Y, Kobayashi S, Ogihara T, de Silva A, Wickremasinghe R, Kato N. Common variants at the GCK, GCKR, G6PC2-ABCB11 and MTNR1B loci are associated with fasting glucose in two Asian populations. *Diabetologia* 2010;53(2):299-308.
325. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, Zeggini E, Huth C, Aulchenko YS, Thorleifsson G, McCulloch LJ, Ferreira T, Grallert H, Amin N, Wu G, Willer CJ, Raychaudhuri S, McCarroll SA, Langenberg C, Hofmann OM, Dupuis J, Qi L, Segre AV, van Hoek M, Navarro P, Ardlie K, Balkau B, Benediktsson R, Bennett AJ, Blagieva R, Boerwinkle E, Bonnycastle LL, Bengtsson Bostrom K, Bravenboer B, Bumpstead S, Burt NP, Charpentier G, Chines PS, Cornelis M, Couper DJ, Crawford G, Doney ASF, Elliott KS, Elliott AL, Erdos MR, Fox CS, Franklin CS, Ganser M, Gieger C, Grarup N, Green T, Griffin S, Groves CJ, Guiducci C, Hadjadj S, Hassanali N, Herder C, Isomaa B, Jackson AU, Johnson PRV, Jorgensen T, Kao WHL, Klopp N, Kong A, Kraft P, Kuusisto J, Lauritzen T, Li M, Lieveise A, Lindgren CM, Lyssenko V, Marre M, Meitinger T, Midthjell K, Morken MA, Narisu N, Nilsson P, Owen KR, Payne F, Perry JRB, Petersen A-K, Platou C,

Proenca C, Prokopenko I, Rathmann W, Rayner NW, Robertson NR, Rocheleau G, Roden M, Sampson MJ, Saxena R, Shields BM, Shrader P, Sigurdsson G, Sparso T, Strassburger K, Stringham HM, Sun Q, Swift AJ, Thorand B, Tichet J, Tuomi T, van Dam RM, van Haeften TW, van Herpt T, van Vliet-Ostaptchouk JV, Walters GB, Weedon MN, Wijmenga C, Witteman J, Bergman RN, Cauchi S, Collins FS, Gloyn AL, Gyllenstein U, Hansen T, Hide WA, Hitman GA, Hofman A, Hunter DJ, Hveem K, Laakso M, Mohlke KL, Morris AD, Palmer CNA, Pramstaller PP, Rudan I, Sijbrands E, Stein LD, Tuomilehto J, Uitterlinden A, Walker M, Wareham NJ, Watanabe RM, Abecasis GR, Boehm BO, Campbell H, Daly MJ, Hattersley AT, Hu FB, Meigs JB, Pankow JS, Pedersen O, Wichmann HE, Barroso I, Florez JC, Frayling TM, Groop L, Sladek R, Thorsteinsdottir U, Wilson JF, Illig T, Froguel P, van Duijn CM, Stefansson K, Altshuler D, Boehnke M, McCarthy MI. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nature Genetics* 2010;42(7):579-89.

326. Saxena R, Hivert MF, Langenberg C, Tanaka T, Pankow JS, Vollenweider P, Lyssenko V, Bouatia-Naji N, Dupuis J, Jackson AU, Kao WH, Li M, Glazer NL, Manning AK, Luan J, Stringham HM, Prokopenko I, Johnson T, Grarup N, Boesgaard TW, Lecoeur C, Shrader P, O'Connell J, Ingelsson E, Couper DJ, Rice K, Song K, Andreasen CH, Dina C, Kottgen A, Le Bacquer O, Pattou F, Taneera J, Steinthorsdottir V, Rybin D, Ardlie K, Sampson M, Qi L, van Hoek M, Weedon MN, Aulchenko YS, Voight BF, Grallert H, Balkau B, Bergman RN, Bielinski SJ, Bonnefond A, Bonnycastle LL, Borch-Johnsen K, Bottcher Y, Brunner E, Buchanan TA, Bumpstead SJ, Cavalcanti-Proenca C, Charpentier G, Chen YD, Chines PS, Collins FS, Cornelis M, G JC, Delplanque J, Doney A, Egan JM, Erdos MR, Firmann M, Forouhi NG, Fox CS, Goodarzi MO, Graessler J, Hingorani A, Isomaa B, Jorgensen T, Kivimaki M, Kovacs P, Krohn K, Kumari M, Lauritzen T, Levy-Marchal C, Mayor V, McAteer JB, Meyre D, Mitchell BD, Mohlke KL, Morken MA, Narisu N, Palmer CN, Pakyz R, Pascoe L, Payne F, Pearson D, Rathmann W, Sandbaek A, Sayer AA, Scott LJ, Sharp SJ, Sijbrands E, Singleton A, Siscovick DS, Smith NL, Sparso T, Swift AJ, Syddall H, Thorleifsson G, Tonjes A, Tuomi T, Tuomilehto J, Valle TT, Waeber G, Walley A, Waterworth DM, Zeggini E, Zhao JH, Illig T, Wichmann HE, Wilson JF, van Duijn C, Hu FB, Morris AD, Frayling TM, Hattersley AT, Thorsteinsdottir U, Stefansson K, Nilsson P, Syvanen AC, Shuldiner AR, Walker M, Bornstein SR, Schwarz P, Williams GH, Nathan DM, Kuusisto J, Laakso M, Cooper C, Marmot M, Ferrucci L, Mooser V, Stumvoll M, Loos RJ, Altshuler D, Psaty BM, Rotter JI, Boerwinkle E, Hansen T, Pedersen O, Florez JC, McCarthy MI, Boehnke M, Barroso I, Sladek R, Froguel P, Meigs JB, Groop L, Wareham NJ, Watanabe RM. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. *Nature Genetics* 2010;42(2):142-8.
327. Wu Y, Li H, Loos RJF, Yu Z, Ye X, Chen L, Pan A, Hu FB, Lin X. Common variants in CDKAL1, CDKN2A/B, IGF2BP2, SLC30A8, and HHEX/IDE genes are associated with type 2 diabetes and impaired fasting glucose in a Chinese Han population. *Diabetes* 2008;57(10):2834-42.
328. Armitage P, Berry G. Sampling. *Statistical Methods in Medical Research*. Oxford: Blackwell Science; 1995. p. 78-92.
329. Burgess S. Online Sample Size and Power Calculator for Mendelian Randomization With a Binary Outcome [1st June 2015]. Available from: <http://spark.rstudio.com/sb452/power/>.

330. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *International Journal of Epidemiology* 2014;43(3):922-9.
331. Ahmad OS, Morris JA, Mujammami M, Forgetta V, Leong A, Li R, Turgeon M, Greenwood CMT, Thanassoulis G, Meigs JB, Sladek R, Richards JB. A Mendelian randomization study of the effect of type-2 diabetes on coronary heart disease. *Nature Communications* 2015;6:doi 10.1038/ncomms8060.
332. Hannou SA, Wouters K, Paumelle R, Staels B. Functional genomics of the CDKN2A/B locus in cardiovascular and metabolic disease: what have we learned from GWASs? *Trends in Endocrinology and Metabolism* 2015;26(4):176-84.
333. Johansen CT, Wang J, Lanktree MB, Cao H, McIntyre AD, Ban MR, Martins RA, Kennedy BA, Hassell RG, Visser ME, Schwartz SM, Voight BF, Elosua R, Salomaa V, O'Donnell CJ, Dallinga-Thie GM, Anand SS, Yusuf S, Huff MW, Kathiresan S, Hegele RA. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nature Genetics* 2010;42(8):684-7.
334. Delgado-Lista J, Perez-Martinez P, García-Rios A, Phillips CM, Williams CM, Gulseth HL, Helal O, Blaak EE, Kiec-Wilk B, Basu S, Drevon CA, Defoort C, Saris WH, Wybranska I, Riserus U, Lovegrove JA, Roche HM, Lopez-Miranda J. Pleiotropic effects of TCF7L2 gene variants and its modulation in the metabolic syndrome: From the LIPGENE study. *Atherosclerosis* 2011;214(1):110-6.
335. 1000 Genomes Project. Initial Phase 3 variant list and phased genotypes. In: Project G, editor. 2015.
336. Ng MCY, Park KS, Oh B, Tam CHT, Cho YM, Shin HD, Lam VKL, Ma RCW, So WY, Cho YS, Kim H-L, Lee HK, Chan JCN, Cho NH. Implication of genetic variants near TCF7L2, SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, and FTO in type 2 diabetes and obesity in 6,719 Asians. *Diabetes* 2008;57(8):2226-33.
337. Hu C, Zhang R, Wang C, Wang J, Ma X, Hou X, Lu J, Yu W, Jiang F, Bao Y, Xiang K, Jia W. Variants from GIPR, TCF7L2, DGKB, MADD, CRY2, GLIS3, PROX1, SLC30A8 and IGF1 are associated with glucose metabolism in the Chinese. *PLoS ONE* 2010;5(11):e15542.
338. Hu C, Zhang R, Wang C, Yu W, Lu J, Ma X, Wang J, Jiang F, Tang S, Bao Y, Xiang K, Jia W. Effects of GCK, GCKR, G6PC2 and MTNR1B variants on glucose metabolism and insulin secretion. *PLoS ONE* 2010;5(7):e11761.
339. Rodriguez S, Gaunt TR, Day IN. Hardy-Weinberg equilibrium testing of biological ascertainment for Mendelian randomization studies. *American Journal of Epidemiology* 2009;169(4):505-14.
340. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Statistics in Medicine* 2008;27(8):1133-63.

341. Bradford Hill A. The Environment and Disease: Association or Causation? *Proceedings of the Royal Society of Medicine* 1965;58(5):295-300.
342. Emerging Risk Factors Collaboration. Diabetes Mellitus, Fasting Glucose, and Risk of Cause-Specific Death. *New England Journal of Medicine* 2011;364(9):829-41.
343. Hofer M. The Bradford Hill considerations on causality: a counterfactual perspective. *Emerging Themes in Epidemiology* 2005;2:11.
344. Lucas RM, McMichael AJ. Association or causation: evaluating links between "environment and disease". *Bulletin of the World Health Organization* 2005;83(10):792-5.
345. Hopper I, Billah B, Skiba M, Krum H. Prevention of diabetes and reduction in major cardiovascular events in studies of subjects with prediabetes: Meta-analysis of randomised controlled clinical trials. *European Journal of Cardiovascular Prevention and Rehabilitation* 2011;18(6):813-23.
346. Li G, Zhang P, Wang J, An Y, Gong Q, Gregg EW, Yang W, Zhang B, Shuai Y, Hong J, Engelgau MM, Li H, Roglic G, Hu Y, Bennett PH. Cardiovascular mortality, all-cause mortality, and diabetes incidence after lifestyle intervention for people with impaired glucose tolerance in the Da Qing Diabetes Prevention Study: a 23-year follow-up study. *Lancet Diabetes and Endocrinology* 2014;2(6):474-80.
347. Diabetes Prevention Program Research Group. Long-term effects of lifestyle intervention or metformin on diabetes development and microvascular complications over 15-year follow-up: the Diabetes Prevention Program Outcomes Study. *Lancet Diabetes & Endocrinology* 2015;3(11):866-75.
348. The Diabetes Prevention Program Outcomes Study Research Group, Orchard TJ, Temprosa M, Barrett-Connor E, Fowler SE, Goldberg RB, Mather KJ, Marcovina SM, Montez M, Ratner RE, Saudek CD, Sherif H, Watson KE. Long-term effects of the Diabetes Prevention Program interventions on cardiovascular risk factors: a report from the DPP Outcomes Study. *Diabetic Medicine* 2013;30(1):46-55.
349. Jansen H, Loley C, Lieb W, Pencina MJ, Nelson CP, Kathiresan S, Peloso GM, Voight BF, Reilly MP, Assimes TL, Boerwinkle E, Hengstenberg C, Laaksonen R, McPherson R, Roberts R, Thorsteinsdottir U, Peters A, Gieger C, Rawal R, Thompson JR, König IR, Vasan RS, Erdmann J, Samani NJ, Schunkert H. Genetic variants primarily associated with type 2 diabetes are related to coronary artery disease risk. *Atherosclerosis* 2015;241(2):419-26.
350. Nohr EA, Olsen J. Commentary: Epidemiologists have debated representativeness for more than 40 years—has the time come to move on? *International Journal of Epidemiology* 2013;42(4):1016-7.
351. Richiardi L, Pizzi C, Pearce N. Commentary: Representativeness is usually not necessary and often should be avoided. *International Journal of Epidemiology* 2013;42(4):1018-22.

352. Ebrahim S, Davey Smith G. Commentary: Should we always deliberately be non-representative? *International Journal of Epidemiology* 2013;42(4):1022-6.
353. Elwood JM. Commentary: On representativeness. *International Journal of Epidemiology* 2013;42(4):1014-5.
354. Chen Z, Peto R, Zhou M, Iona A, Smith M, Yang L, Guo Y, Chen Y, Bian Z, Lancaster G, Sherliker P, Pang S, Wang H, Su H, Wu M, Wu X, Chen J, Collins R, Li L. Contrasting male and female trends in tobacco-attributed mortality in China: evidence from successive nationwide prospective cohort studies. *Lancet* 2015;386(10002):1447-56.
355. Rose G. Sick individuals and sick populations. *International Journal of Epidemiology* 2001;30(3):427-32.
356. Bezerra DC, Sharrett AR, Matsushita K, Gottesman RF, Shibata D, Mosley Jr TH, Coresh J, Szklo M, Carvalho MS, Selvin E. Risk factors for lacune subtypes in the Atherosclerosis Risk in Communities (ARIC) Study. *Neurology* 2012;78(2):102-8.
357. Vigneri P, Frasca F, Sciacca L, Pandini G, Vigneri R. Diabetes and cancer. *Endocrine-Related Cancer* 2009;16(4):1103-23.

Appendix A. List of papers authored during DPhil

1. **Bragg F.**, Li L., Smith M., Guo Y., Chen Y., Millwood I., Bian Z., Walters R., Chen J., Yang L., Collins R., Peto R., Lu Y., Yu B., Xie X., Lei Y., Luo G., Chen Z., on behalf of the China Kadoorie Biobank collaborative group. Associations of blood glucose and prevalent diabetes with risk of cardiovascular disease in 500 000 adult Chinese: the China Kadoorie Biobank. *Diabetic Medicine* 2014; 31(5): 540-51.
2. Lewington S., **Bragg F.**, Clarke R., A review on metaanalysis of biomarkers: promises and pitfalls. *Clinical Chemistry* 2012;58(8):1192-204.
3. **Bragg F.**, Li L., Yang L., Guo Y., Chen Y., Bian Z., Chen J., Collins R., Peto R., Wang C., Dong C., Pan R., Zhou J., Xu X., Chen Z., Emerging burden of diabetes-related cardiovascular disease in China: a 7-year prospective study of 0.5 million adults. (Submitted)
4. **Bragg F.**, Li L., Bennett D., Guo Y., Lewington S., Bian Z., Yang L., Chen J., Chen Y., Zhu B., Yin J., Hu X., Zhou L., Pan Y., Collins R., Peto R., Chen Z., Associations of plasma glucose with risk of cardiovascular disease among 467 000 Chinese adults without diabetes: a 7-year prospective study. (Submitted)
5. **Bragg F.***, Holmes M.V.*, Li L., Iona A., Guo Y., Yang L., Bian Z., Chen Y., Chen J., Collins R., Peto R., Chen Z., Diabetes and cause-specific mortality in China: analysis of 0.5 million adults. (*Joint first authors) (In preparation)

Appendix B. Additional China Kadoorie Biobank methods

Table B.1. Type 2 diabetes-related SNPs genotyped in the CKB

SNP ID	Chromosome	Coordinate	Reference allele	Other allele	Nearby genes
rs10010131	4	6292915	A	G	WFS1
rs10811661	9	22134094	T	C	CDKN2A/B
rs10830963	11	92708710	C	G	MTNR1B
rs10906115	10	12314997	A	G	CDC123, CAMK1D
rs10923931	1	120517959	G	T	NOTCH2-ADAM30
rs1111875	10	94462882	C	T	HHEX
rs11634397	15	80432222	A	G	ZFAND6
rs1169303	12	121436376	A	C	HNF1A
rs11708067	3	123065778	A	G	ADCY5
rs13266634	8	118184783	C	T	SLC30A8
rs13292136	9	81952128	C	T	CHCHD9
rs1359790	13	80717156	G	A	SPRY2
rs1470579	3	185529080	A	C	IGF2BP2
rs1531343	12	66174894	G	C	HMG2A
rs1535500	6	39284050	G	T	KCNK16
rs1552224	11	72433098	A	C	CENTD2
rs16861329	3	186666461	C	T	ST6GAL1
rs17584499	9	8879118	C	T	PTPRD
rs1802295	10	70931474	C	T	VPS26A
rs2028299	15	90374257	C	A	AP3S2
rs2191349	7	15064309	G	T	DGKB
rs2237892	11	2839751	C	T	KCNQ1
rs243021	2	60584819	G	A	BCL11A
rs2943641	2	227093745	T	C	IRS1
rs340874	1	214159256	T	C	PROX1
rs3786897	19	33893008	A	G	PEPD
rs3923113	2	165501849	A	C	GRB14
rs4430796	17	36098040	G	A	HNF1B
rs4457053	5	76424949	G	A	ZBED3
rs4523957	17	2208899	G	T	SRR
rs4607103	3	64711904	C	T	ADAMTS9
rs4607517	7	44235668	G	A	GCK
rs4752781	11	1702055	A	T	INS/IGF2
rs5215	11	17408630	C	T	KCNJ11
rs5945326	X	152899922	A	G	DUSP9
rs6017317	20	42946966	T	G	FITM2-R3HDML- HNF4A
rs6467136	7	127164958	A	G	GCC1-PAX4
rs6780569	3	23198484	G	A	UBE2E2
rs6815464	4	1309901	C	G	MAEA
rs7041847	9	4287466	A	G	GLIS3
rs7172432	15	62396389	A	G	C2CD4A-C2CD4B
rs7178572	15	77747190	A	G	HMG20A

Table B.1 continued

SNP ID	Chromosome	Coordinate	Reference allele	Other allele	Nearby genes
rs7578597	2	43732823	T	C	THADA
rs7593730	2	161171454	T	C	RBMS1
rs7649970	3	12392272	C	T	PPARG
rs7754840	6	20661250	G	C	CDKAL1
rs780094	2	27741237	T	C	GCKR
rs7901695	10	114754088	T	C	TCF7L2
rs7961581	12	71663102	C	T	TSPAN8-LGR5
rs8042680	15	91521337	C	A	PRC1
rs831571	3	64048297	T	C	PSMD6
rs864745	7	28180556	T	C	JAZF1
rs896854	8	95960511	T	C	TP53INP1
rs9470794	6	38106844	T	C	ZFAND3
rs972283	7	130466854	A	G	KLF14
rs9939609	16	53820527	T	A	FTO

Appendix C. Additional methods and results: Intra-individual variation in plasma glucose levels

Table C1. Age and sex-specific regression dilution ratios based on baseline and first resurvey in participants without self-reported diabetes at baseline or first resurvey, estimated by different methods

Age at baseline, years	No. of participants	Mean baseline to resurvey interval (years)	Regression dilution ratio		
			MacMahon's	Rosner's	Self-correlation ^a
Men					
30-59	5124	2.6	0.58	0.42	0.22
60-69	1332	2.6	0.72	0.56	0.25
70-79	422	2.6	0.52	0.52	0.28
All	6878	2.6	0.61	0.47	0.24
Women					
30-59	8608	2.6	0.67	0.54	0.28
60-69	1626	2.6	0.64	0.56	0.27
70-79	437	2.6	0.65	0.56	0.34
All	10671	2.6	0.66	0.55	0.29
Total- unadjusted	17549	2.6	0.64	0.52	0.27
Total- age- and sex-adjusted	17549	2.6	N/A	0.50	0.26

^aCalculated using Spearman's rank correlation coefficient.

Appendix Note C.1. Intra-individual variation in plasma glucose levels stratified by baseline and resurvey fasting period

Background

Repeat RPG measurements on a randomly selected sample of the CKB study population were used to estimate the extent of intra-individual variation in RPG levels. Intra-individual variation in exposures typically results from a range of factors, including random measurement error, short-term biologic variability and longer-term systematic changes. In the context of RPG, short-term biologic variability includes the effects of fasting period, as demonstrated by the inverse association between fasting period and plasma glucose (Section 4.3.3). Varying self-reported fasting periods within individuals at baseline and resurvey in the CKB might therefore contribute to the extent of within-person variation in plasma glucose levels. The aims of these analyses were to gain a better understanding of the influence of fasting period on intra-individual variation in plasma glucose levels, and to investigate the feasibility of accounting for within-person variation in fasting time when adjusting for regression dilution bias.

Methods

Analyses investigating intra-individual variation in plasma glucose levels using MacMahon's method, as described in Section 4.2.4, were repeated in four population subgroups defined by strata of fasting period at baseline and first resurvey; <8 hour fasting period at baseline and first resurvey, ≥8 hour fasting period at baseline and first resurvey, <8 hour fasting period at baseline and ≥8 hour fasting period at first resurvey, and ≥8 hour fasting period at baseline and <8 hour fasting period at first resurvey. These analyses were based on first, but not second, resurvey data as the main analyses examining intra-individual variation in RPG levels were based on first resurvey data only (Sections 4.2.4 and 4.3.5), and stratifying by baseline, first resurvey and second resurvey fasting times would add an unnecessary level of complexity.

Results

The majority of participants included in these analyses (71.7%) fasted for <8 hours at both baseline and first resurvey (Table C.2), with between 8.5% and 10.2% of participants in the remaining three subgroups. There was a modest increase in mean plasma glucose level between baseline and first resurvey in participants with a fasting period <8 hours at both baseline and resurvey, from 5.8 mmol/L to 6.4 mmol/L. In those with a fasting period ≥8 hours at both baseline and resurvey the mean plasma glucose remained constant between surveys (5.3 mmol/L). Among individuals with a fasting period <8 hours at baseline and ≥8 hours at resurvey there was a modestly lower mean plasma glucose at resurvey (5.6 mmol/L) than at baseline (6.1 mmol/L), and an opposite trend was seen in participants with a baseline fasting period ≥8 hours and resurvey fasting period <8 hours in whom there was a marked increase in mean plasma glucose from 5.3 mmol/L to 6.6 mmol/L. Regression dilution ratios varied between the four population subgroups. The extremes of this variation were among individuals who reported a fasting period ≥8 hours at baseline and <8 hours at resurvey, in whom the regression dilution ratio was highest (0.81) and in those with a fasting period <8 hours at baseline and ≥8 hours at resurvey, in whom the calculated regression dilution ratio was 0.35.

Discussion

Regression dilution ratios closer to zero imply greater intra-individual variation in plasma glucose levels. Perhaps unsurprisingly, the lowest regression dilution ratio, indicating the greatest level of intra-individual variation, was found in individuals with a baseline fasting period <8 hours but a resurvey fasting period ≥8 hours, whilst the highest regression dilution ratio was in individuals with opposite fasting period patterns, that is a baseline fasting period ≥8 hours and resurvey fasting period <8 hours.

Addressing the second aim of these analyses, it would be theoretically possible to apply regression dilution ratios of 0.35 and 0.45 to plasma glucose-associated CVD risk estimates in individuals with baseline fasting periods of <8 hours and ≥8 hours, respectively, in order

to estimate CVD risks associated with incremental changes in usual FPG levels. This would be a theoretically attractive approach to addressing varying fasting periods and the potential influence of this on plasma glucose levels, on intra-individual variation in plasma glucose levels and on the association of plasma glucose with CVD. Unfortunately, however, the relatively small proportion of participants with a baseline fasting period ≥ 8 hours (21.2%) produced imprecise CVD risk estimates with less power to identify any association (Section 5.3.3). Furthermore, the regression dilution ratio estimates for the population subgroups with a resurvey fasting period ≥ 8 hours were based on small numbers of participants, with greater susceptibility to error and imprecision.

Conclusion

In summary, these analyses show differences in the extent of within-person variation in plasma glucose levels according to fasting time in a manner which would be expected given the association between fasting period and plasma glucose levels. The proposed application of this understanding, to account for within-person variation in fasting time when adjusting for regression dilution bias, was not feasible, in practice, in particular due to the relatively small proportion of study participants with a baseline fasting period ≥ 8 hours.

Table C2. Mean baseline and resurvey plasma glucose levels and regression dilution ratios by baseline and resurvey fasting period in participants without self-reported diabetes at baseline

Baseline fasting period (hours)	Resurvey fasting period							
	<8 hours				≥8 hours			
	n (%)	Baseline plasma glucose (mmol/L)	Resurvey plasma glucose (mmol/L)	Regression dilution ratio	n (%)	Baseline plasma glucose (mmol/L)	Resurvey plasma glucose (mmol/L)	Regression dilution ratio
<8	12813 (71.7)	5.8	6.4	0.60	1514 (8.5)	6.1	5.6	0.35
≥8	1829 (10.2)	5.3	6.6	0.81	1707 (9.6)	5.3	5.3	0.45

Appendix D. Additional results: Diabetes, plasma glucose and prevalent cardiovascular disease

Figure D.1. Causal diagram demonstrating the relationships of diabetes and plasma glucose with cardiovascular disease

Causal relationships were identified through review of relevant literature.

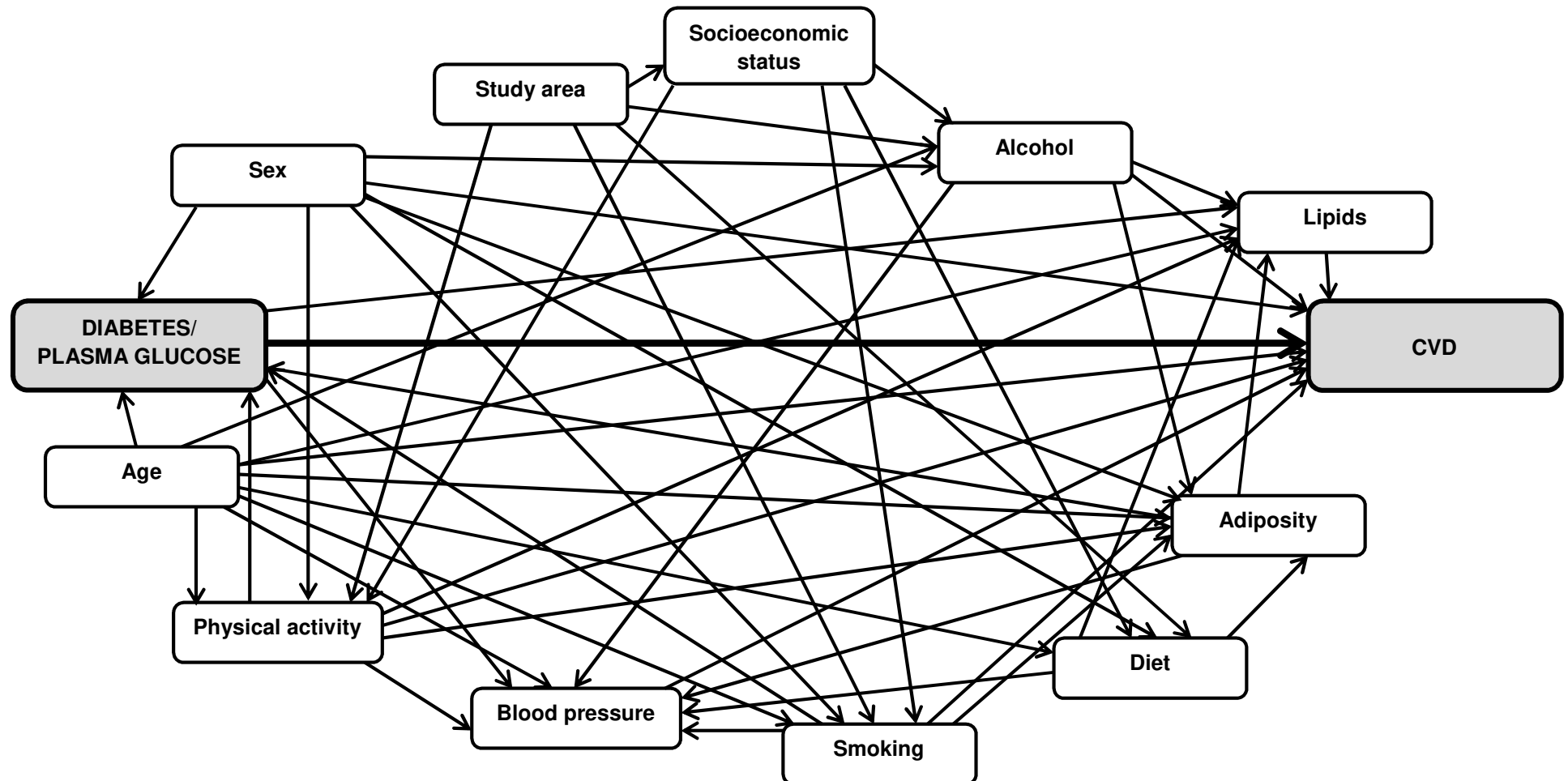


Table D.1. Association of baseline ischaemic heart disease prevalence with lifestyle characteristics

Lifestyle characteristic	No. of events	OR for ischaemic heart disease		
		Minimal adjustment ^a	Intermediate adjustment	Full adjustment ^b
Education^c				
No formal schooling	2384	1.00	-	1.00
Primary school	4938	1.30	-	1.31
Middle school	3791	1.52	-	1.53
High school	2638	1.74	-	1.78
College or university	1661	1.93	-	1.96
Smoking^d				
Never	10036	1.00	1.00	1.00
Occasional	686	0.96	1.00	1.02
Ex-regular	1898	1.49	1.44	1.41
Current regular	2792	0.90	0.94	1.05
Alcohol^e				
Never regular	7810	1.00	1.00	1.00
Occasional	645	1.99	1.85	1.75
Ex-regular	4549	0.72	0.72	0.75
Reduced intake	314	0.58	0.58	0.61
Monthly	609	1.31	1.27	1.18
Weekly	1485	0.61	0.61	0.63
Physical activity (MET hours/day)				
<10	2567	1.00	-	1.00
10-19.9	9003	0.92	-	0.95
20-29.9	2497	0.70	-	0.74
30-39.9	830	0.56	-	0.63
≥40	515	0.47	-	0.53

^aAdjusted for age, sex, education and study area; ^bAdditionally adjusted for alcohol, smoking, systolic blood pressure, physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, meat consumption, dairy product consumption, eggs consumption, wheat consumption, tea consumption; ^cAdjusted for variables as above excluding education; ^dIntermediate adjustment adjusted for age, sex, education, study area and alcohol consumption; ^eIntermediate adjustment adjusted for age, sex, education, study area and smoking.
MET, metabolic equivalent of task; OR, odds ratio.

Table D.2. Association of baseline ischaemic heart disease prevalence with dietary consumption

Dietary consumption	No. of events	OR for ischaemic heart disease	
		Minimal adjustment ^a	Full adjustment ^b
Fresh fruit			
Rarely/never	1094	1.00	1.00
Monthly	3694	0.86	0.90
1-3 days/week	4081	0.88	0.88
4-6 days/week	1288	0.94	0.90
Daily	5255	0.93	0.88
Dairy products			
Rarely/never	7058	1.00	1.00
Monthly	1938	1.18	1.22
1-3 days/week	1945	1.23	1.28
4-6 days/week	520	1.40	1.49
Daily	3951	1.44	1.50
Eggs			
Rarely/never	1270	1.00	1.00
Monthly	2127	0.80	0.83
1-3 days/week	6577	0.73	0.77
4-6 days/week	1598	0.75	0.77
Daily	3840	0.83	0.84
Meat			
Rarely/never	891	1.00	1.00
Monthly	1809	0.93	1.02
1-3 days/week	5162	0.91	1.00
4-6 days/week	2019	0.86	0.93
Daily	5531	0.84	0.90
Wheat			
Rarely/never	675	1.00	1.00
Monthly	1453	1.02	1.02
1-3 days/week	3580	1.21	1.17
4-6 days/week	1382	1.36	1.29
Daily	8322	1.26	1.18
Tea			
Never or almost never	6823	1.00	1.00
Occasionally	3243	0.89	0.92
Only at certain seasons	308	0.82	0.82
Monthly but less than weekly	504	0.86	0.86
At least once a week	4534	0.84	0.85

^aAdjusted for age, sex, education and study area; ^bAdditionally adjusted for alcohol, smoking, systolic blood pressure, physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, meat consumption, dairy product consumption, eggs consumption, wheat consumption, tea consumption.

OR, odds ratio.

Figure D.2. Association of baseline ischaemic heart disease prevalence with adiposity

Minimal adjustment: adjusted for age sex, study area and education; Intermediate adjustment: BMI additionally adjusted for waist circumference and hip circumference; waist circumference additionally adjusted for hip circumference and BMI; hip circumference additionally adjusted for waist circumference and BMI; waist-to-hip ratio additionally adjusted for BMI; body fat percent additionally adjusted for BMI, waist circumference and hip circumference.

Squares represent the odds ratio with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

Data for body fat percent missing for 236 participants. BMI, body mass index.

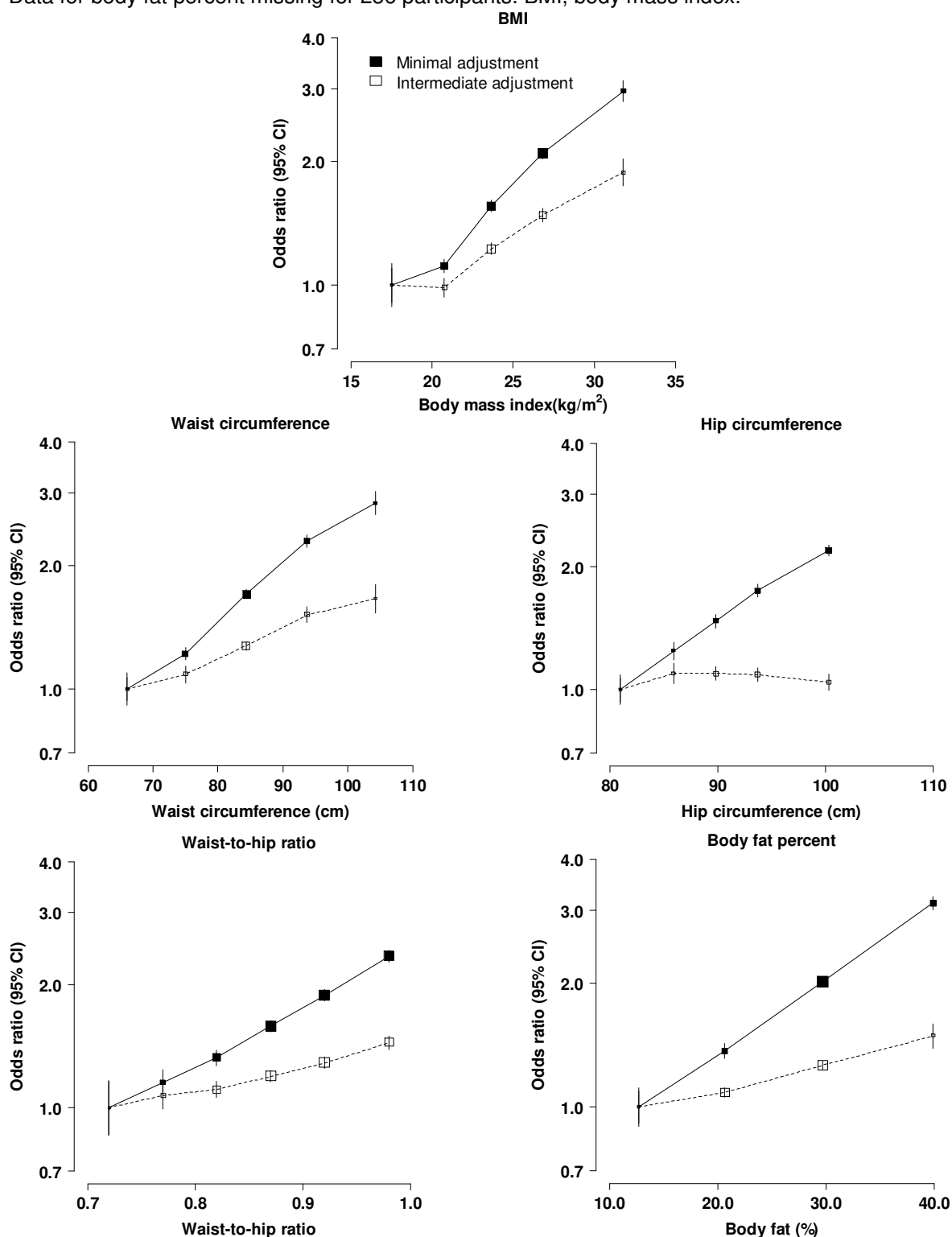


Figure D.3. Association of baseline ischaemic heart disease prevalence with blood pressure

Minimal adjustment: age, sex, education and study area; Full adjustment: additionally adjusted for alcohol, smoking, systolic blood pressure (for diastolic blood pressure analyses only), physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, meat consumption, dairy product consumption, egg consumption, wheat consumption, tea consumption.

Squares represent the odds ratio with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

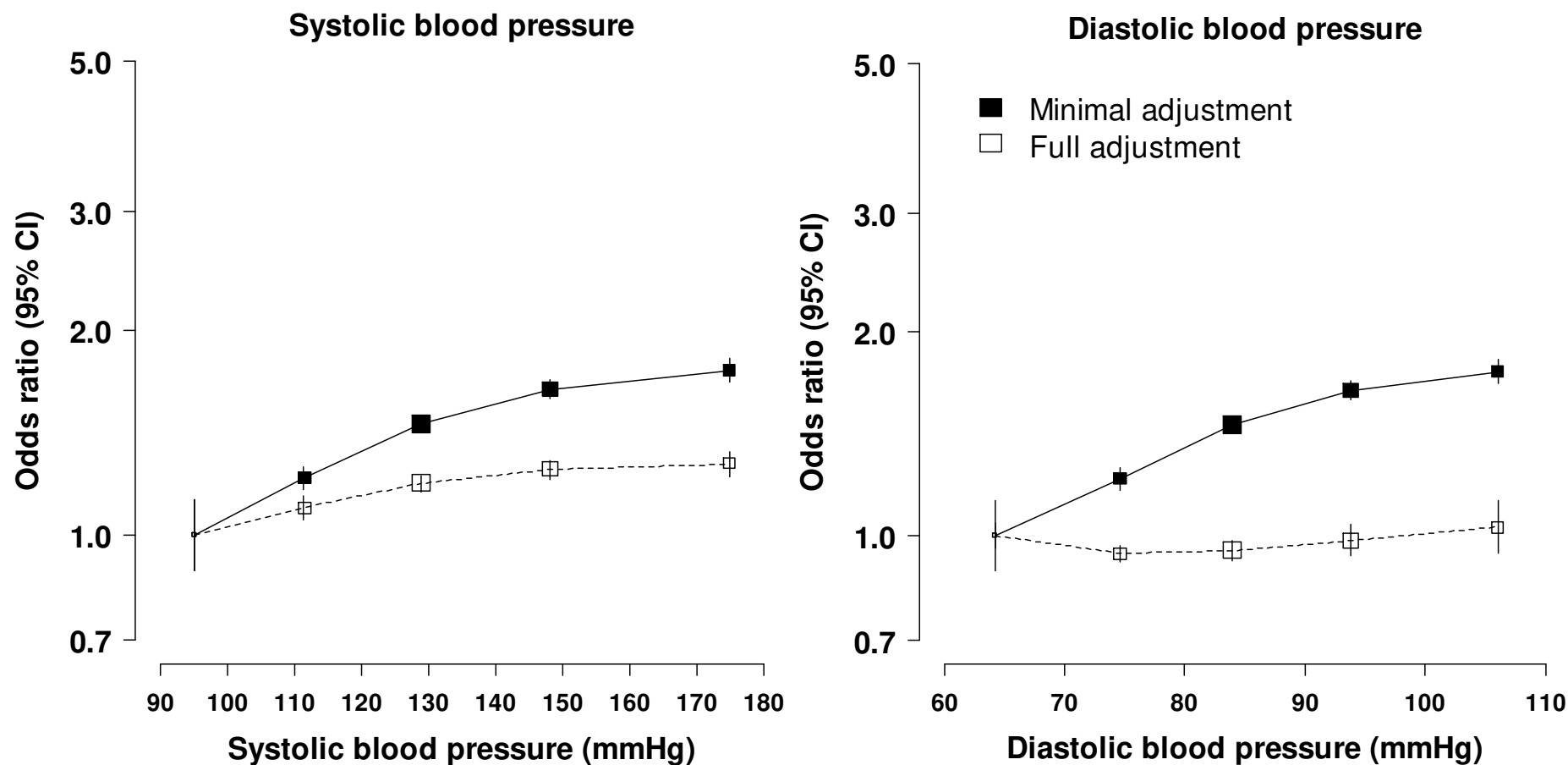


Table D.3. Association of baseline stroke/transient ischaemic attack prevalence with lifestyle characteristics

Lifestyle characteristic	No. of events	OR for stroke/TIA		
		Minimal adjustment ^a	Intermediate adjustment	Full adjustment ^b
Education^c				
No formal schooling	1515	1.00	-	1.00
Primary school	3039	1.06	-	1.17
Middle school	2319	1.13	-	1.32
High school	1290	1.01	-	1.24
College or university	697	0.84	-	1.11
Smoking^d				
Never	4441	1.00	1.00	1.00
Occasional	517	0.88	0.96	1.01
Ex-regular	1611	1.66	1.52	1.52
Current regular	2291	0.91	0.98	1.13
Alcohol^e				
Never regular	4007	1.00	1.00	1.00
Occasional	753	2.99	2.74	2.56
Ex-regular	2392	0.53	0.54	0.59
Reduced intake	172	0.35	0.35	0.39
Monthly	568	1.27	1.23	1.20
Weekly	968	0.44	0.44	0.46
Physical activity (MET hours/day)				
<10	2415	1.00	-	1.00
10-19.9	4511	0.55	-	0.58
20-29.9	1221	0.35	-	0.40
30-39.9	402	0.24	-	0.29
≥40	311	0.20	-	0.24

^aAdjusted for age, sex, education and study area; ^bAdditionally adjusted for alcohol, smoking, systolic blood pressure, physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, meat consumption, dairy product consumption, eggs consumption, wheat consumption, tea consumption; ^cAdjusted for variables as above excluding education; ^dIntermediate adjustment: adjusted for age, sex, education, study area and alcohol consumption; ^eIntermediate adjustment: adjusted for age, sex, education, study area and smoking.

MET, metabolic equivalent of task; OR, odds ratio; TIA, transient ischaemic attack.

Table D.4. Association of baseline stroke/transient ischaemic attack prevalence with dietary consumption

Dietary consumption	No. of events	OR for stroke/TIA	
		Minimal adjustment ^a	Full adjustment ^b
Fresh fruit			
Rarely/never	895	1.00	1.00
Monthly	2758	0.90	1.02
1-3 days/week	2401	0.84	0.93
4-6 days/week	578	0.72	0.81
Daily	2228	0.79	0.88
Dairy products			
Rarely/never	4857	1.00	1.00
Monthly	1137	1.16	1.24
1-3 days/week	926	1.08	1.19
4-6 days/week	222	1.10	1.28
Daily	1718	1.22	1.35
Eggs			
Rarely/never	963	1.00	1.00
Monthly	1456	0.73	0.82
1-3 days/week	4016	0.66	0.79
4-6 days/week	775	0.64	0.80
Daily	1650	0.65	0.78
Meat			
Rarely/never	792	1.00	1.00
Monthly	1352	0.75	0.83
1-3 days/week	3015	0.69	0.79
4-6 days/week	961	0.54	0.65
Daily	2740	0.61	0.70
Wheat			
Rarely/never	400	1.00	1.00
Monthly	893	0.97	0.98
1-3 days/week	2017	1.01	1.03
4-6 days/week	659	0.97	0.99
Daily	4891	1.12	1.13
Tea			
Never or almost never	4479	1.00	1.00
Occasionally	1842	0.80	0.90
Only at certain seasons	168	0.63	0.69
Monthly but less than weekly	287	0.82	0.91
At least once a week	2084	0.68	0.75

^aAdjusted for age, sex, education and study area; ^bAdditionally adjusted for alcohol, smoking, systolic blood pressure, physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, meat consumption, dairy product consumption, eggs consumption, wheat consumption, tea consumption.

OR, odds ratio; TIA, transient ischaemic attack.

Figure D.4. Association of baseline stroke/transient ischaemic attack prevalence with adiposity

Minimal adjustment: adjusted for age, sex, study area and education; Intermediate adjustment: BMI additionally adjusted for waist circumference and hip circumference; waist circumference additionally adjusted for hip circumference and BMI; hip circumference additionally adjusted for waist circumference and BMI; waist-to-hip ratio additionally adjusted for BMI; body fat percent additionally adjusted for BMI, waist circumference and hip circumference.

Squares represent the odds ratio with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

Data for body fat percent missing for 236 participants. BMI, body mass index.

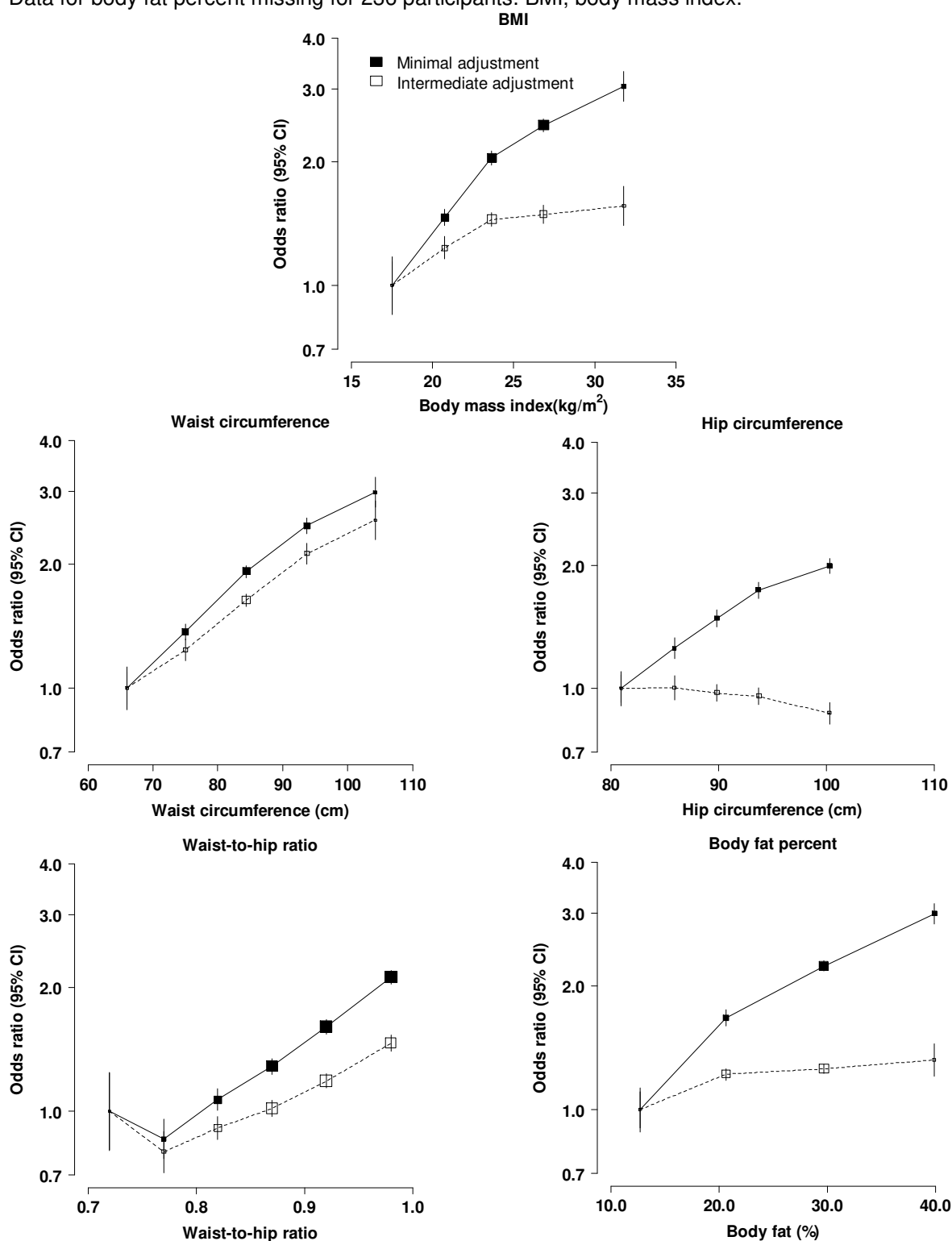


Figure D.5. Association of baseline stroke/transient ischaemic attack prevalence with blood pressure

Minimal adjustment: age, sex, education and study area; Full adjustment: additionally adjusted for alcohol, smoking, systolic blood pressure (for diastolic blood pressure analyses only), physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, meat consumption, dairy product consumption, egg consumption, wheat consumption, tea consumption.

Squares represent the odds ratio with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

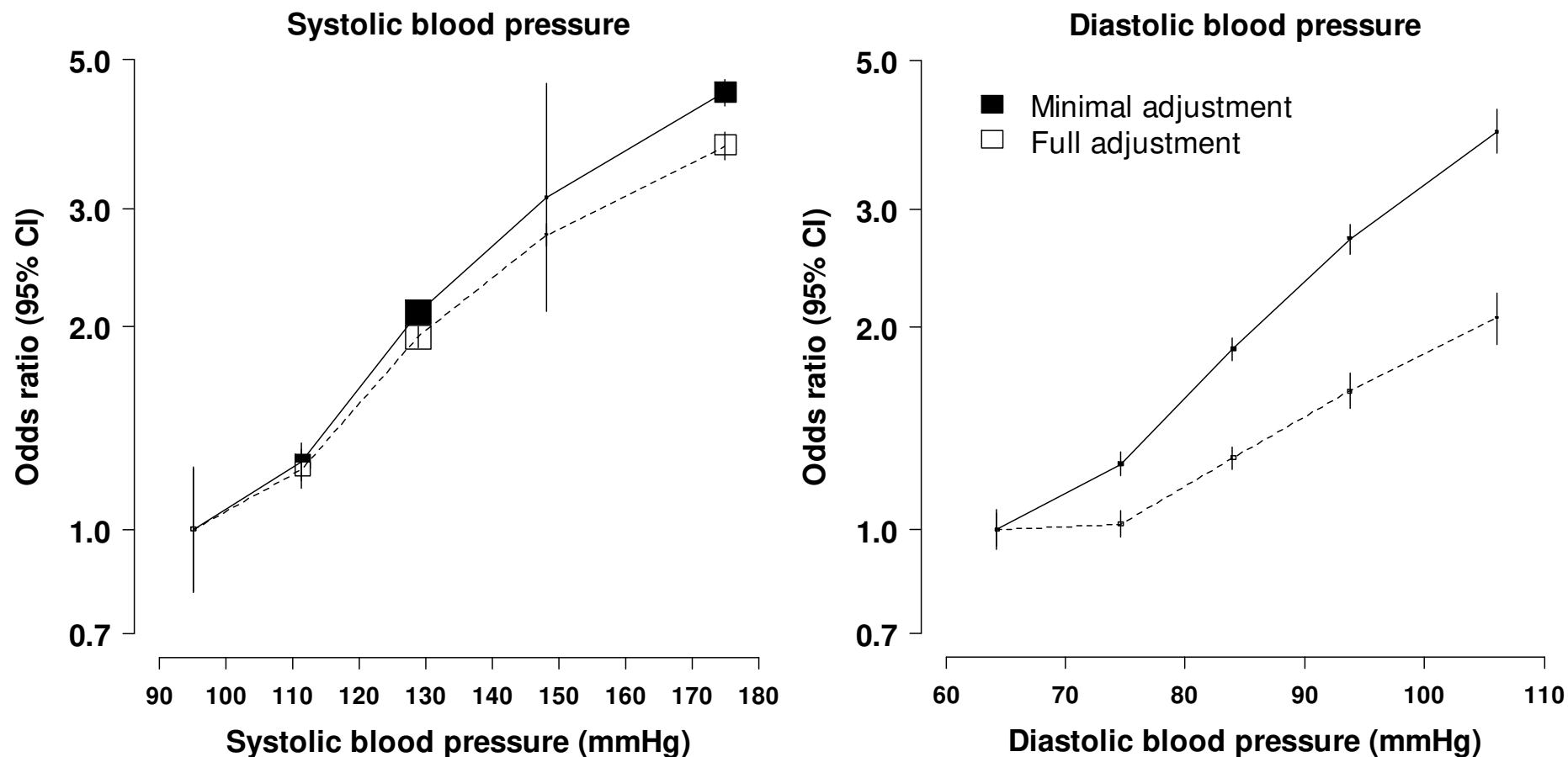
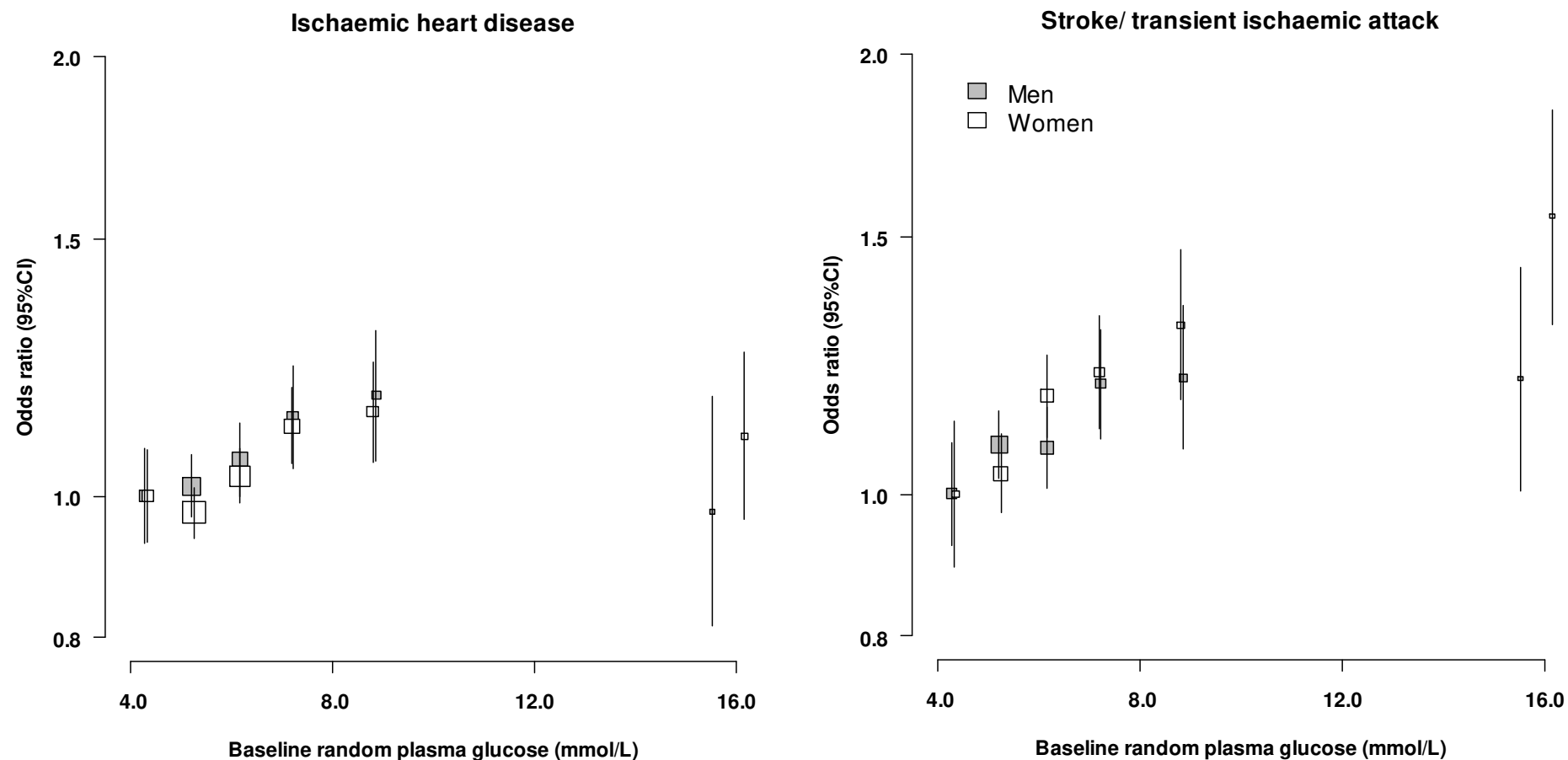


Figure D.6. Adjusted odds ratios for prevalent cardiovascular disease by baseline random plasma glucose in participants without self-reported diabetes

Adjusted for age, sex, study area, education, smoking, alcohol consumption, physical activity and systolic blood pressure.

Squares represent the odds ratios with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).



Appendix E. Additional methods and results: Diabetes and the risk of incident cardiovascular disease

Table E.1. Cardiovascular disease outcome definitions

Stroke outcomes

Outcome	Definition
First non-fatal total stroke	• I60, I61, I63, I64 in CDC morbidity or HI datasets with or without a CT or MR scan. <i>Additional conditions:</i> • No earlier I60, I61, I63 or I64 stroke event recorded with or without an associated scan in any dataset. • No death within 28 days from any cause.
First fatal total stroke	• I60, I61, I63, I64 in CDC mortality dataset with or without a record of a CT or MR scan or autopsy. <i>Or</i> • I60, I61, I63, I64 in CDC morbidity or HI datasets with or without a CT or MR scan with death from any cause within 28 days of the event. <i>Additional conditions:</i> • No earlier I60, I61, I63 or I64 stroke event recorded in HI or CDC morbidity dataset with or without an associated scan.
First fatal or non-fatal total stroke	• I60, I61, I63, I64 in CDC mortality dataset or in CDC morbidity or HI datasets with or without a CT or MR scan or autopsy. <i>Additional conditions:</i> • No earlier I60, I61, I63 or I64 stroke event recorded with or without an associated scan in any dataset.

Outcome	Definition
First non-fatal ischaemic stroke	• I63 in CDC morbidity or HI datasets with or without a CT or MR scan. <i>Additional conditions:</i> • No earlier I60, I61, I63 or I64 stroke event recorded with or without an associated scan in any dataset. • No death within 28 days from any cause. • No record of ICH (I61) or SAH (I60) within 28 days of the event in any dataset.
First fatal ischaemic stroke	• I63 in CDC mortality dataset with or without a record of a CT or MR scan or autopsy. <i>Or</i> • I63 in CDC morbidity or HI datasets with or without a CT or MR scan with death from any cause within 28 days of the event. <i>Additional conditions:</i> • No earlier I60, I61, I63 or I64 stroke event recorded in HI or CDC morbidity dataset with or without an associated scan.
First fatal or non-fatal ischaemic stroke	• I63 in CDC mortality dataset or in CDC morbidity or HI datasets with or without a CT or MR scan or autopsy. <i>Additional conditions:</i> • No earlier I60, I61, I63 or I64 stroke event recorded with or without an associated scan in any dataset. • No record of ICH (I61) or SAH (I60) within 28 days of the event in any dataset.

Outcome	Definition
First non-fatal intracerebral haemorrhage	• I61 in CDC morbidity or HI datasets with or without of CT or MR scan. <i>Additional conditions:</i> • No earlier I60, I61, I63 or I64 stroke event recorded with or without an associated scan in any dataset. • No death within 28 days from any cause. • No record of IS (I63) or SAH (I60) within 28 days of the event in any dataset.
First fatal intracerebral haemorrhage	• I61 in CDC mortality dataset with or without a record of a CT or MR scan or autopsy. <i>Or</i> • I61 in CDC morbidity or HI datasets with or without a CT or MR scan with death from any cause within 28 days of the event. <i>Additional conditions:</i> • No earlier I60, I61, I63 or I64 stroke event recorded in HI or CDC morbidity dataset with or without an associated scan.
First fatal or non-fatal intracerebral haemorrhage	• I61 in CDC mortality dataset or in CDC morbidity or HI datasets with or without a CT or MR scan or autopsy. <i>Additional conditions:</i> • No earlier I60, I61, I63 or I64 stroke event recorded with or without an associated scan in any dataset. • No record of IS (I63) or SAH (I60) within 28 days of the event in any dataset.

IHD outcomes

Outcome	Definition
First non-fatal MI	- I21-23 in CDC morbidity or HI datasets. <i>Additional conditions:</i> - No earlier MI (I21-I23) event recorded in any dataset. - No death within 28 days from any cause.
First fatal MI	- I21-23 in CDC mortality dataset <i>Or</i> - I21-23 in CDC morbidity or HI datasets with death from any cause within 28 days. <i>Additional conditions:</i> - No earlier MI (I21-I23) event recorded in any dataset.
First fatal or non-fatal MI	- I21-23 in CDC mortality, CDC morbidity or HI datasets. <i>Additional conditions:</i> - No earlier MI (I21-I23) event recorded in any dataset.
Major coronary event	- I21-23 in CDC morbidity or HI datasets. <i>Or</i> - I20-25 in CDC mortality dataset. <i>Or</i> - I20-25 in CDC morbidity or HI datasets with death from any cause within 28 days. <i>Additional conditions:</i> - No earlier MI (I21-I23) or fatal IHD (I20-I25) event recorded in any dataset.

Composite CVD outcomes

Outcome	Definition
Major occlusive vascular disease	• I21-23 or I61 in CDC morbidity or HI datasets. <i>Or</i> • I20-25 or I61 in CDC mortality dataset. <i>Or</i> • I20-25 or I61 in CDC morbidity or HI datasets with death from any cause within 28 days. <i>Additional conditions:</i> • No MI (I21-I23) or fatal IHD (I20-I25) event recorded in any dataset prior to included MI or IHD event. • No I60, I61, I63 or I64 stroke event recorded with or without an associated scan in any dataset prior to included IS event. • No record of ICH (I61) or SAH (I60) within 28 days of included IS event in any dataset.
Cardiovascular disease mortality	• I00-25, I27-88 or I95-99 in CDC mortality dataset <i>Or</i> • I00-25, I27-88 or I95-99 in CDC morbidity or HI datasets with death from any cause within 28 days.

Table E.2. Self-reported cardiovascular disease medication use in men and women with self-reported diabetes at baseline

Medication	Men^a, % (n=5046)	Women^b, % (n=8238)	p value
ACE-inhibitor	4.5	4.1	0.6
Aspirin	2.2	2.1	0.5
Beta-blocker	3.1	3.7	0.1
Calcium channel antagonist	7.8	8.9	0.001
Diuretic	1.0	1.2	0.2
Statin	0.9	0.9	0.5
Any of above	15.6	17.2	0.002

^aData missing for 779 participants. ^bData missing for 956 participants.
Standardised to the age and study area structure of the study population.
ACE, angiotensin converting enzyme.

Table E.3. Association of major occlusive vascular disease with lifestyle characteristics

Lifestyle characteristic	No. of events	HRs by different levels of adjustment		
		Minimal adjustment ^a	Intermediate adjustment	Full adjustment ^b
Education^c				
No formal schooling	5225	1.00	-	1.00
Primary school	8055	0.96	-	1.00
Middle school	5760	0.97	-	1.03
High school	3461	0.91	-	1.03
College or university	1669	0.82	-	0.99
Smoking^d				
Never	13663	1.00	1.00	1.00
Occasional	1365	1.01	1.04	1.04
Ex-regular	2164	1.12	1.11	1.08
Current regular	6978	1.20	1.22	1.27
Alcohol^e				
Never regular	10865	1.00	1.00	1.00
Occasional	672	1.35	1.32	1.27
Ex-regular	7521	0.83	0.82	0.86
Reduced intake	709	0.75	0.73	0.78
Monthly	756	1.08	1.05	1.01
Weekly	3647	0.88	0.85	0.83
Physical activity (MET hours/day)				
<10	10115	1.00	-	1.00
10-19.9	8358	0.85	-	0.88
20-29.9	2808	0.78	-	0.82
30-39.9	1694	0.74	-	0.79
≥40	1195	0.69	-	0.73

^aAdjusted for age, sex, education and study area; ^bAdditionally adjusted for alcohol, smoking, systolic blood pressure, physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, meat consumption, dairy product consumption, eggs consumption, wheat consumption, tea consumption; ^cAdjusted for variables as above excluding education; ^dIntermediate adjustment: adjusted for age, sex, education, study area and alcohol consumption; ^eIntermediate adjustment: adjusted for age, sex, education, study area and smoking.
HR, hazard ratio; MET, metabolic equivalent of task.

Table E.4. Association of major occlusive vascular disease with dietary consumption

Dietary consumption	No. of events	HRs by different levels of adjustment	
		Minimal adjustment ^a	Full adjustment ^b
Fresh fruit			
Rarely/never	2359	1.00	1.00
Monthly	8407	0.90	0.94
1-3 days/week	6685	0.84	0.89
4-6 days/week	1870	0.82	0.87
Daily	4849	0.76	0.81
Dairy products			
Rarely/never	14983	1.00	1.00
Monthly	3050	1.02	1.06
1-3 days/week	2232	1.00	1.07
4-6 days/week	609	1.06	1.17
Daily	3296	0.97	1.07
Eggs			
Rarely/never	2420	1.00	1.00
Monthly	4691	0.94	0.98
1-3 days/week	10880	0.90	0.97
4-6 days/week	2326	0.88	0.98
Daily	3853	0.82	0.92
Meat			
Rarely/never	1670	1.00	1.00
Monthly	3632	0.96	0.98
1-3 days/week	8000	0.92	0.96
4-6 days/week	3440	0.87	0.91
Daily	7428	0.96	1.00
Wheat			
Rarely/never	1509	1.00	1.00
Monthly	3087	1.00	1.01
1-3 days/week	5334	1.05	1.04
4-6 days/week	2013	1.07	1.06
Daily	12227	1.05	1.04
Tea			
Never or almost never	10448	1.00	1.00
Occasionally	5012	0.93	0.95
Only at certain seasons	586	0.95	0.96
Monthly but less than weekly	803	0.93	0.94
At least once a week	7321	0.98	0.96

^aAdjusted for age, sex, education and study area; ^bAdditionally adjusted for alcohol, smoking, systolic blood pressure, physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, meat consumption, dairy product consumption, eggs consumption, wheat consumption, tea consumption.

HR, hazard ratio.

Figure E.1. Association of major occlusive vascular disease with adiposity

Minimal adjustment: stratified by age, sex and study area and adjusted for education; Intermediate adjustment: BMI additionally adjusted for waist circumference and hip circumference; waist circumference additionally adjusted for hip circumference and BMI; hip circumference additionally adjusted for waist circumference and BMI; waist-to-hip ratio additionally adjusted for BMI; body fat percent additionally adjusted for BMI, waist circumference and hip circumference.

Squares represent the hazard ratio with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

Data for body fat percent missing for 236 participants. BMI, body mass index.

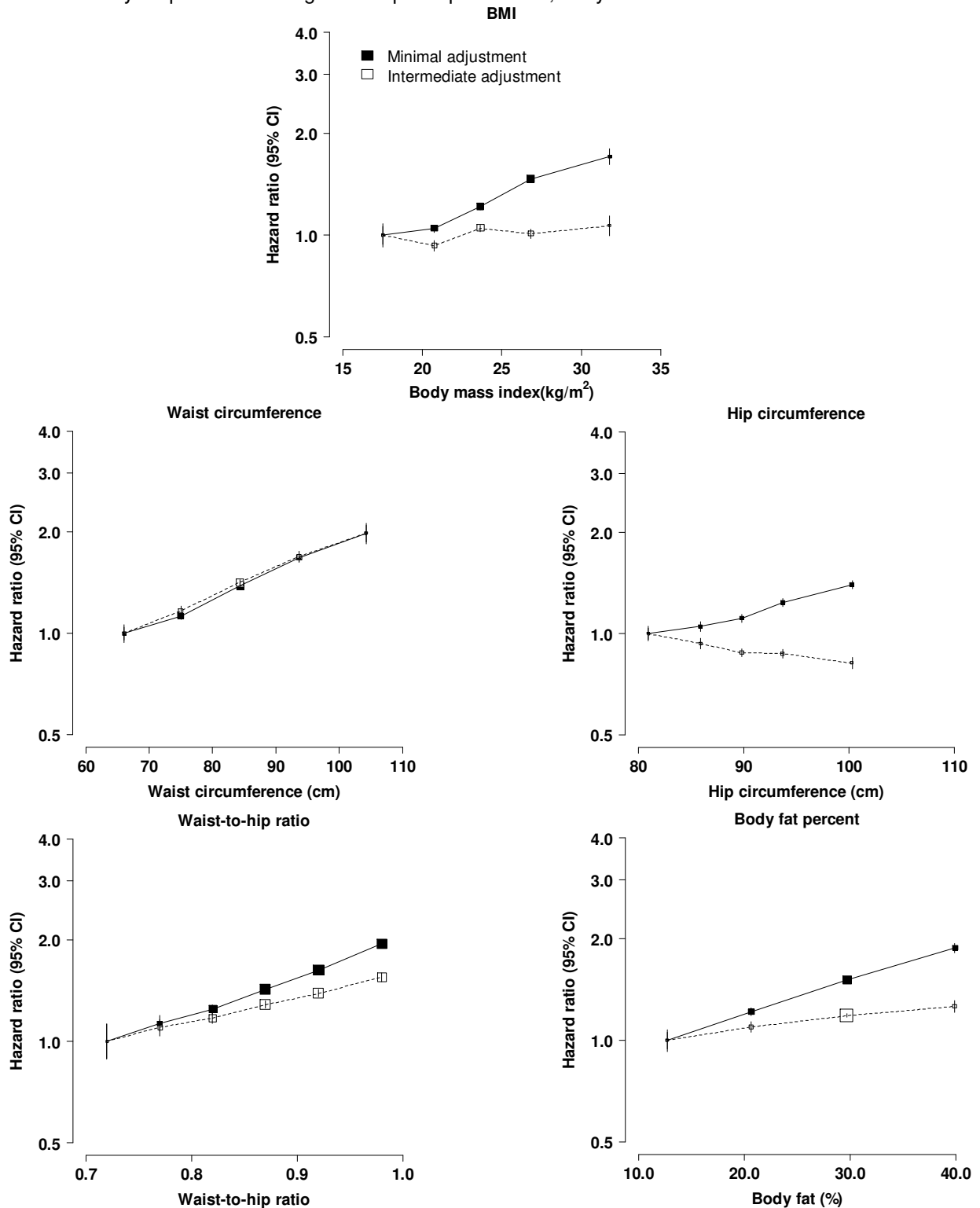


Figure E.2. Association of major occlusive vascular disease with blood pressure

Minimal adjustment: stratified by age, sex, education and study area; Full adjustment: additionally adjusted for alcohol, smoking, systolic blood pressure (for diastolic blood pressure analyses only), physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, meat consumption, dairy product consumption, eggs consumption, wheat consumption, tea consumption. Squares represent the hazard ratio with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

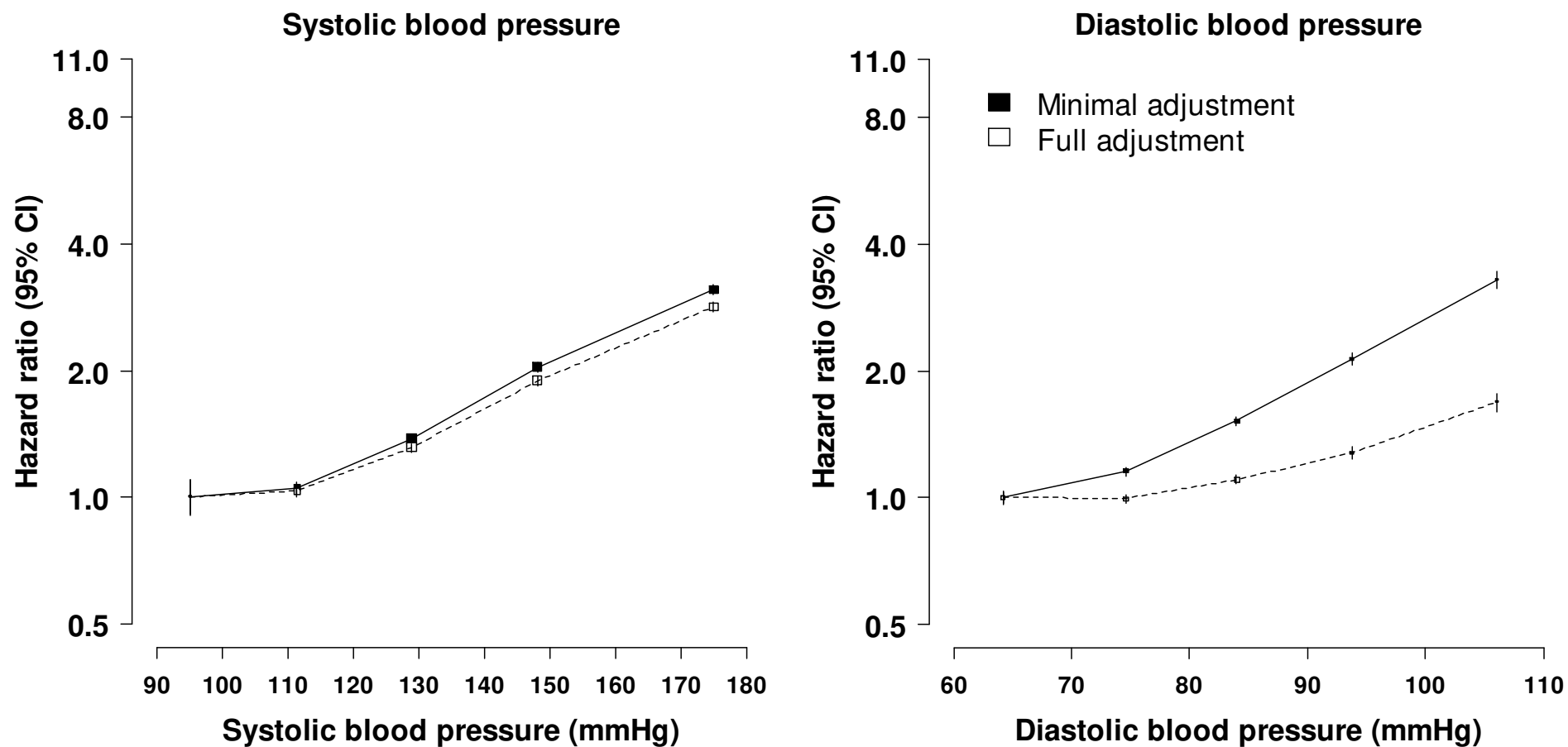


Table E.5. Association of intra-cerebral haemorrhage with lifestyle characteristics

Lifestyle characteristic	No. of events	HRs by different levels of adjustment		
		Minimal adjustment ^a	Intermediate adjustment	Full adjustment ^b
Education^c				
No formal schooling	1413	1.00	-	1.00
Primary school	1906	0.85	-	0.92
Middle school	849	0.75	-	0.85
High school	341	0.59	-	0.71
College or university	103	0.48	-	0.65
Smoking^d				
Never	2317	1.00	1.00	1.00
Occasional	300	1.03	1.06	1.06
Ex-regular	404	1.11	1.08	1.09
Current regular	1591	1.02	1.02	1.04
Alcohol^e				
Never regular	2255	1.00	1.00	1.00
Occasional	182	1.27	1.25	1.20
Ex-regular	1145	0.81	0.81	0.90
Reduced intake	142	0.94	0.93	1.06
Monthly	139	0.99	0.98	0.99
Weekly	749	0.97	0.97	0.95
Physical activity (MET hours/day)				
<10	1639	1.00	-	1.00
10-19.9	1429	0.77	-	0.80
20-29.9	723	0.69	-	0.72
30-39.9	476	0.70	-	0.75
≥40	345	0.76	-	0.80

^aAdjusted for age, sex, education and study area; ^bAdditionally adjusted for alcohol, smoking, systolic blood pressure, physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, meat consumption, dairy product consumption, eggs consumption, wheat consumption, tea consumption; ^cAdjusted for variables as above excluding education; ^dIntermediate adjustment: adjusted for age, sex, education, study area and alcohol consumption; ^eIntermediate adjustment: adjusted for age, sex, education, study area and smoking.
HR, hazard ratio; MET, metabolic equivalent of task.

Table E.6. Association of intra-cerebral haemorrhage with dietary consumption

Dietary consumption	No. of events	HRs by different levels of adjustment	
		Minimal adjustment ^a	Full adjustment ^b
Fresh fruit			
Rarely/never	500	1.00	1.00
Monthly	2144	0.89	0.98
1-3 days/week	1296	0.78	0.89
4-6 days/week	282	0.69	0.81
Daily	390	0.63	0.75
Dairy products			
Rarely/never	3602	1.00	1.00
Monthly	510	0.98	1.06
1-3 days/week	224	0.83	0.95
4-6 days/week	39	0.54	0.65
Daily	237	0.64	0.76
Eggs			
Rarely/never	605	1.00	1.00
Monthly	1042	0.85	0.91
1-3 days/week	1976	0.75	0.86
4-6 days/week	510	0.68	0.83
Daily	479	0.62	0.76
Meat			
Rarely/never	410	1.00	1.00
Monthly	897	0.90	0.94
1-3 days/week	1790	0.83	0.90
4-6 days/week	676	0.77	0.86
Daily	839	0.77	0.87
Wheat			
Rarely/never	480	1.00	1.00
Monthly	828	0.88	0.91
1-3 days/week	822	0.91	0.94
4-6 days/week	297	0.96	0.97
Daily	2185	0.91	0.93
Tea			
Never or almost never	1602	1.00	1.00
Occasionally	947	0.97	1.01
Only at certain seasons	126	0.86	0.91
Monthly but less than weekly	142	0.93	0.98
At least once a week	1795	0.85	0.86

^aAdjusted for age, sex, education and study area; ^bAdditionally adjusted for alcohol, smoking, systolic blood pressure, physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, meat consumption, dairy product consumption, eggs consumption, wheat consumption, tea consumption.
HR, hazard ratio.

Figure E.3. Association of intra-cerebral haemorrhage with adiposity

Minimal adjustment: stratified by age, sex and study area and adjusted for education; Intermediate adjustment: BMI additionally adjusted for waist circumference and hip circumference; waist circumference additionally adjusted for hip circumference and BMI; hip circumference additionally adjusted for waist circumference and BMI; waist-to-hip ratio additionally adjusted for BMI; body fat percent additionally adjusted for BMI, waist circumference and hip circumference.

Squares represent the hazard ratio with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

Data for body fat percent missing for 236 participants. BMI, body mass index.

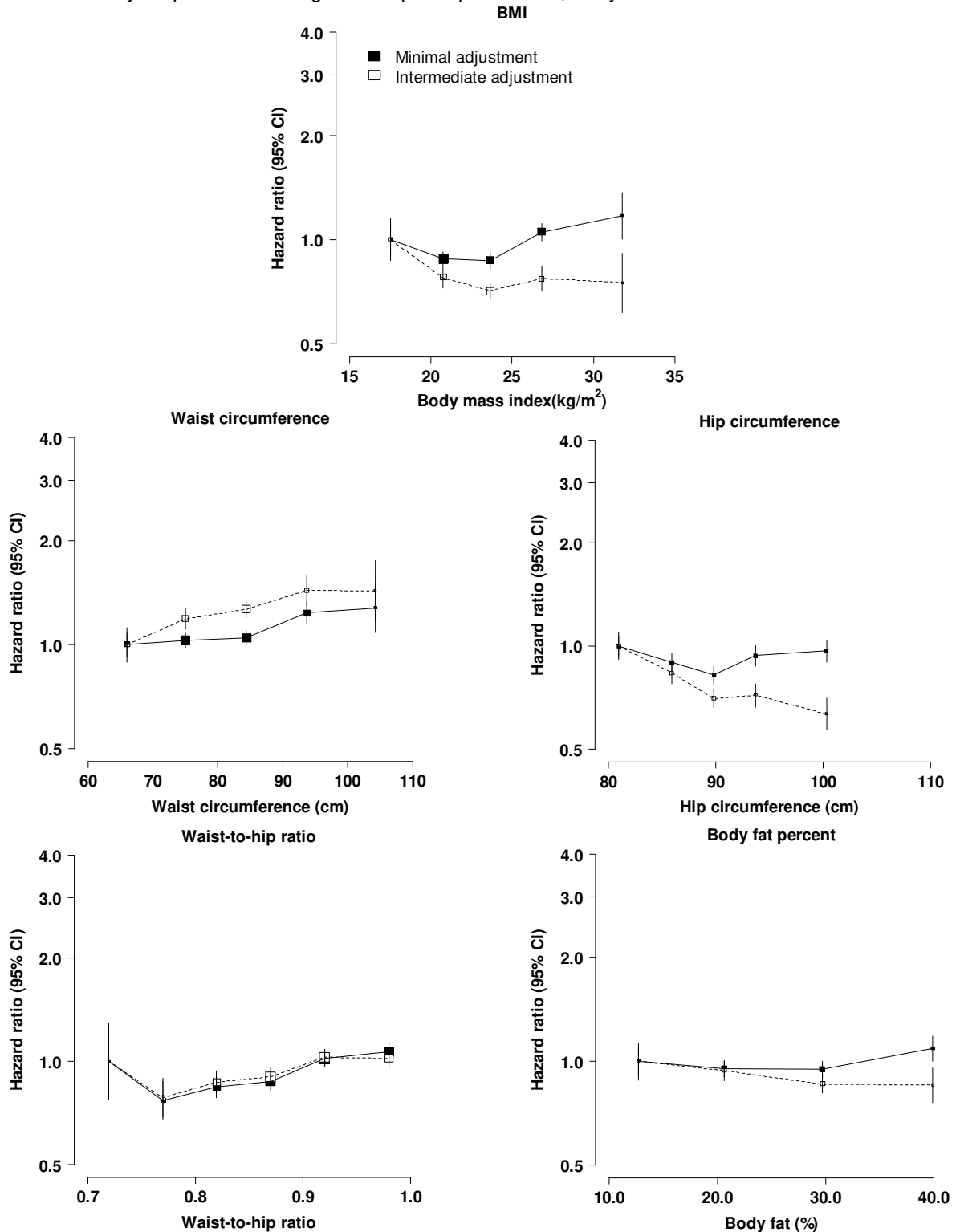


Figure E.4. Association of intra-cerebral haemorrhage with blood pressure

Minimal adjustment: stratified by age, sex, education and study area; Full adjustment: additionally adjusted for alcohol, smoking, systolic blood pressure (for diastolic blood pressure analyses only), physical activity, body mass index, waist circumference, hip circumference, fresh fruit consumption, meat consumption, dairy product consumption, eggs consumption, wheat consumption, tea consumption.

Squares represent the hazard ratio with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

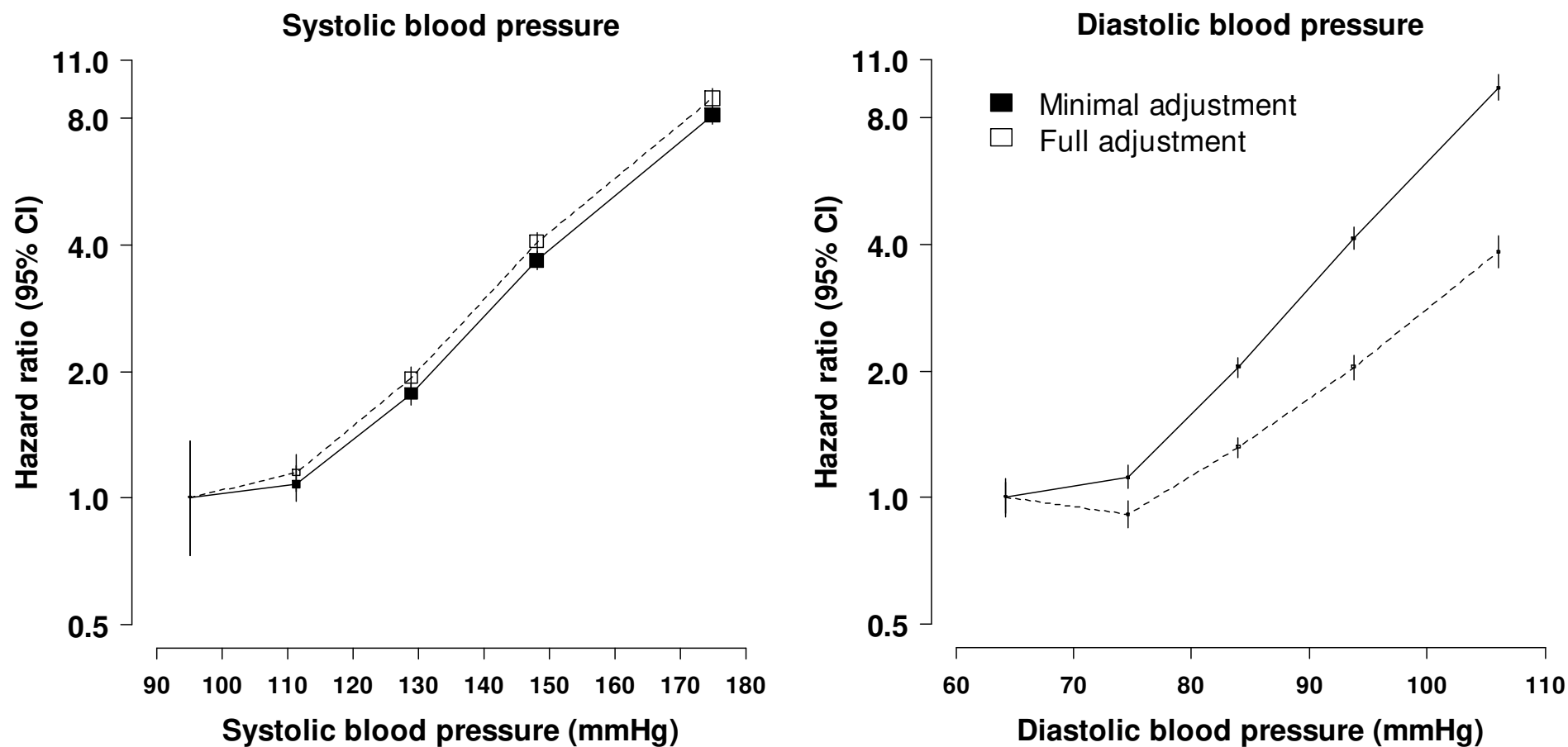


Table E.7. Adjusted hazard ratios for incident cardiovascular diseases by self-reported diabetes status, excluding individuals with diabetes newly diagnosed during follow-up from the comparator group

Outcomes	Diabetes		No diabetes		Model A ^b		Model B ^c		Model C ^d	
	Events	Rate per 100 000 person years ^a	Events	Rate per 100 000 person years ^a	HR	95%CI	HR	95%CI	HR	95%CI
Cardiovascular disease mortality	571		6239		2.48	2.27-2.71	2.05	1.88-2.24	2.05	1.87-2.24
Major occlusive vascular disease	1938		19923		2.09	1.99-2.19	1.82	1.74-1.91	1.76	1.68-1.85
Ischaemic heart disease										
Fatal myocardial infarction	127		1221		3.00	2.48-3.62	2.63	2.18-3.18	2.54	2.10-3.08
Non-fatal myocardial infarction	122		1020		2.66	2.19-3.22	2.31	1.90-2.80	2.23	1.83-2.72
Major coronary event	368		2999		2.84	2.54-3.18	2.47	2.20-2.76	2.40	2.15-2.70
Stroke										
<i>Ischaemic stroke</i>										
Fatal	43		409		2.65	1.92-3.66	2.19	1.58-3.04	2.20	1.58-3.06
Non-fatal	1589		16855		1.97	1.87-2.07	1.72	1.63-1.82	1.66	1.57-1.75
Any	1632		17264		1.98	1.88-2.09	1.73	1.64-1.82	1.67	1.58-1.76
<i>Intracerebral haemorrhage</i>										
Fatal	100		1869		2.00	1.63-2.46	1.58	1.29-1.94	1.60	1.30-1.97
Non-fatal	92		2275		1.23	1.00-1.52	1.00	0.81-1.24	1.00	0.81-1.24
Any	192		4144		1.54	1.33-1.79	1.24	1.07-1.44	1.25	1.08-1.45
<i>Total stroke</i>										
Fatal	168		2541		2.19	1.87-2.57	1.76	1.50-2.07	1.79	1.52-2.11
Non-fatal	1749		20158		1.90	1.81-1.99	1.64	1.56-1.73	1.59	1.51-1.67
Any	1917		22699		1.92	1.83-2.01	1.65	1.57-1.73	1.61	1.53-1.69

^aAge-, sex- and study area- standardised rates; ^bModel A: stratified by age, sex and study area; ^cModel B: additionally adjusted for education, smoking, alcohol, physical activity and systolic blood pressure; ^dModel C: additionally adjusted for waist-to-hip ratio.

Events classified as fatal if death from any cause recorded within 28 days of disease onset.

CI, confidence interval; HR, hazard ratio.

Figure E.5. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease by screen-detected diabetes status, by study area

Stratified by age and sex and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall HR. The open diamond represents the overall HR and its 95% CI.

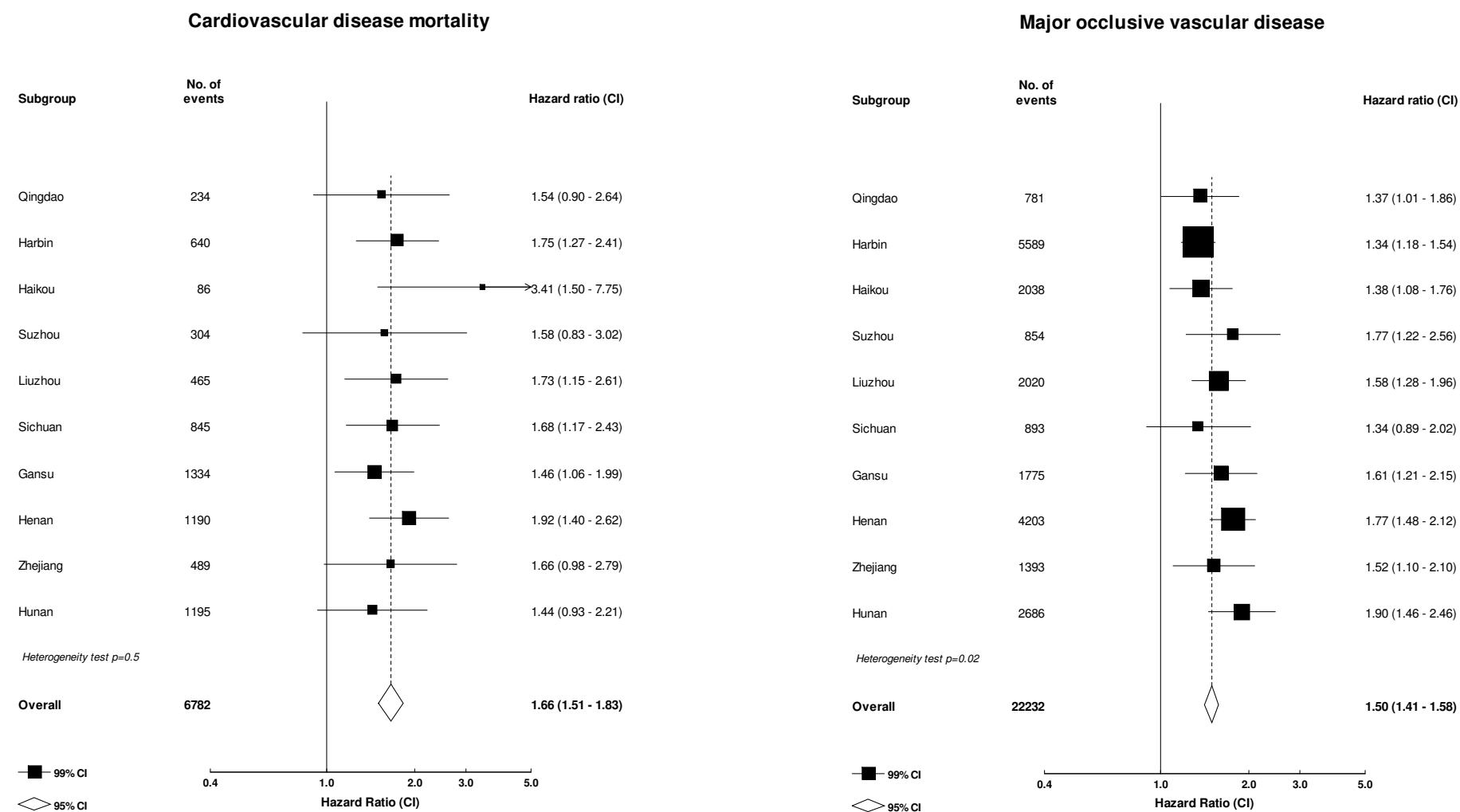


Figure E.6. Adjusted hazard ratios for major coronary event and ischaemic stroke by screen-detected diabetes status, by study area

Stratified by age and sex and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall HR. The open diamond represents the overall HR and its 95% CI.

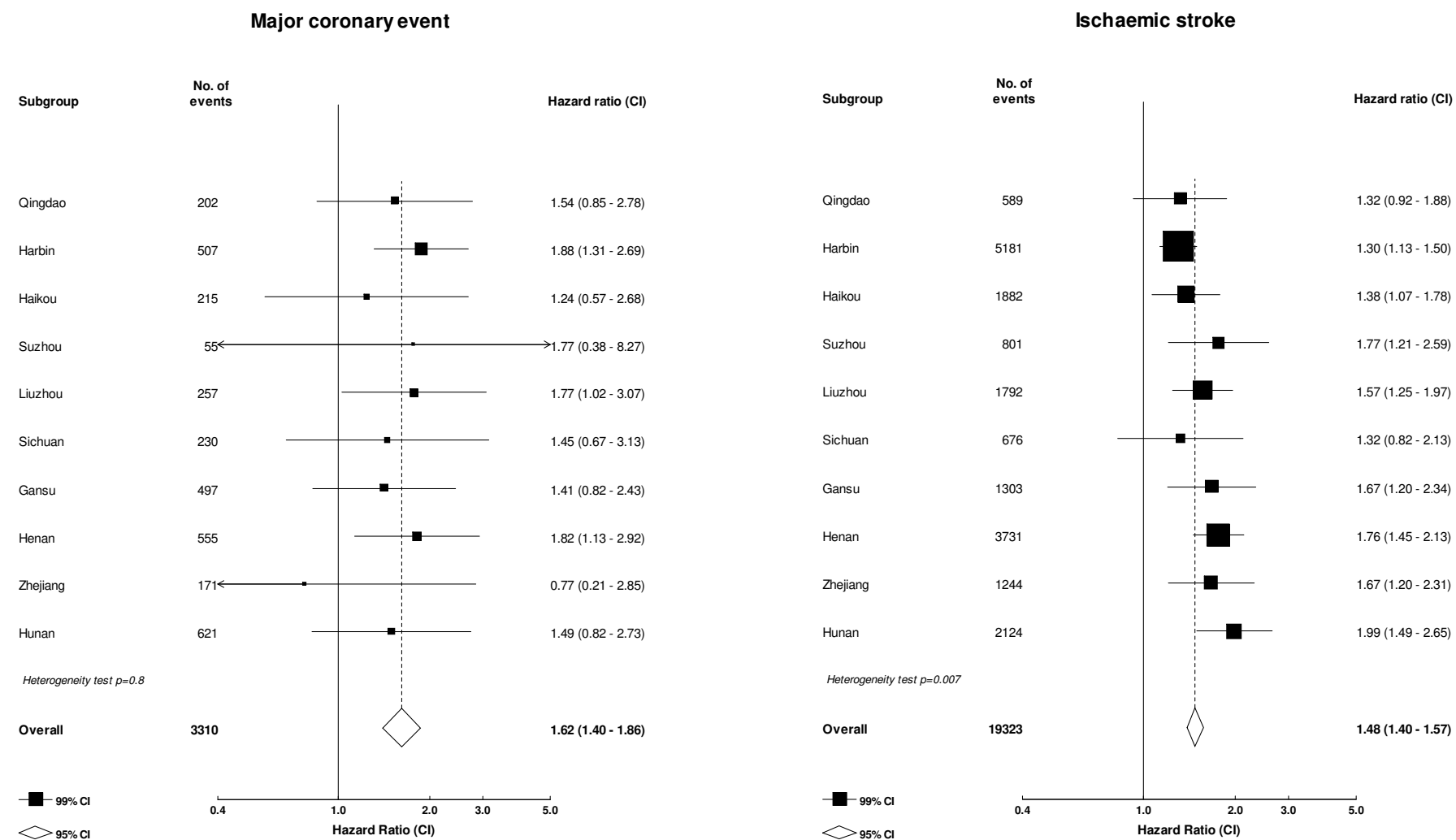


Figure E.7. Adjusted hazard ratios for intracerebral haemorrhage by screen-detected diabetes status, by study area

Stratified by age and sex and adjusted for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratio (HR) with area inversely proportional to the variance. Horizontal lines represent the corresponding 99% confidence intervals (CIs). The dotted line indicates the overall HR. The open diamond represents the overall HR and its 95% CI.

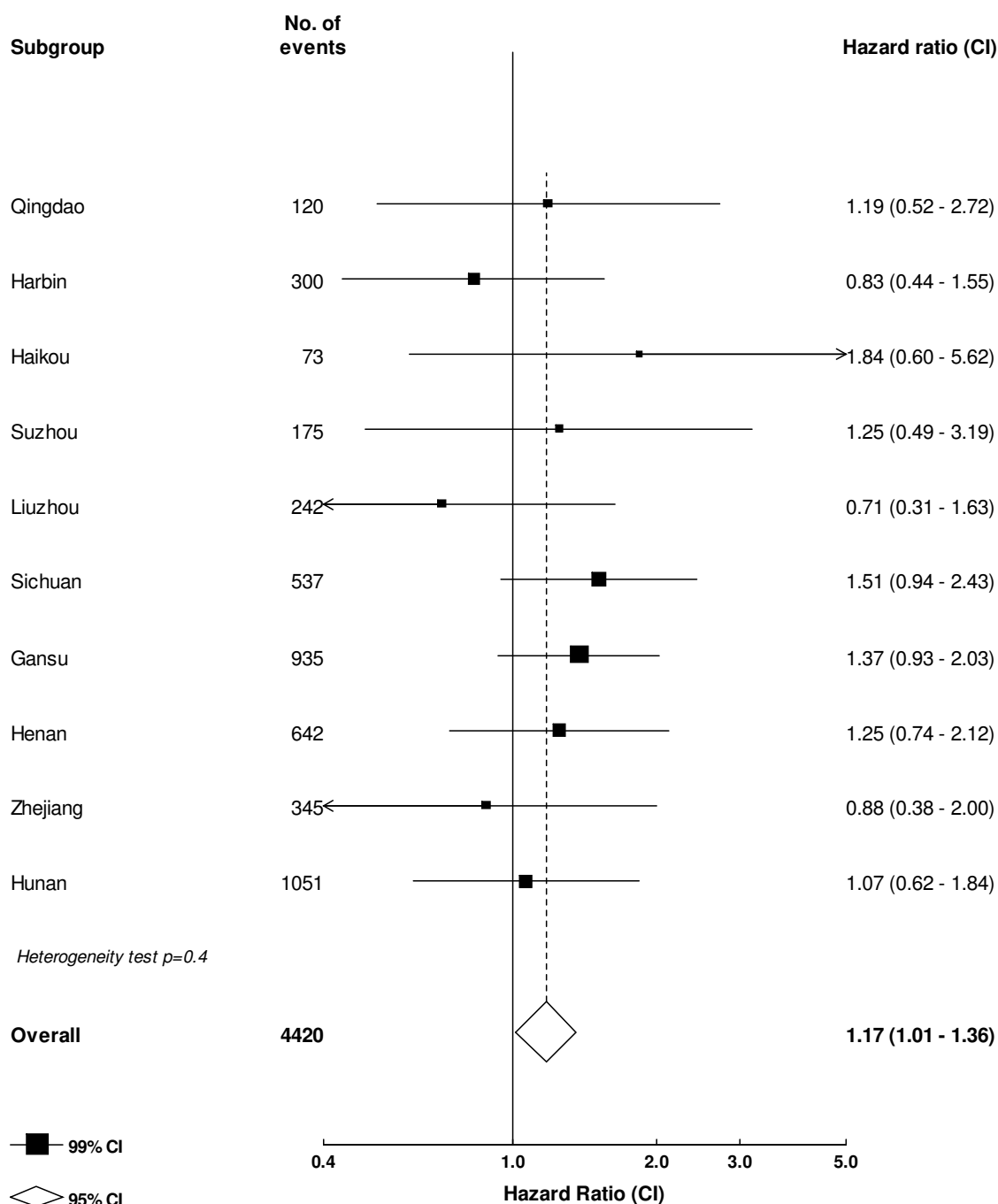


Figure E.8. Adjusted hazard ratios for major occlusive vascular disease by screen-detected diabetes status in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratios (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

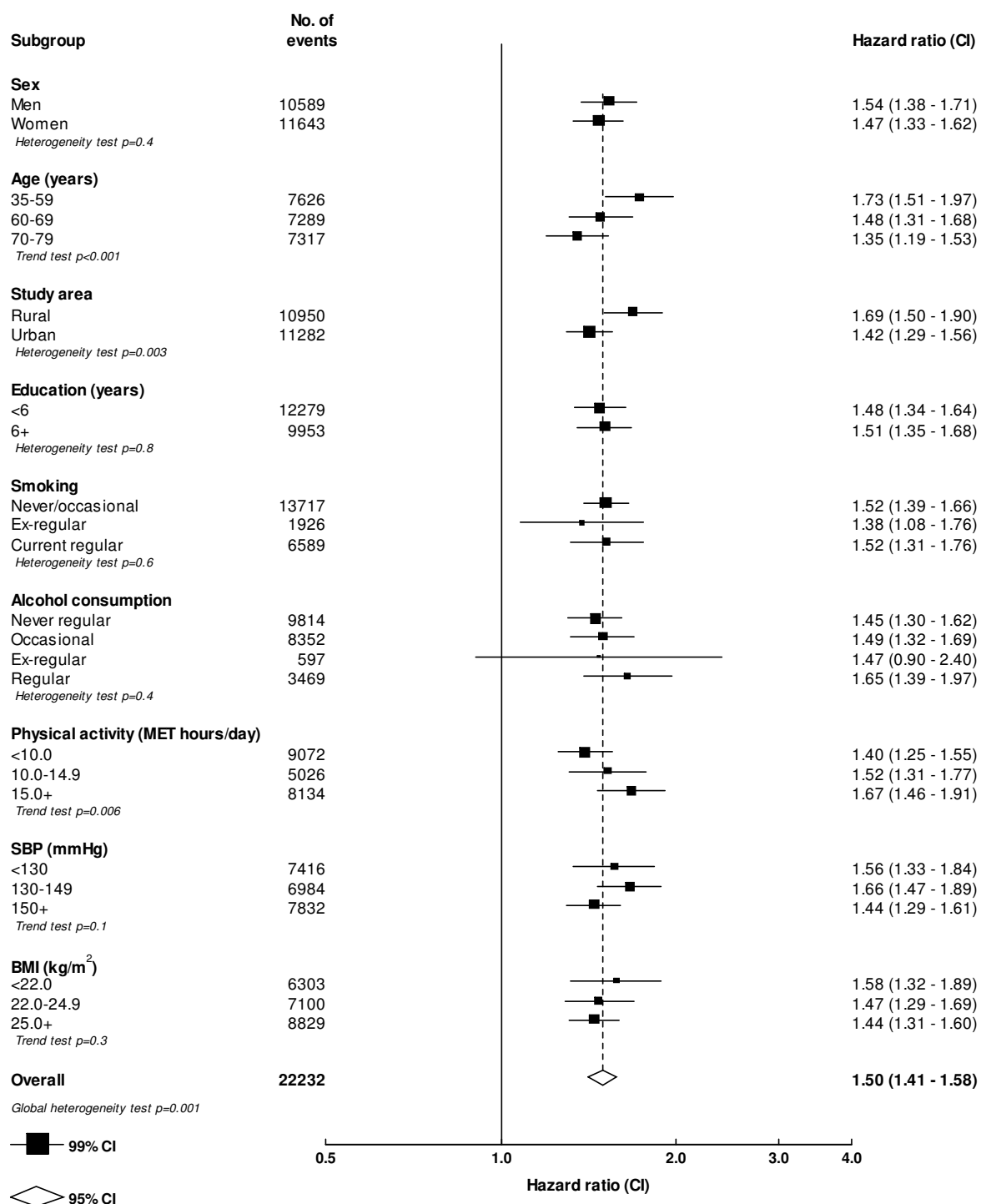


Figure E.9. Adjusted hazard ratios for ischaemic stroke by screen-detected diabetes status in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratios (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

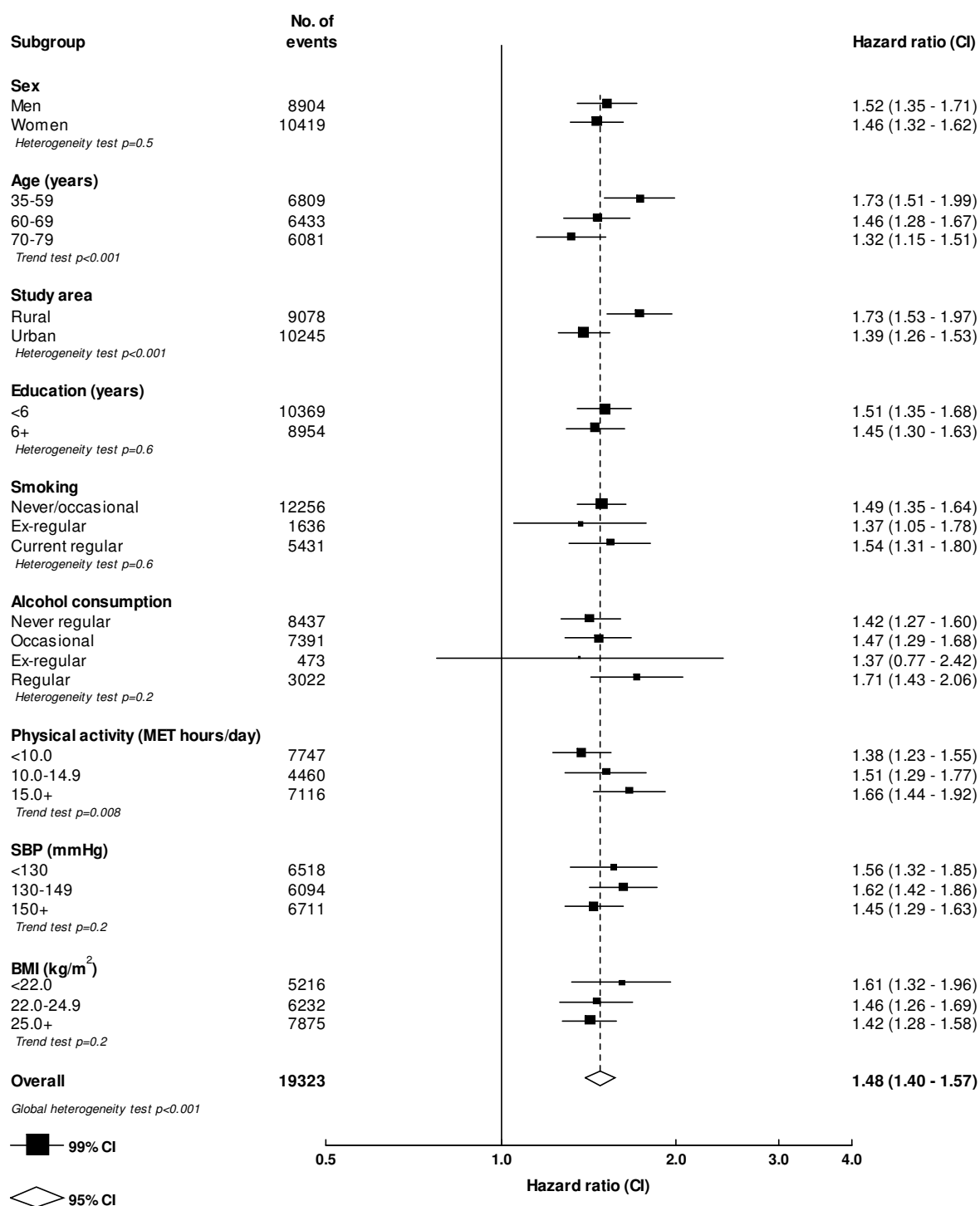


Figure E.10. Adjusted hazard ratios for cardiovascular disease mortality by screen-detected diabetes status in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratios (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

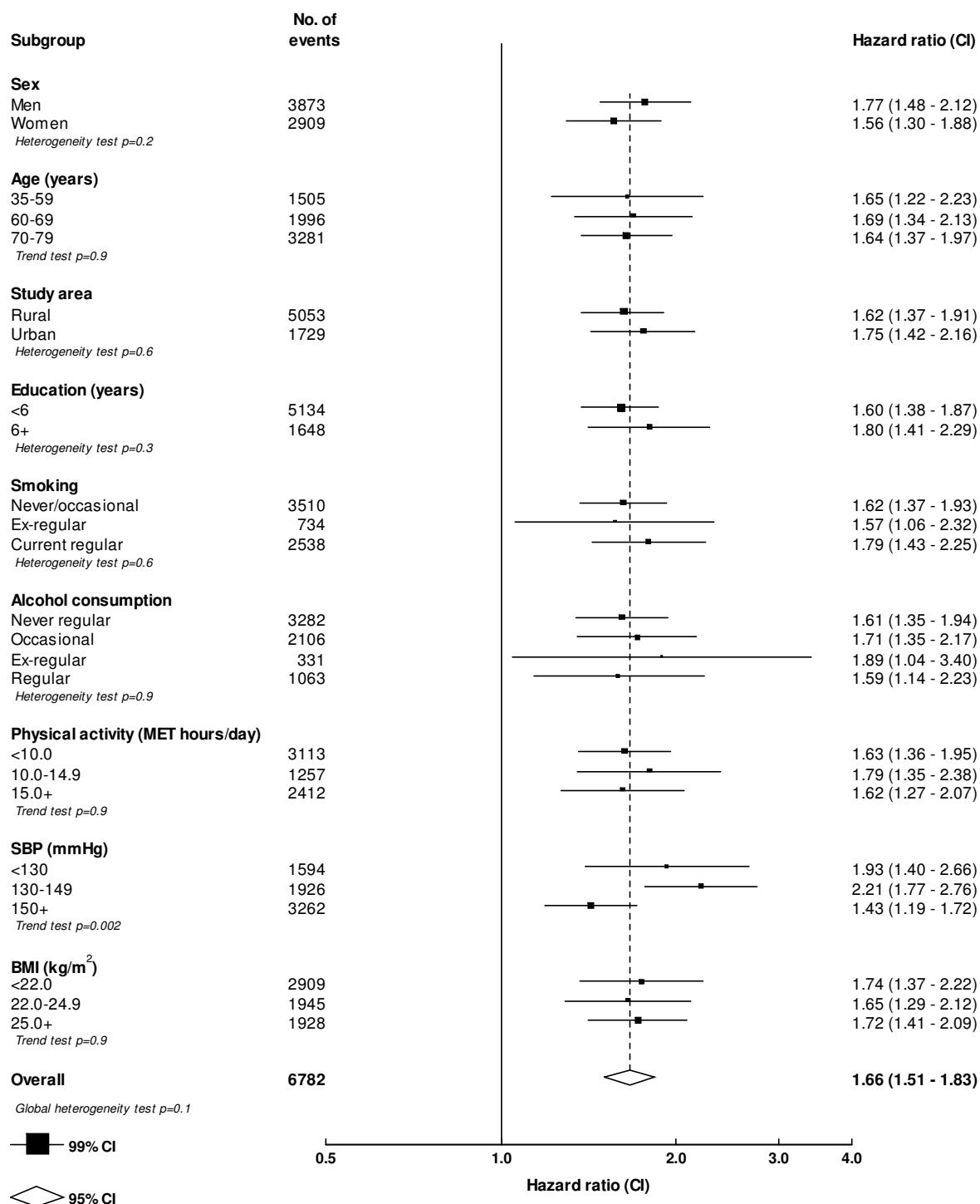


Figure E.11. Adjusted hazard ratios for major coronary event by screen-detected diabetes status in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratios (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

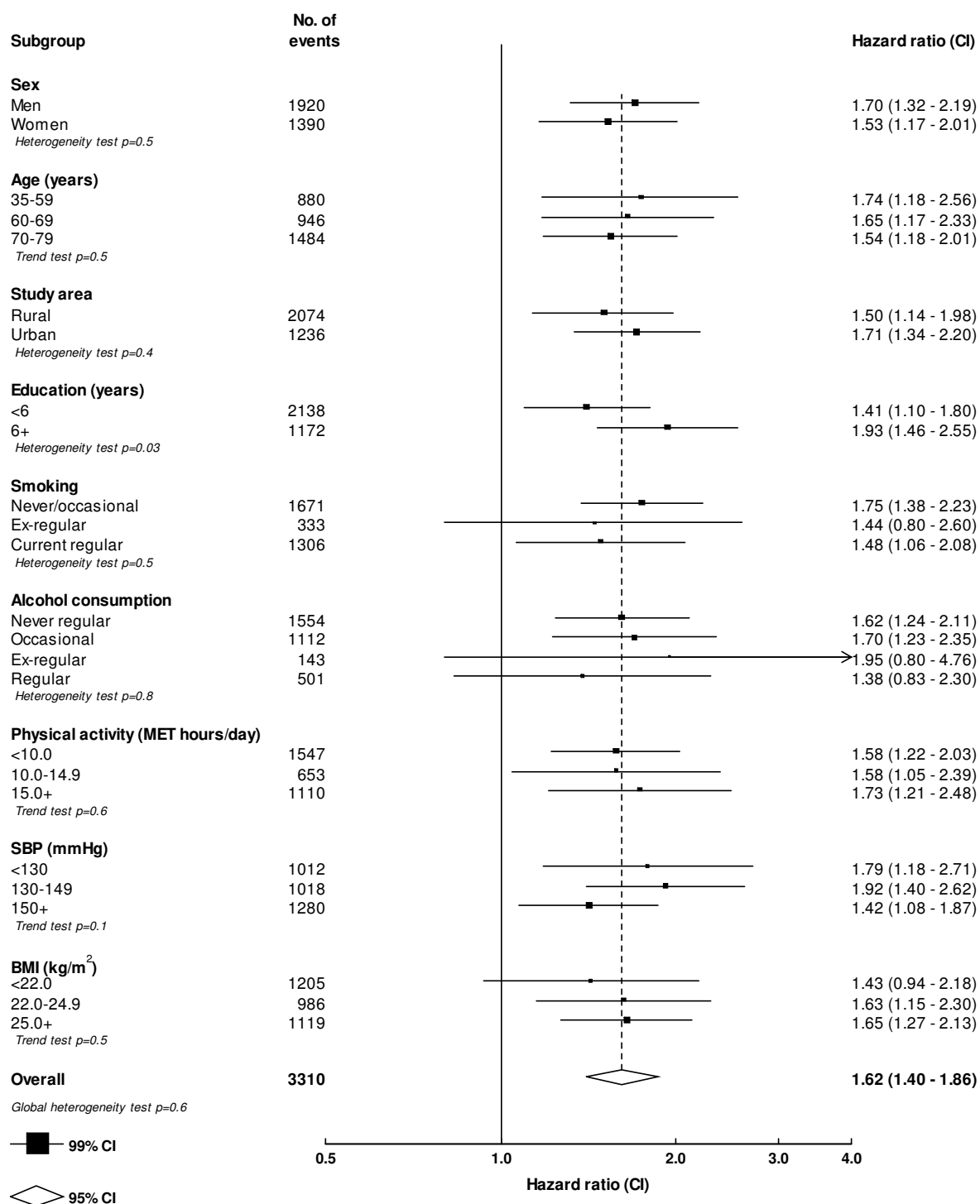


Figure E.12. Adjusted hazard ratios for intracerebral haemorrhage by screen-detected diabetes status in population subgroups

Stratified by age, sex and study area and adjusted (except where it is the variable of interest) for education, smoking, alcohol consumption, physical activity and systolic blood pressure. Shaded squares represent the hazard ratios (HR) with area inversely proportional to the variance of the log HR. Horizontal lines represent the corresponding 99% confidence intervals (CI). The dotted line represents the overall HR. The open diamond represents the overall HR and its 95% CI. BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

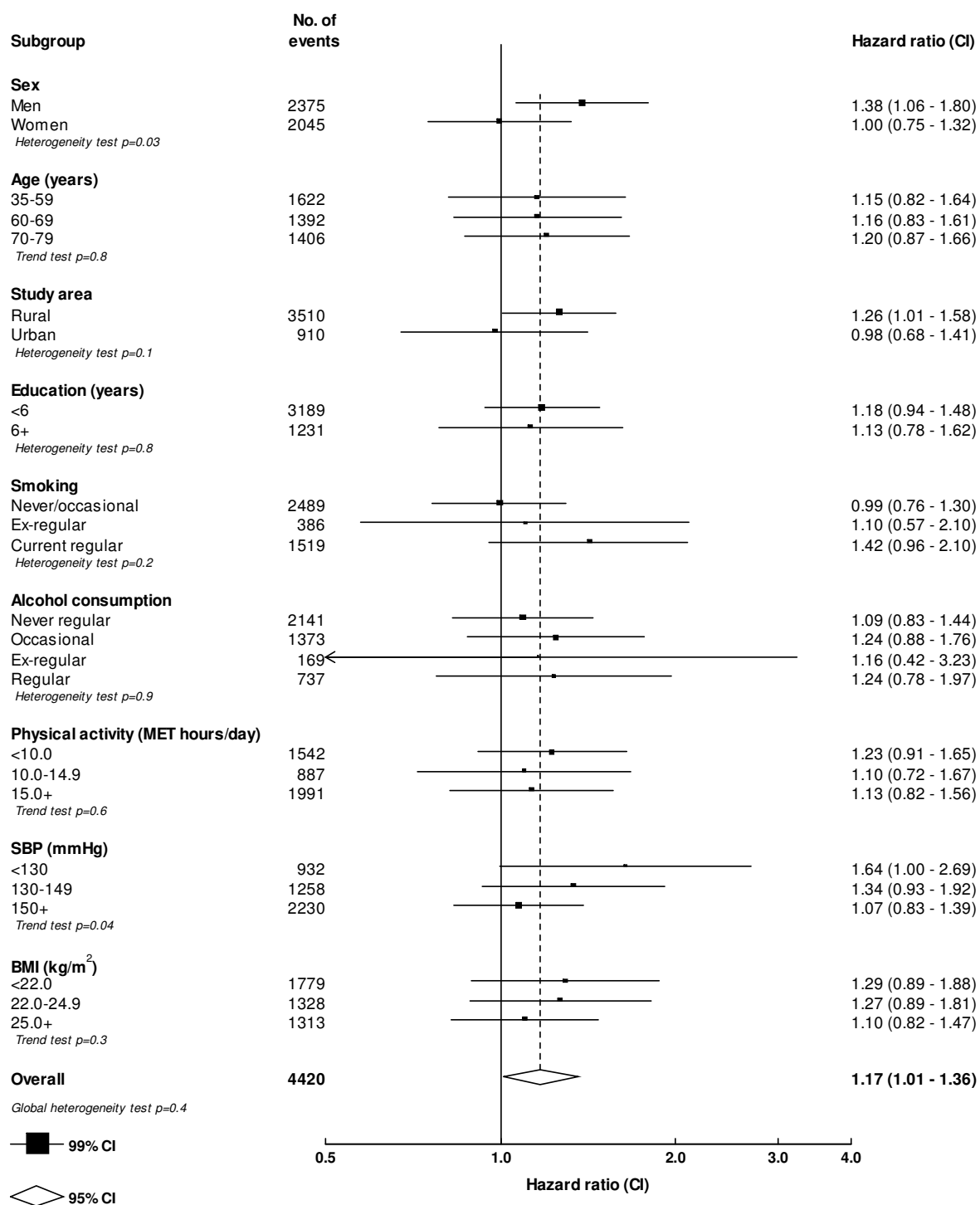


Table E.8. Cardiovascular disease risk factors in men and women with and without self-reported diabetes at baseline

Cardiovascular disease risk factor	Sex	No diabetes		Diabetes		Difference in mean or proportion	p value
		Mean	SE	Mean	SE		
Current smoking (%)	Men	62.4	0.1	54.8	1.0	-7.6	
	Women	2.4	0.0	2.3	0.2	-0.1	<0.001
Physical activity (MET hours/day)	Men	23.0	0.0	19.2	0.2	-3.8	
	Women	20.6	0.0	18.5	0.1	-2.1	<0.001
SBP (mmHg)	Men	131.6	0.0	135.6	0.3	4.1	
	Women	129.6	0.0	137.8	0.2	8.1	<0.001
BMI (kg/m²)	Men	23.3	0.0	24.3	0.0	1.0	
	Women	23.7	0.0	24.6	0.0	0.8	<0.001

Standardised to the age and study area structure of the study population.

BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure.

Appendix F. Additional methods and results: Plasma glucose and the risk of incident cardiovascular disease

Figure F.1. Incidence of cardiovascular diseases by age and sex in prospective plasma glucose analyses population (n=467 508)

Standardised to study area and education structure of study population.
Squares represent cardiovascular disease incidence and vertical lines represent corresponding 95% confidence intervals.

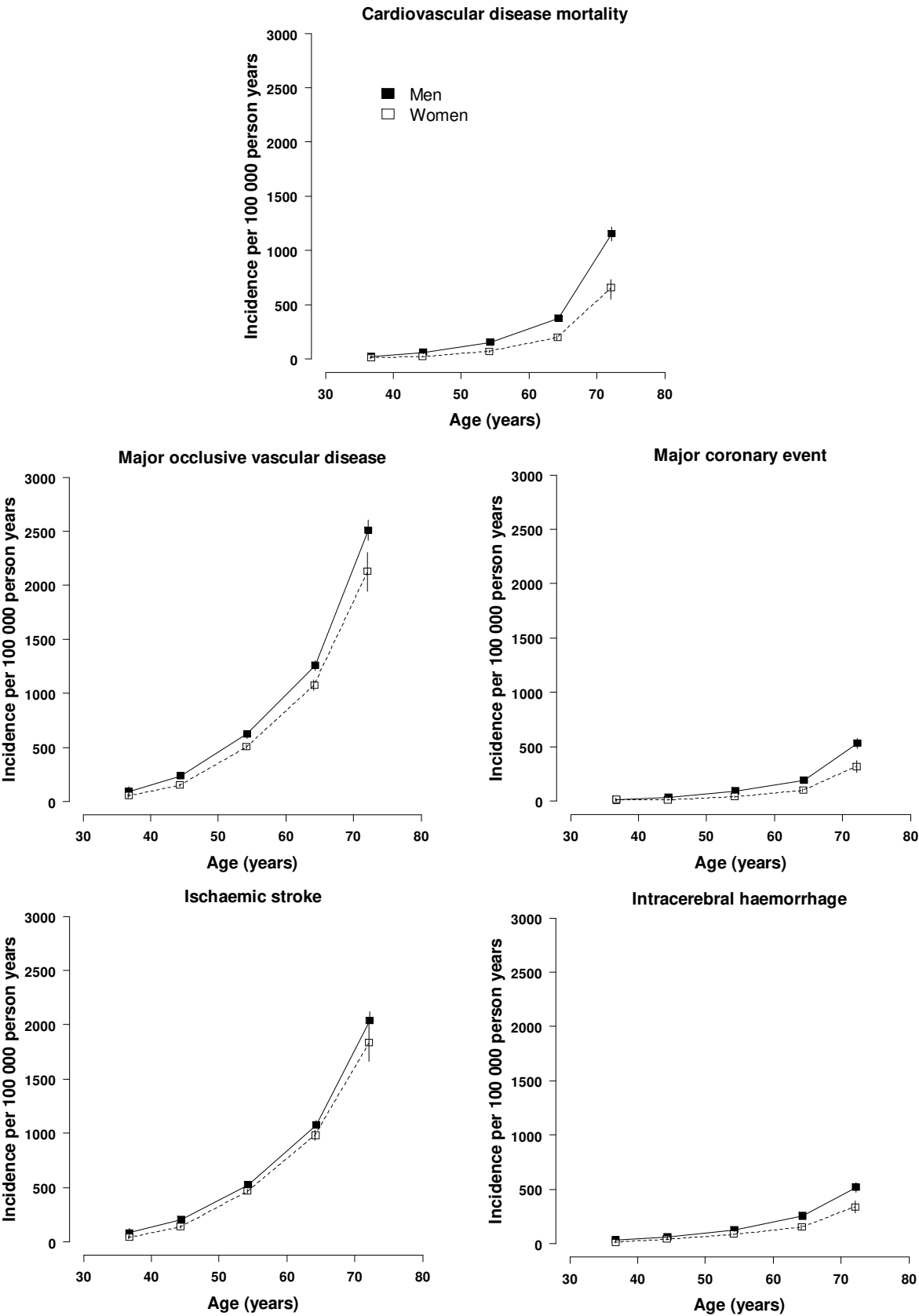


Figure F.2. Incidence of cardiovascular disease mortality by study area in prospective plasma glucose analyses population (n=467 508)

Standardised to age and education structure of study population.
Error bars represent 95% confidence intervals.
R, rural; U, urban.

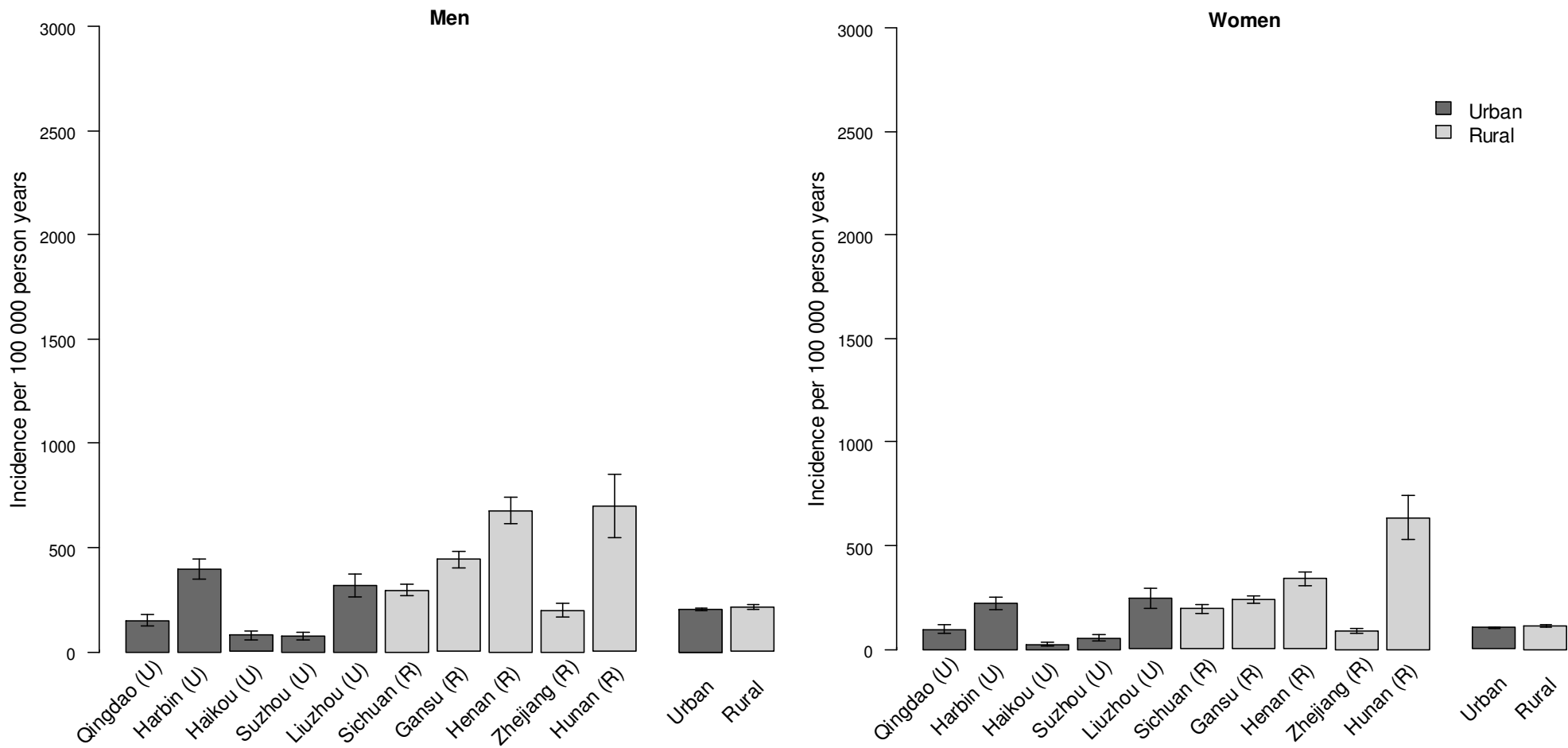


Figure F.3. Incidence of major occlusive vascular disease by study area in prospective plasma glucose analyses population (n=467 508)

Standardised to age and education structure of study population.
Error bars represent 95% confidence intervals.
R, rural; U, urban.

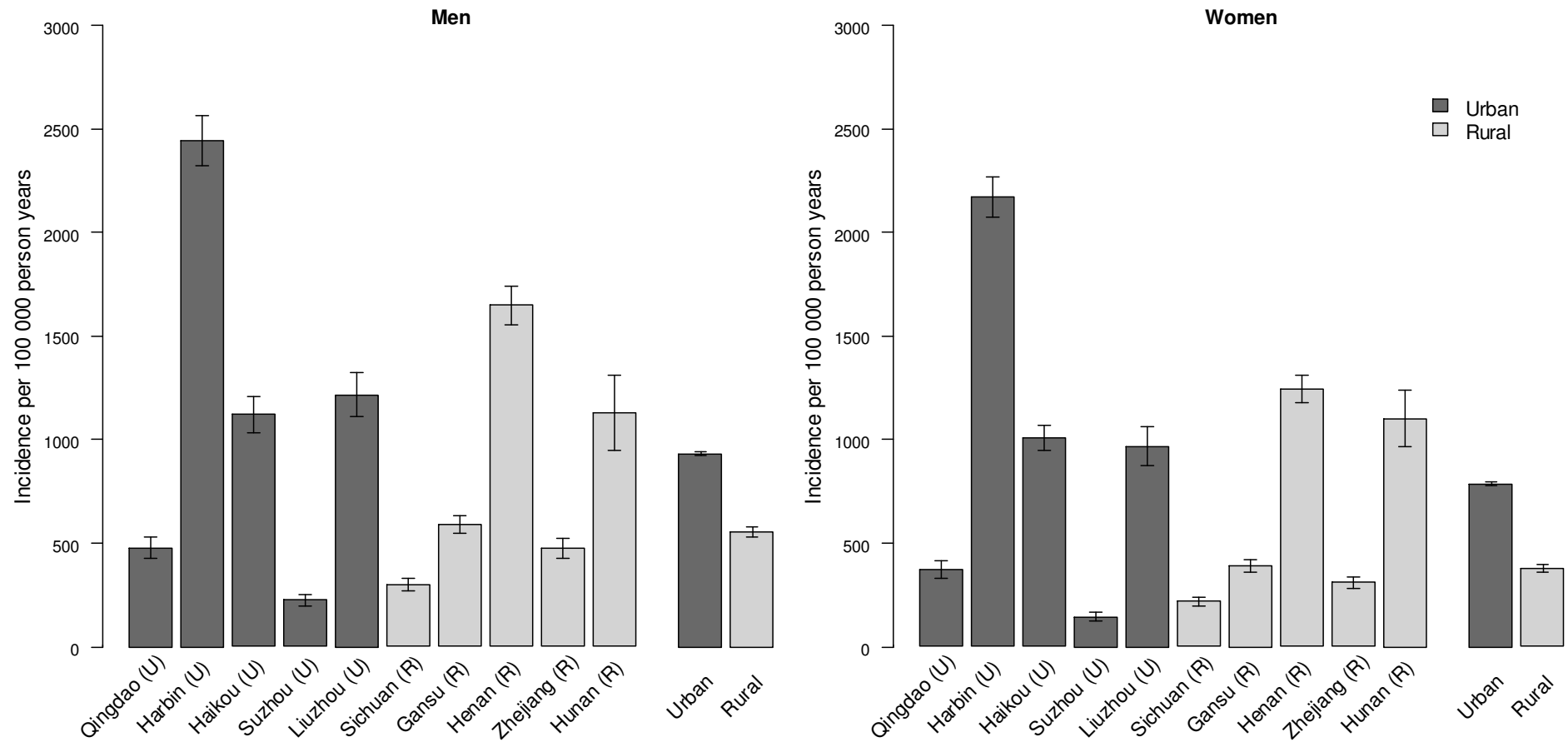


Figure F.4. Incidence of major coronary event by study area in prospective plasma glucose analyses population (n=467 508)

Standardised to age and education structure of study population.

Error bars represent 95% confidence intervals.

R, rural; U, urban.

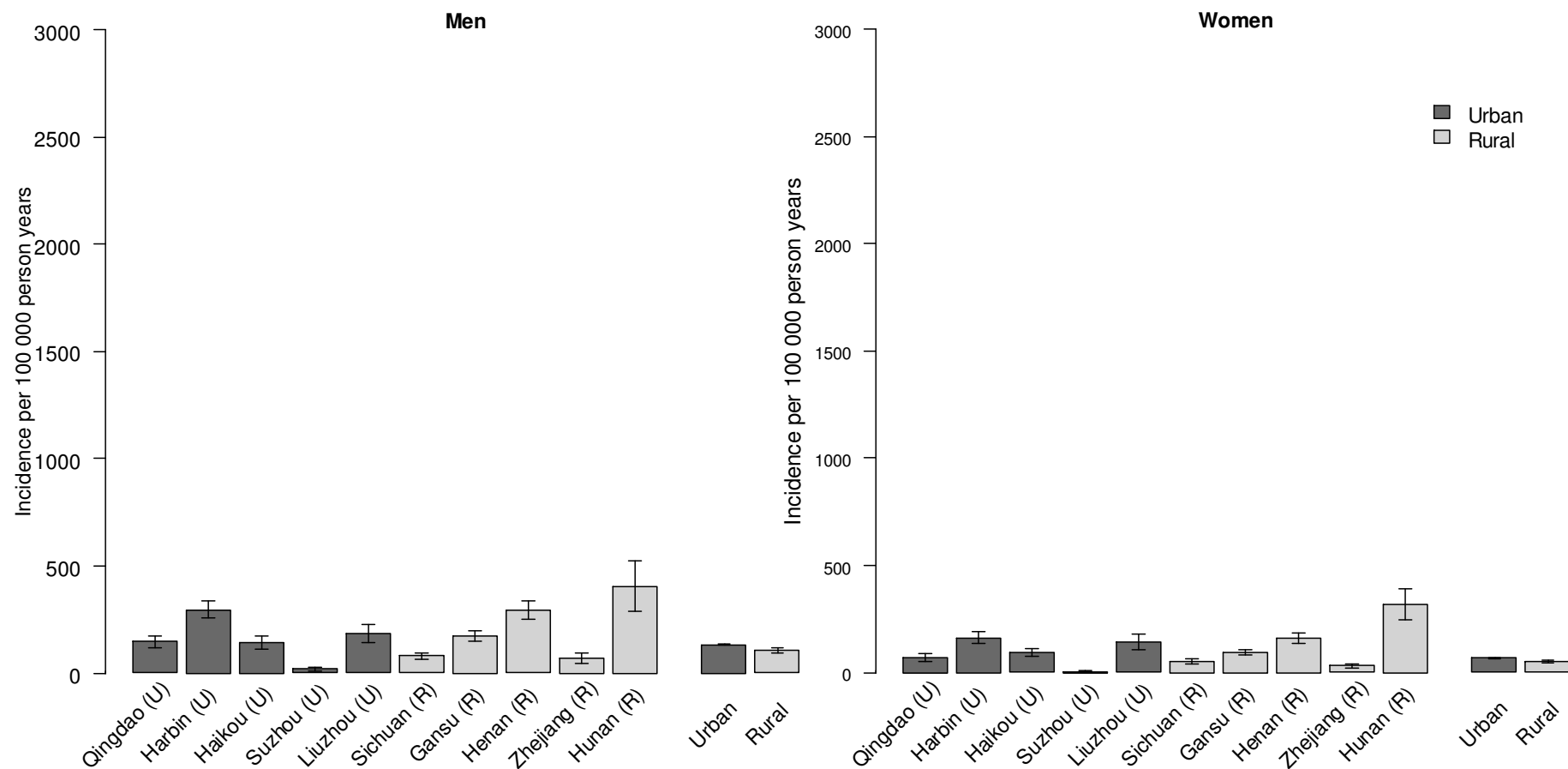


Figure F.5. Incidence of ischaemic stroke by study area in prospective plasma glucose analyses population (n=467 508)

Standardised to age and education structure of study population.

Error bars represent 95% confidence intervals.

R, rural; U, urban.

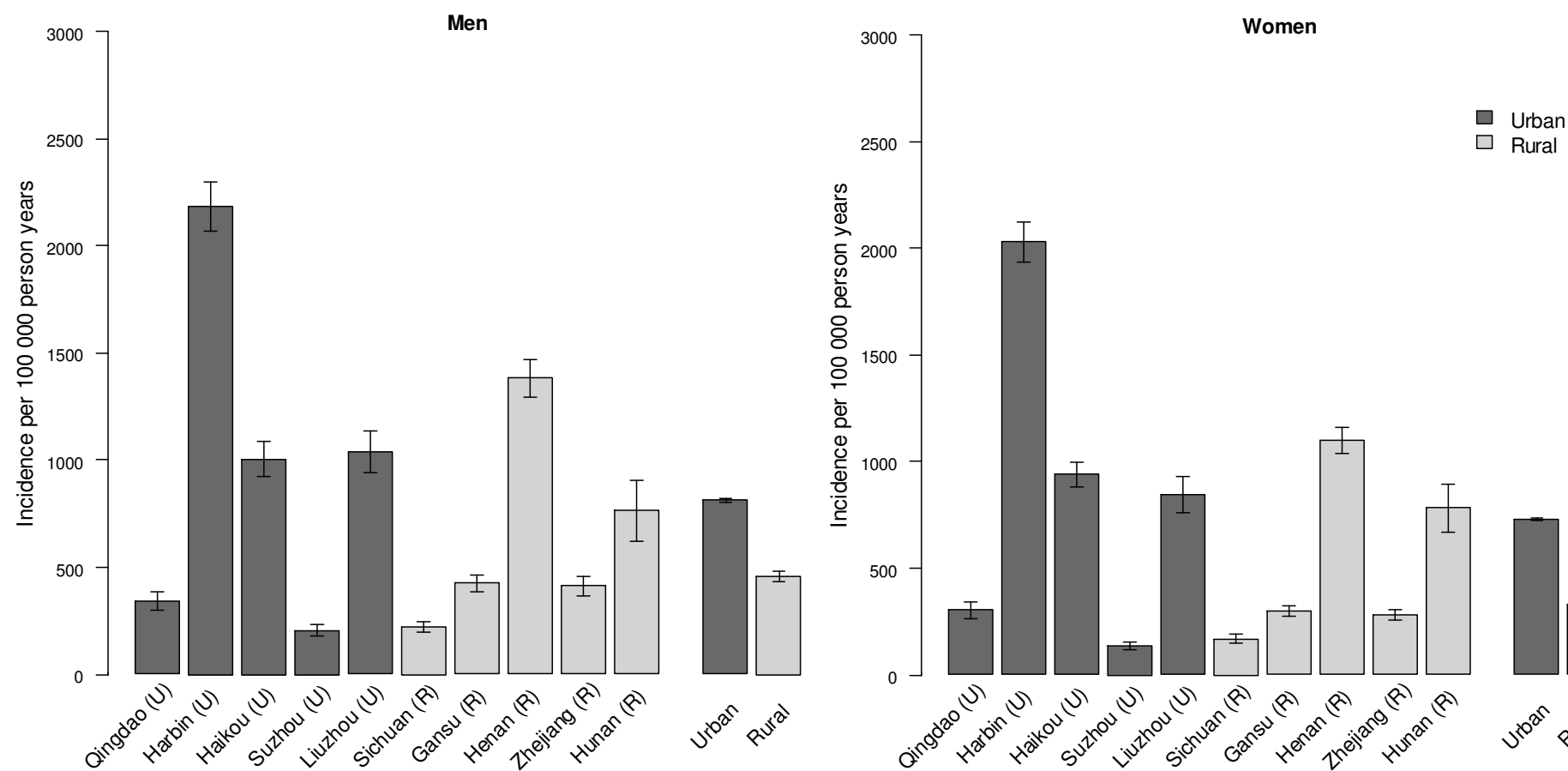


Figure F.6. Incidence of intracerebral haemorrhage by study area in prospective plasma glucose analyses population (n=467 508)

Standardised to age and education structure of study population.
Error bars represent 95% confidence intervals.
R, rural; U, urban.

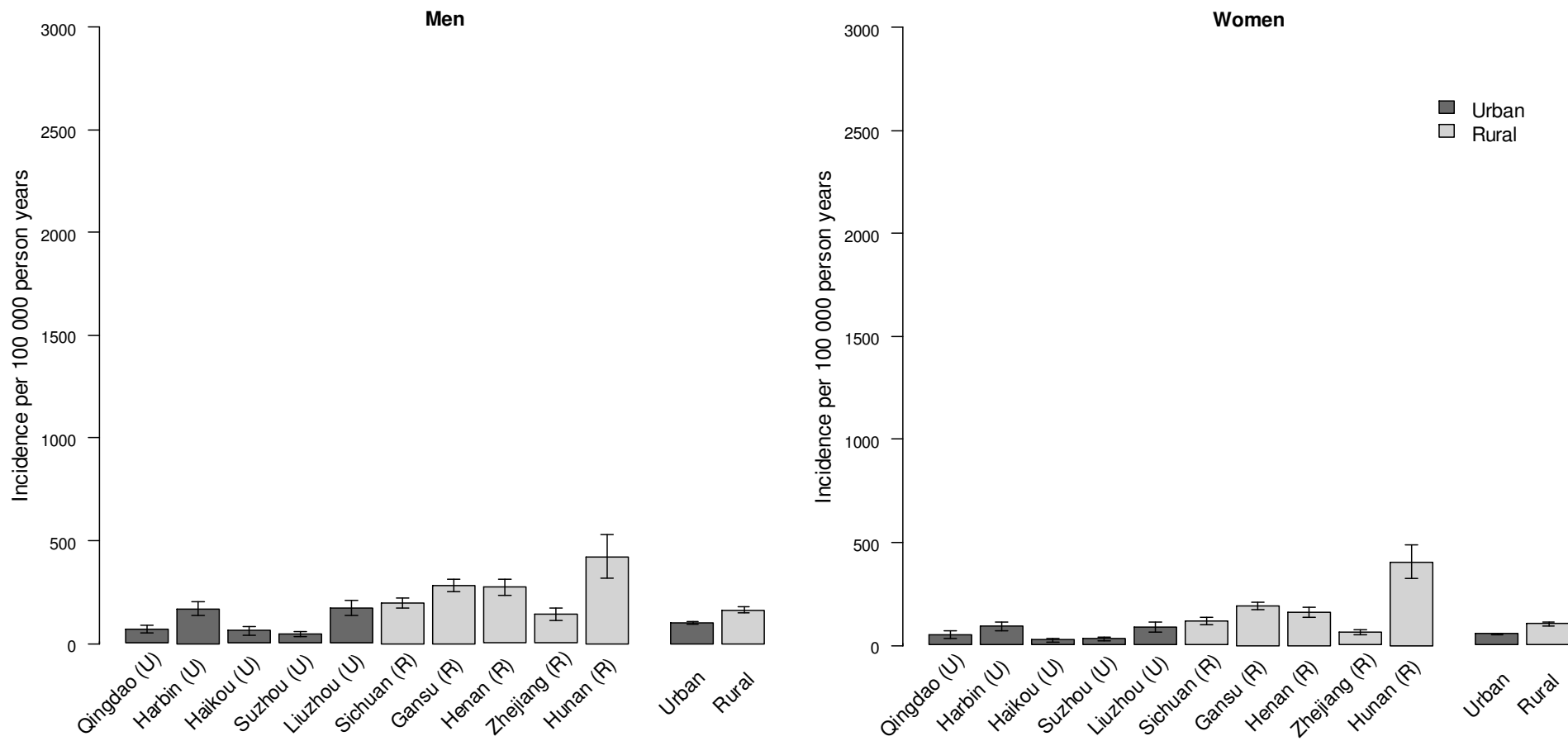


Figure F.7. Adjusted hazard ratios for cardiovascular disease mortality and major occlusive vascular disease by baseline random plasma glucose in men and women

Stratified by age and study area, and adjusted for education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HR) are plotted against mean baseline random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

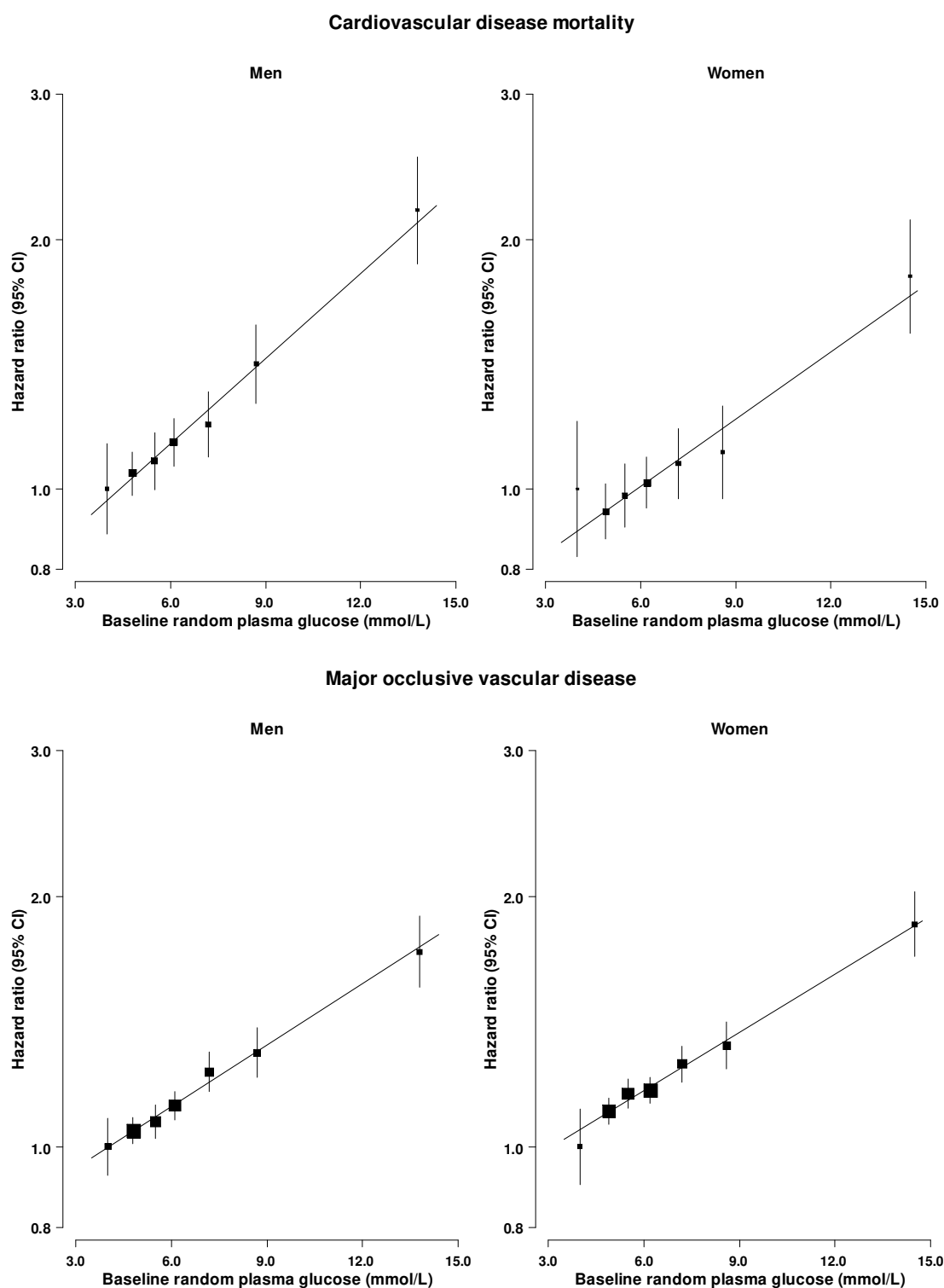


Figure F.8. Adjusted hazard ratios for major coronary event and ischaemic stroke by baseline random plasma glucose in men and women

Stratified by age and study area, and adjusted for education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HR) are plotted against mean baseline random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

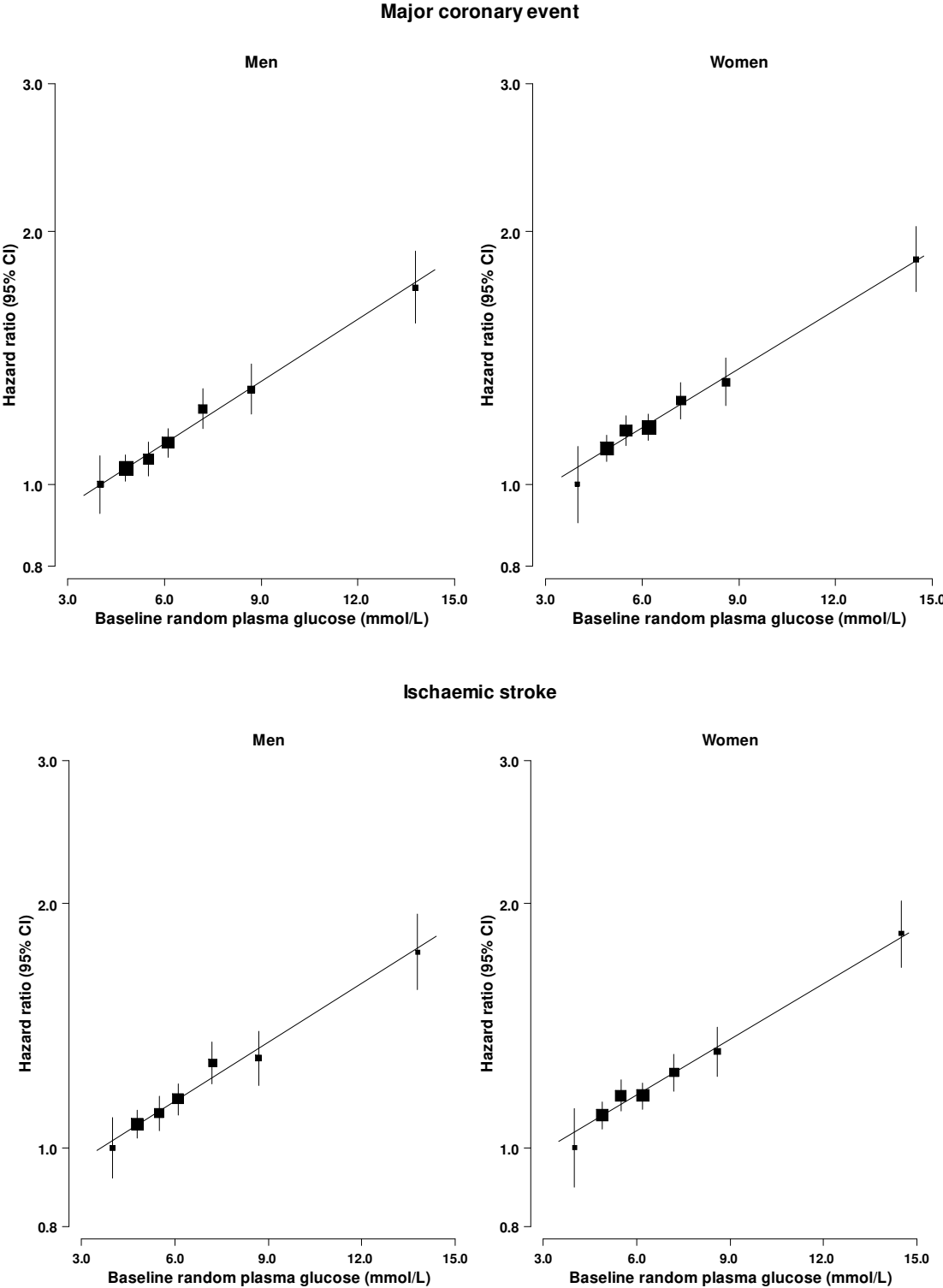


Figure F.9. Adjusted hazard ratios for intracerebral haemorrhage by baseline random plasma glucose in men and women

Stratified by age and study area, and adjusted for education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HR) are plotted against mean baseline random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs).

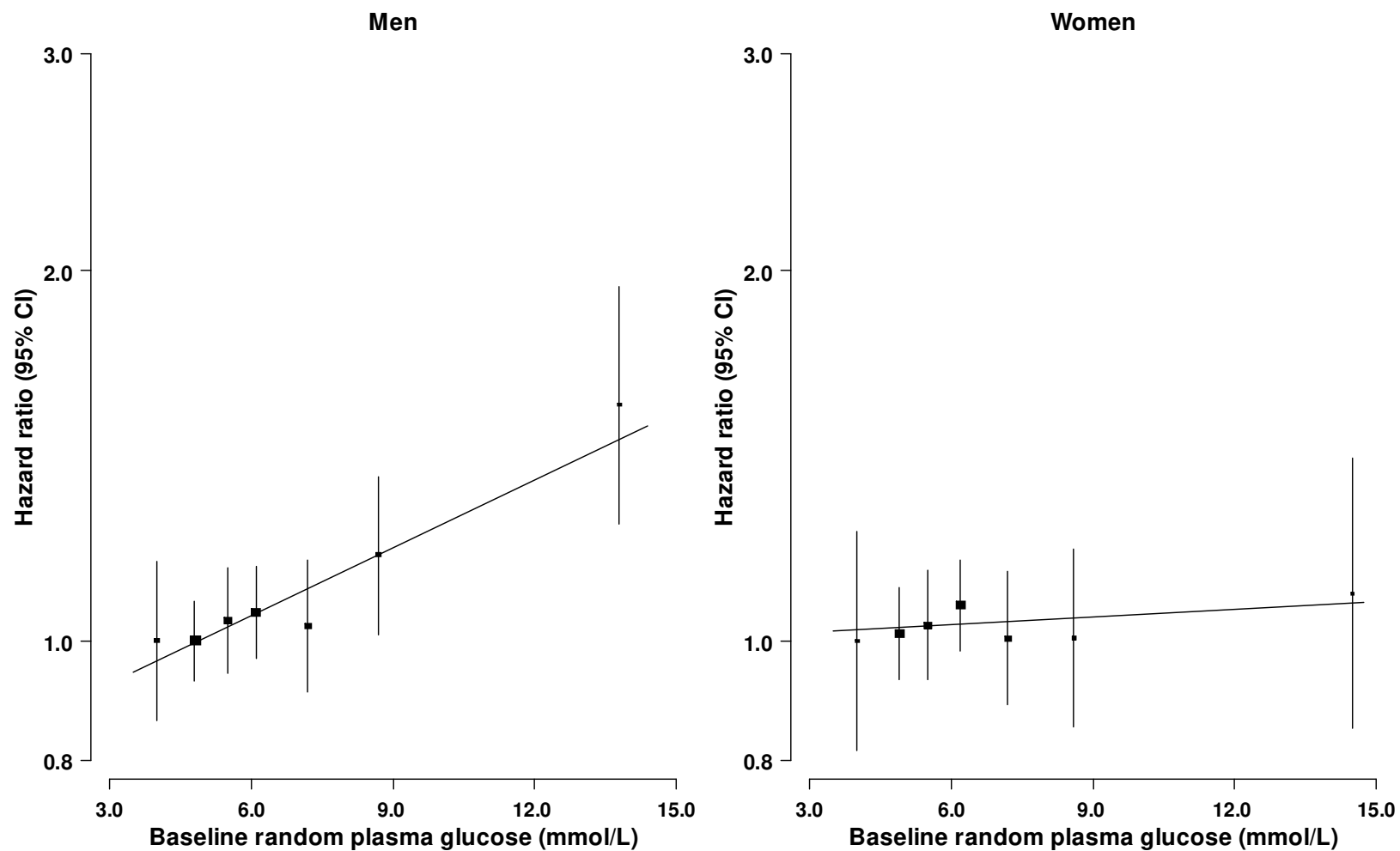


Figure F.10. Adjusted hazard ratios for cardiovascular diseases by baseline random plasma glucose

Stratified by age, sex and study area, and adjusted for education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HR) are plotted against mean baseline random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs). Numbers above the CIs are the HRs and numbers below the CIs are of cardiovascular disease events.

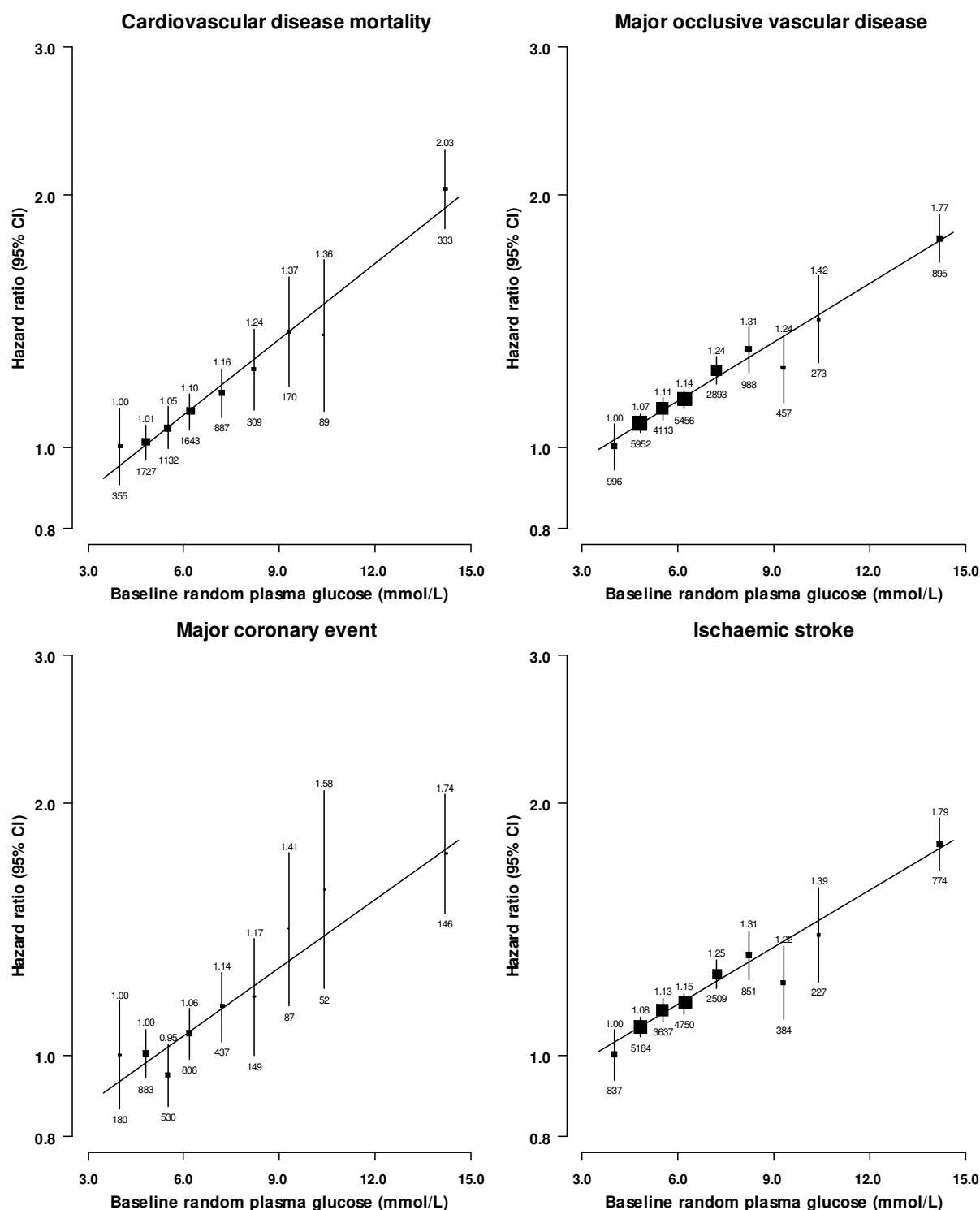
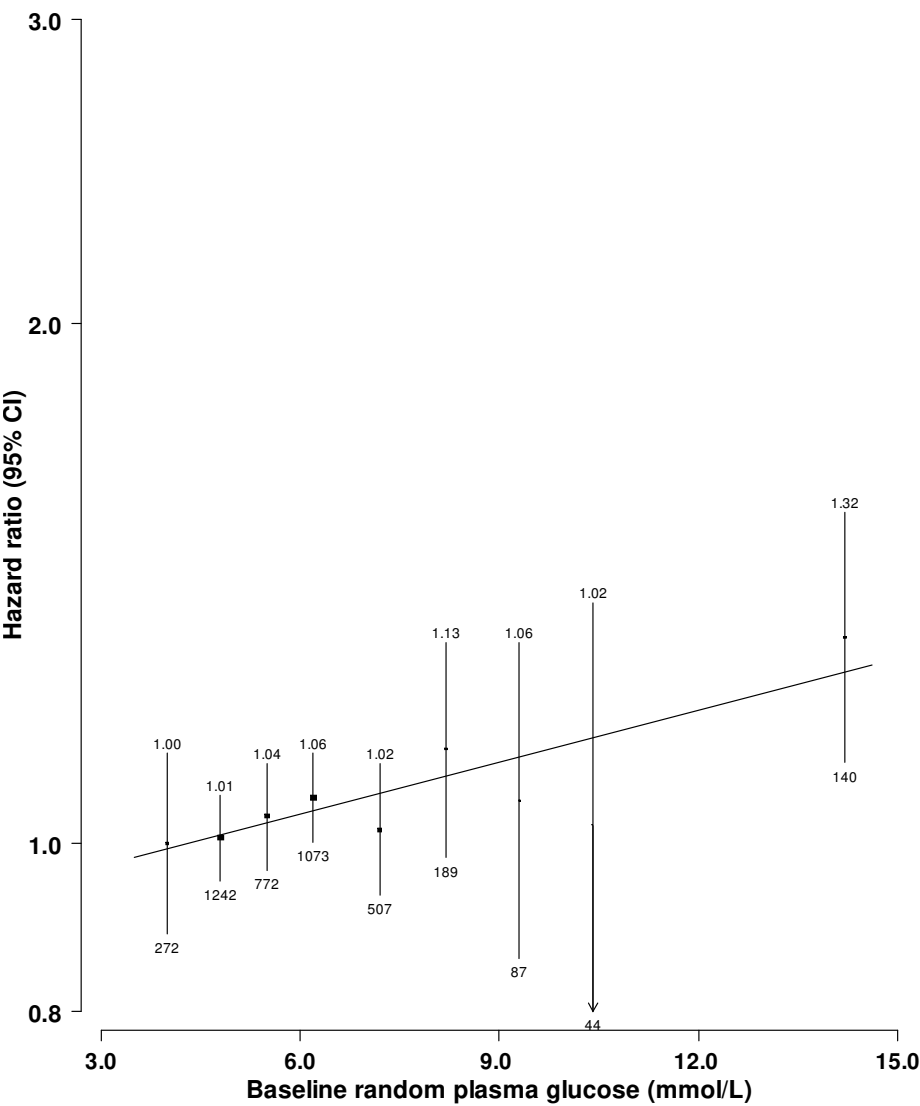


Figure F.11. Adjusted hazard ratios for intracerebral haemorrhage by baseline random plasma glucose

Stratified by age, sex and study area, and adjusted for education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HR) are plotted against mean baseline random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs). Numbers above the CIs are the HRs and numbers below the CIs are of intracerebral haemorrhage events.



Appendix Note F.1. Fasting time-adjusted plasma glucose and the risk of incident cardiovascular disease

Background

As described, plasma glucose levels are influenced by dietary intake, both the composition of food or drink consumed and the duration of time passed since it was consumed.^{215, 217, 226} With the exception of one study area (Zhejiang), plasma glucose levels measured in the CKB were random, that is were measured without specifying a prior fasting time. Data on fasting time were, however, collected from participants. The relevance of fasting time to the association between RPG levels and the risk of CVDs is unclear, and analyses were therefore undertaken to investigate whether adjustment of plasma glucose levels for self-reported fasting time, in order to at least partially control for the effects of recent dietary intake, would enable more robust estimates of the true association of plasma glucose levels with risk of CVDs.

Methods

As described in Section 4.3.3, there was an initial inverse association between fasting period and mean age-, sex-, study area- and education-adjusted plasma glucose level, which flattened off at longer fasting periods (Figure F.12). The linearity of the association between plasma glucose levels and fasting time was examined separately for fasting periods <8 hours and ≥8 hours. This 8 hour cut-point was selected based on the observed association of plasma glucose levels with fasting time in the CKB, and on guidance for fasting plasma glucose measurement,¹⁸³ by testing the type 3 chi-square for fasting period as a categorical variable in a model also containing fasting period as a continuous variable. Although this showed significant deviation from linearity, this may merely have reflected the large sample size. Given the appearance of a reasonably linear association of plasma glucose with fasting time at fasting periods <8 hours and ≥8 hours when plotted, and the exploratory nature of these analyses, linear regression models were used to calculate residuals of RPG regressed on fasting time in individuals with fasting periods <8 hours and

separately in individuals with fasting periods ≥ 8 hours. Mean RPG levels for the two fasting categories (< 8 hours and ≥ 8 hours) were added to the residuals for the respective fasting category to produce a fasting time-adjusted plasma glucose variable. Based on the population in which analyses of the association of RPG with incident CVDs were undertaken (Section 7.2.4), Cox proportional hazards models were used to estimate HRs and associated CIs for the associations of fasting time-adjusted RPG (< 4.3 , 4.3 - 5.2 , 5.3 - 5.7 , 5.8 - 6.7 , 6.8 - 7.7 , 7.8 - 11.0 , ≥ 11.1 mmol/L) with incident MOVD, MCE, IS and ICH and CVD mortality, excluding age-at-risk < 35 and > 79 years. The Cox proportional hazards models were stratified by age-at-risk, sex and study area and adjusted simultaneously for education (no formal education, primary school, middle school, high school, college/university), smoking (never, occasional, ex-regular, current regular), alcohol (never, occasional, ex-regular, reduced intake, weekly intake), SBP (< 100 , 100 - 109 , 110 - 119 , 120 - 129 , 130 - 139 , 140 - 149 , 150 - 159 , 160 - 169 , ≥ 170 mmHg) and physical activity (< 10 , 10 - 19.9 , 20 - 29.9 , 30 - 39.9 , ≥ 40 MET hours/day). Fasting time-adjusted RPG was also investigated as a continuous variable.

Results

There was evidence of a positive log-linear association of baseline fasting time-adjusted RPG with the risk of CVD mortality and incident MOVD, MCE and IS, with no evidence of a threshold in the association (Figure F.13). There was a weaker positive association with the risk of incident ICH (Figure F.14). When baseline fasting time-adjusted RPG was examined as a continuous variable, each 1 mmol/L increment was associated with a 6% (95% CI 5-7%) increased risk of CVD mortality, a 4% (4-5%) increased risk of incident MOVD and IS, a 5% (4-7%) increased risk of MCE and a 3% (1-4%) increased risk of ICH (Table F.1).

Discussion

The presented analyses demonstrate positive, log-linear associations between fasting time-adjusted RPG levels and risk of CVD mortality, and incident MOVD, MCE, IS and ICH in the CKB. The shape and strength of these associations were comparable with the

associations of baseline RPG not adjusted for fasting time. Thus, accounting for the influence of fasting time on RPG levels in this way makes very little difference to the association of plasma glucose levels with risks of major incident CVDs. Whilst this could reflect an inappropriate assumption of a linear association between RPG levels and self-reported fasting period in fasting categories <8 hours and ≥8 hours, it more likely reflects the limited influence of fasting time on RPG levels in the CKB, as described previously (Section 4.3.3). Data on the nature of food or drink most recently consumed are not available in the CKB, but it is possible that accounting for this, and not just for the time passed since food was last consumed, may have altered the findings.

Conclusion

These analyses demonstrate that adjusting RPG levels for fasting time in the CKB does not appreciably alter the strength or shape of the association between plasma glucose levels and the risk of major incident CVDs. Given the additional complexities involved in these analyses and in interpretation of the association of fasting time-adjusted RPG with incident CVD risk, as well as the comparability of the findings, it was deemed more appropriate not to adjust RPG levels for fasting time in the main analyses.

Figure F.12. Mean baseline plasma glucose levels by self-reported fasting time

Adjusted for age, sex, study area and education. Squares represent mean plasma glucose and vertical lines represent corresponding 95% confidence intervals.

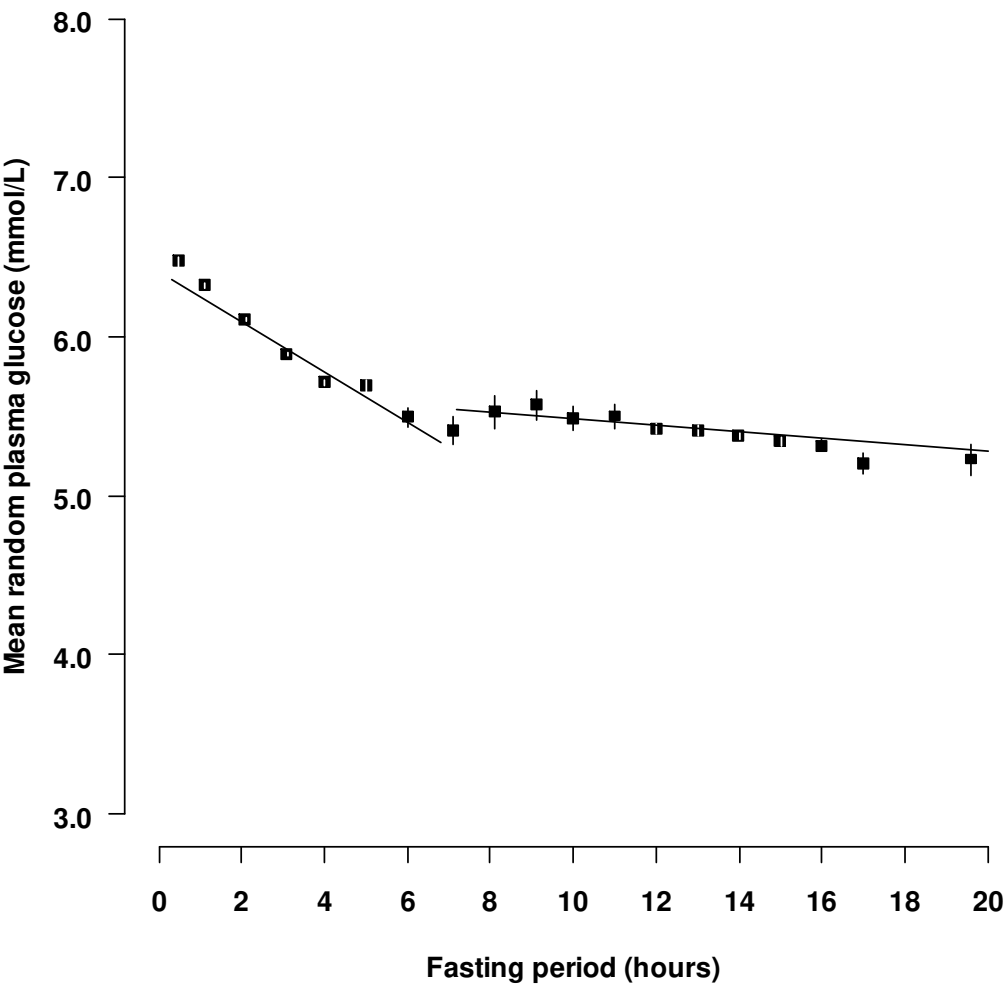


Figure F.13. Adjusted hazard ratios for cardiovascular diseases by fasting time-adjusted baseline random plasma glucose

Stratified by age, sex and study area and adjusted for, education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HRs) are plotted against mean baseline fasting time-adjusted random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs). Numbers above the CIs are the HRs and numbers below the CIs are of cardiovascular disease events.

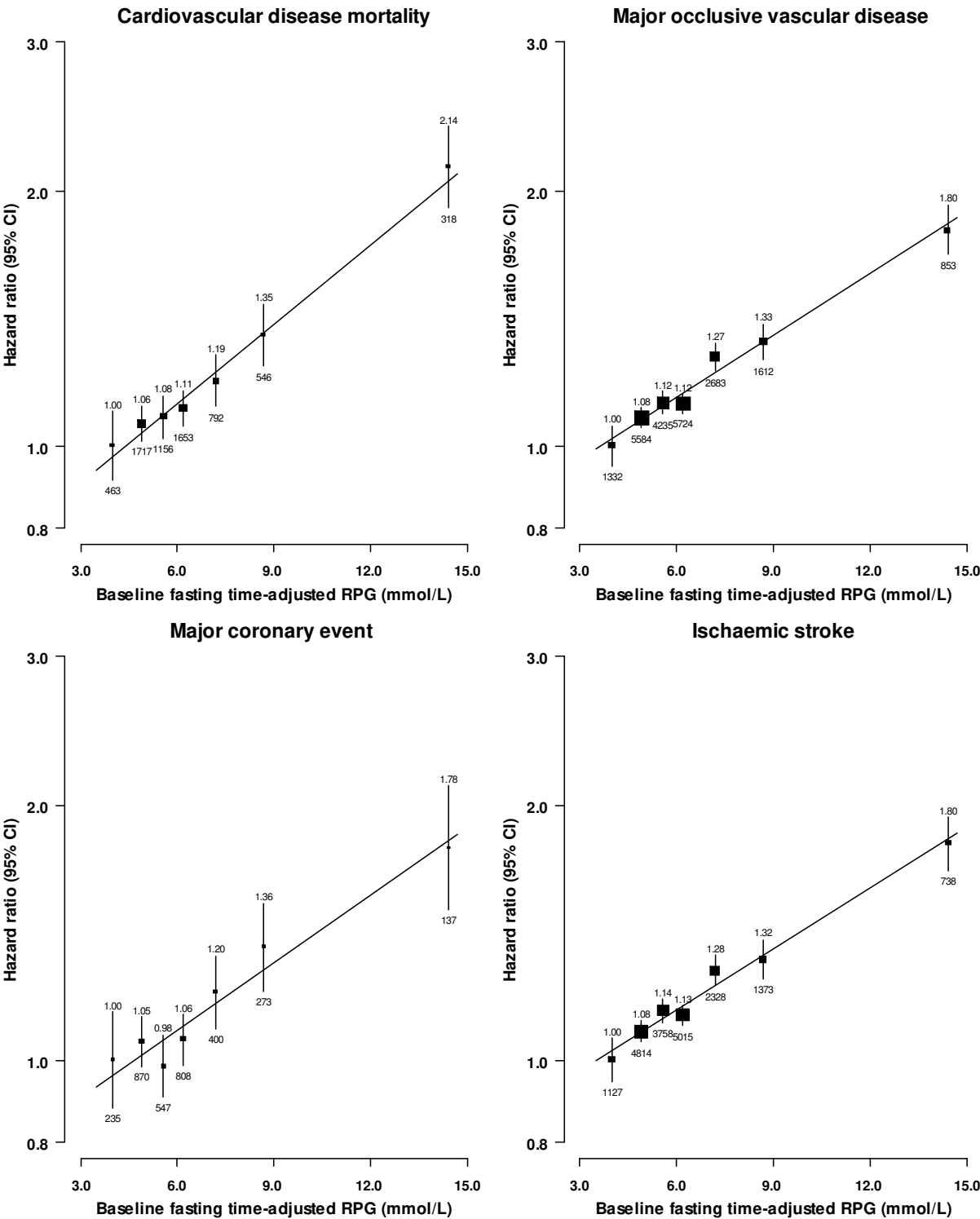


Figure F.14. Adjusted hazard ratios for intracerebral haemorrhage by fasting time-adjusted baseline random plasma glucose

Stratified by age, sex and study area and adjusted for, education, smoking, alcohol, systolic blood pressure and physical activity. Hazard ratios (HRs) are plotted against mean baseline fasting time-adjusted random plasma glucose level in each category. Squares represent the HR with area inversely proportional to the variance. Vertical lines represent the corresponding 95% confidence intervals (CIs). Numbers above the CIs are the HRs and numbers below the CIs are of intracerebral haemorrhage events.

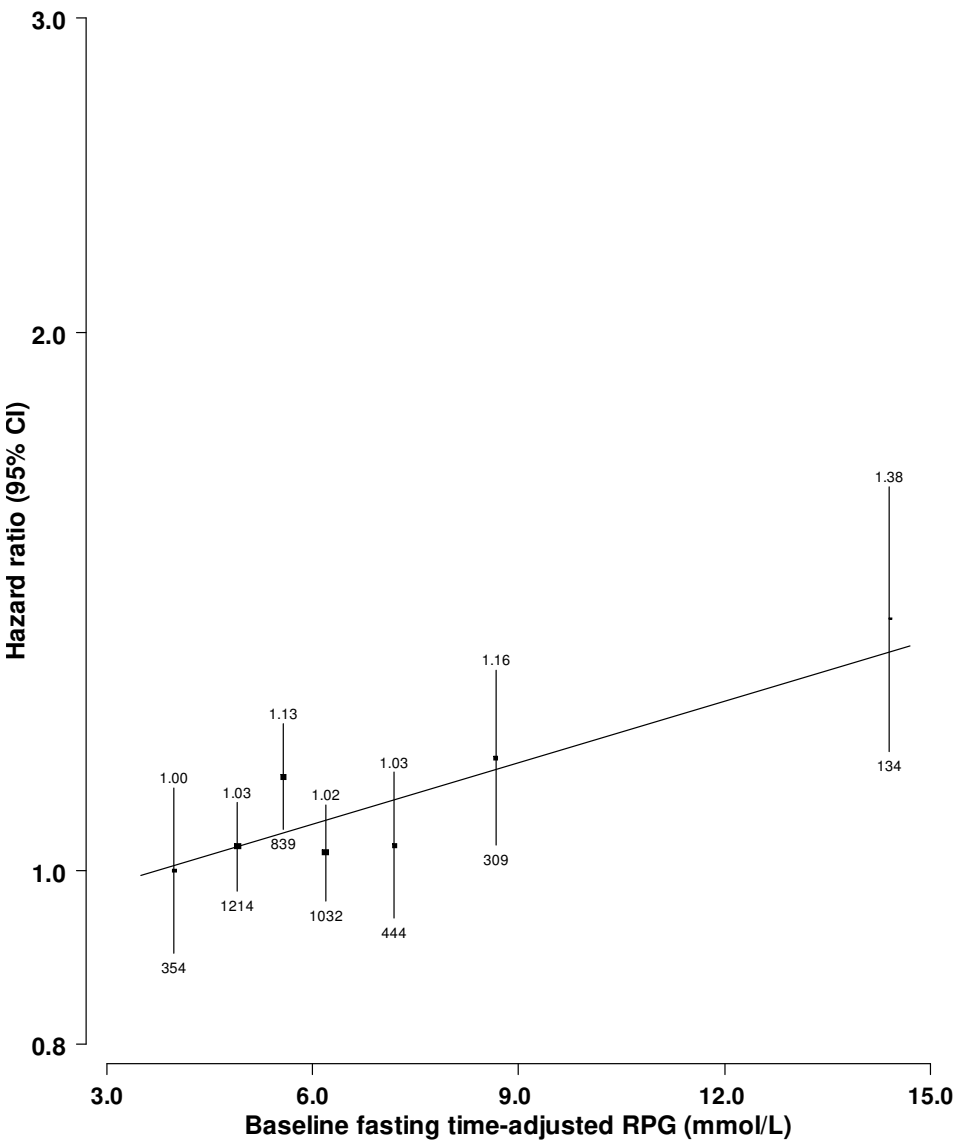


Table F.1. Adjusted hazard ratios for incident cardiovascular diseases by baseline fasting time-adjusted baseline plasma glucose

Disease outcome	Events	Baseline random plasma glucose			
		Not adjusted for fasting time		Adjusted for fasting time	
		HR	95% CI	HR	95% CI
Cardiovascular disease mortality	6645	1.06	1.05-1.07	1.06	1.05-1.07
Major occlusive vascular disease	22023	1.04	1.04-1.05	1.04	1.04-1.05
Major coronary event	3270	1.06	1.04-1.07	1.05	1.04-1.07
Ischaemic stroke	19153	1.04	1.04-1.05	1.04	1.04-1.05
Intracerebral haemorrhage	4326	1.03	1.01-1.04	1.03	1.01-1.04

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, physical activity and systolic blood pressure.

CI, confidence interval; HR, hazard ratio.

Appendix G. Additional methods and results: Assessing the causal association between random plasma glucose and cardiovascular disease: a Mendelian randomisation study

Figure G.1. Participant exclusions for genetic analyses

*Some participants had missing or outlying data for >1 baseline variable.
BMI, body mass index; DBP, diastolic blood pressure; SBP, systolic blood pressure.

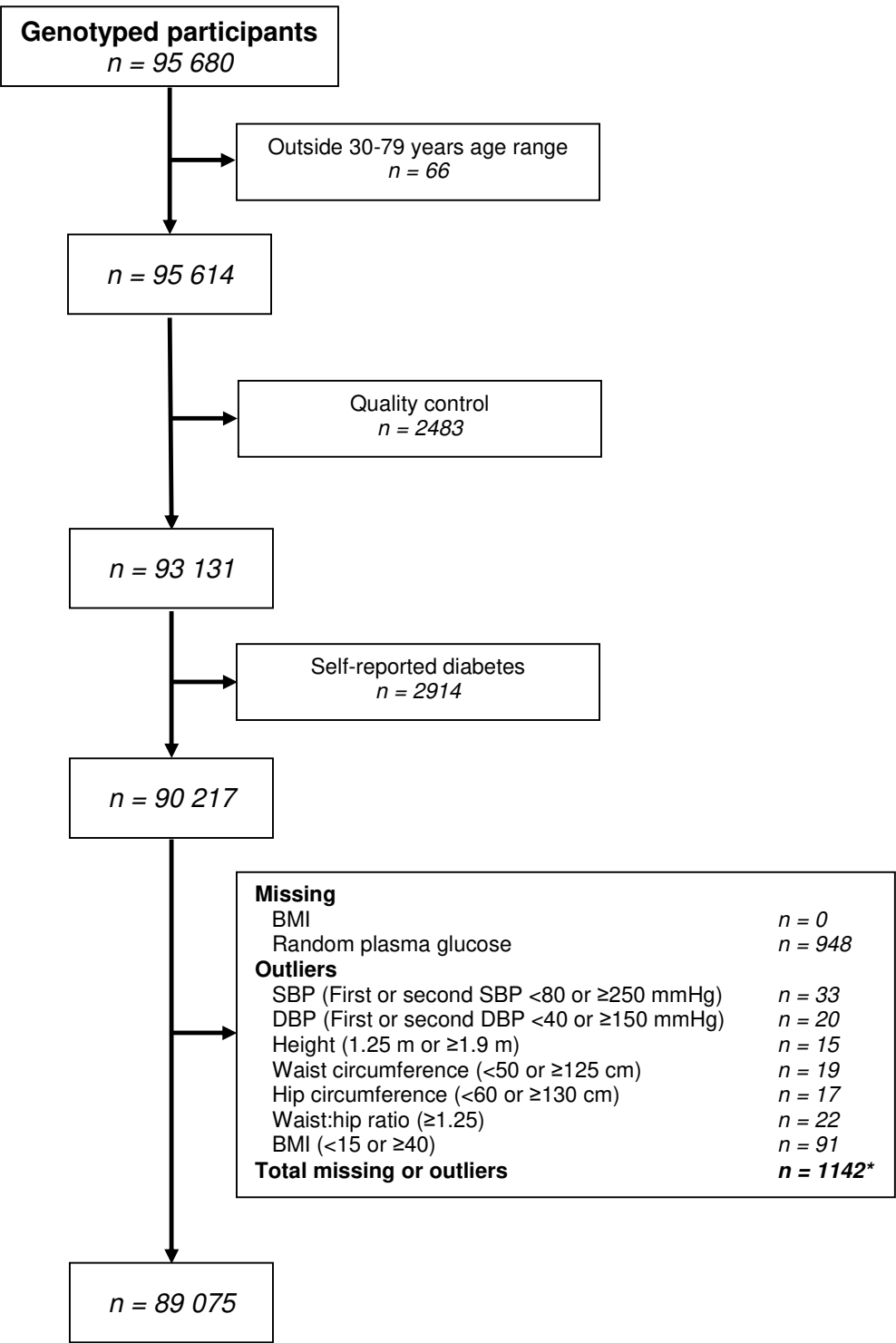


Table G.1. Adjusted hazard ratios for incident cardiovascular diseases per 1 mmol/L higher random plasma glucose in genotyped population (n=89 075)

Disease outcome	Events	Baseline random plasma glucose		Usual random plasma glucose	
		HR	95% CI	HR	95% CI
Major occlusive vascular disease	4697	1.04	1.03-1.05	1.07	1.05-1.09
Major coronary event	761	1.05	1.02-1.08	1.09	1.04-1.14
Ischaemic stroke	4054	1.04	1.03-1.05	1.07	1.04-1.09

Stratified by age, sex and study area and adjusted for education, smoking, alcohol, physical activity and systolic blood pressure.

CI, confidence interval; HR, hazard ratio.

Table G.2. Association of random plasma glucose-related SNPs with baseline characteristics

SNP	No. of participants	% men ^a	Mean (SD) age (years) ^b	% living in urban area	% ≥6 years education	% current regular smokers	% regular alcohol consumption	Mean (SD) physical activity (MET hours/day)	Mean (SD) SBP (mmHg)	Mean (SD) BMI (kg/m ²)	Mean (SD) waist-to-hip ratio
Total	89075	40.5	51.1 (10.6)	40.6	48.1	26.5	14.8	21.8 (14.0)	131.0 (21.0)	23.6 (3.4)	0.88 (0.07)
rs10811661 (<i>CDKN2A/B</i>)											
CC	18638	41.1	51.1 (10.6)	40.3	51.7	26.7	14.8	21.8 (11.8)	130.9 (19.3)	23.7 (3.3) ^e	0.88 (0.06)
CT	44376	40.3	51.1 (10.5)	40.7	52.1	26.4	15.0	21.8 (11.8)	131.0 (19.3)	23.6 (3.2) ^e	0.88 (0.06)
TT	25993	40.6	51.2 (10.5)	40.6	51.9	26.4	14.5	21.7 (11.9)	131.1 (19.4)	23.6 (3.1) ^e	0.88 (0.06)
rs10830963 (<i>MTNR1B</i>)											
GG	29226	40.8	51.0 (10.5)	40.3	51.7	26.3	15.0 ^c	21.8 (11.7)	130.9 (19.3)	23.6 (3.2)	0.88 (0.06)
GC	43523	40.1	51.2 (10.5)	40.6	51.9	26.6	14.6 ^c	21.8 (11.9)	131.0 (19.3)	23.6 (3.2)	0.88 (0.06)
CC	16302	41.2	51.1 (10.5)	40.9	52.3	26.5	14.9 ^c	21.8 (11.9)	131.1 (19.4)	23.6 (3.2)	0.88 (0.06)
rs11708067 (<i>ADCY5</i>)											
GG	3	40.9	53.6 (-)	0.0 ^e	78.0	22.0	22.0	36.6 (-)	122.8 (-)	21.3 (-)	0.72 (-)
GA	623	39.4	51.4 (12.6)	30.9 ^e	48.0	28.0	14.6	20.6 (11.6)	130.3 (20.8)	23.6 (3.2)	0.88 (0.06)
AA	88445	40.5	51.1 (10.5)	40.7 ^e	52.0	26.5	14.8	21.8 (11.8)	131.0 (19.3)	23.6 (3.2)	0.88 (0.06)
rs13266634 (<i>SLC30A8</i>)											
TT	22269	40.6	50.9 (10.5)	41.6 ^d	52.0	26.4	14.8	21.8 (12.0)	131.1 (19.6)	23.6 (3.2)	0.88 (0.06)
TC	37562	40.6	51.1 (10.5)	40.2 ^d	51.9	26.4	14.8	21.8 (11.8)	131.1 (19.2)	23.6 (3.2)	0.88 (0.06)
CC	28680	40.4	51.0 (10.5)	40.3 ^d	52.0	26.7	14.8	21.7 (11.8)	130.9 (19.2)	23.6 (3.2)	0.88 (0.06)
rs1470579 (<i>IGF2BP2</i>)											
AA	6023	40.5	50.9 (10.5)	39.5	51.7	26.9	15.3	21.7 (11.6)	131.3 (19.2)	23.6 (3.2)	0.88 (0.06)
AC	33670	40.1	51.2 (10.5)	40.4	52.1	26.4	14.8	21.8 (11.9)	131.1 (19.5)	23.6 (3.2)	0.88 (0.06)
CC	49359	40.8	51.2 (10.5)	40.8	51.9	26.5	14.8	21.8 (11.8)	130.9 (19.1)	23.6 (3.2)	0.88 (0.06)
rs1552224 (<i>CENTD2</i>)											
CC	74756	40.6	51.1 (10.5)	40.9 ^e	52.0	26.5	14.8	21.8 (11.8)	130.9 (19.2) ^c	23.6 (3.2)	0.88 (0.06) ^c
CA	13624	40.5	51.1 (10.7)	39.1 ^e	51.6	26.4	14.8	21.9 (12.0)	131.4 (19.7) ^c	23.7 (3.2)	0.88 (0.06) ^c
AA	694	37.9	51.2 (11.4)	39.4 ^e	52.7	28.4	15.4	21.7 (13.6)	130.8 (19.7) ^c	23.7 (3.1)	0.88 (0.08) ^c
rs340874 (<i>PROX1</i>)											
TT	33056	40.8	51.1 (10.5)	40.9	51.7	26.4	14.9	21.7 (11.8)	131.0 (19.1)	23.6 (3.2) ^c	0.88 (0.06)
TC	42468	40.6	51.1 (10.5)	40.3	52.1	26.4	14.7	21.9 (11.9)	131.1 (19.4)	23.6 (3.2) ^c	0.88 (0.06)
CC	13528	39.9	51.1 (10.5)	40.7	52.2	26.8	14.9	21.6 (11.8)	130.8 (19.6)	23.6 (3.2) ^c	0.88 (0.06)

Adjusted for age, sex and study area. ^aAdjusted for age and study area; ^bAdjusted for sex and study area; adjusted for age and sex. P for heterogeneity across genotypes >0.05 unless otherwise indicated. ^c0.01 < p ≤ 0.05; ^d0.001 < p ≤ 0.01; ^ep ≤ 0.001.

BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure; SNP, single nucleotide polymorphism.

Table G.2. continued

SNP	No. of participants	% men ^a	Mean (SD) age (years) ^b	% living in urban area	% ≥6 years education	% current regular smokers	% regular alcohol consumption	Mean (SD) physical activity (MET hours/day)	Mean (SD) SBP (mmHg)	Mean (SD) BMI (kg/m ²)	Mean (SD) waist-to-hip ratio
rs4607517 (<i>GCK</i>)											
GG	55566	40.5	51.2 (10.5) ^c	40.4 ^c	51.9	26.4	14.8	21.9 (11.9) ^c	131.0 (19.4)	23.6 (3.2)	0.88 (0.06)
GA	29432	40.7	51.1 (10.5) ^c	40.7 ^c	51.9	26.7	14.9	21.6 (11.7) ^c	131.0 (19.2)	23.6 (3.2)	0.88 (0.06)
AA	4060	39.9	51.0 (10.7) ^c	42.5 ^c	52.1	27.0	14.2	21.8 (11.8) ^c	131.0 (20.1)	23.7 (3.2)	0.88 (0.06)
rs780094 (<i>GCKR</i>)											
CC	21312	40.5	51.1 (10.6)	42.7 ^e	52.2 ^d	26.6	15.4	21.8 (12.2)	130.7 (19.6) ^e	23.6 (3.2)	0.88 (0.06)
CT	44069	40.2	51.1 (10.5)	40.0 ^e	52.0 ^d	26.4	14.8	21.8 (11.8)	130.9 (19.2) ^e	23.6 (3.2)	0.88 (0.06)
TT	23683	41.1	51.2 (10.7)	39.7 ^e	51.6 ^d	26.4	14.4	21.7 (11.6)	131.4 (19.7) ^e	23.6 (3.2)	0.88 (0.06)
rs7901695 (<i>TCF7L2</i>)											
TT	83053	40.5	51.1 (10.4)	40.8 ^e	52.0	26.4	14.8	21.8 (11.8)	131.0 (19.3)	23.6 (3.2) ^c	0.88 (0.06)
TC	5179	40.6	51.1 (10.8)	37.6 ^e	51.6	27.4	14.3	21.6 (11.9)	130.4 (19.2)	23.6 (3.3) ^c	0.88 (0.06)
CC	114	41.0	49.3 (9.4)	36.3 ^e	50.4	24.9	19.4	21.9 (9.7)	130.8 (9.7)	23.0 (1.7) ^c	0.87 (0.04)

Adjusted for age, sex and study area. ^aAdjusted for age and study area; ^bAdjusted for sex and study area; adjusted for age and sex. P for heterogeneity across genotypes >0.05 unless otherwise indicated. ^c0.01 < p ≤ 0.05; ^d0.001 < p ≤ 0.01; ^ep ≤ 0.001.

BMI, body mass index; MET, metabolic equivalent of task; SBP, systolic blood pressure; SNP, single nucleotide polymorphism.

Table G.3. Power estimates for detection of an association between genetically determined random plasma glucose and risk of incident cardiovascular diseases^{329, 330}

Disease outcome	HR per 1 mmol/L higher usual random plasma glucose	Power (%)
Major occlusive vascular disease		
<i>Participants= 84 858</i>	1.04	4.1
<i>Events= 4804</i>	1.08	6.4
	1.12	9.4
	1.16	13.2
	1.20	17.7
	1.24	23.0
Major coronary event		
<i>Participants= 80 852</i>	1.04	3.1
<i>Events= 798</i>	1.08	3.8
	1.12	4.6
	1.16	5.4
	1.20	6.4
	1.24	7.4
Ischaemic stroke		
<i>Participants= 84 188</i>	1.04	4.0
<i>Events= 4130</i>	1.08	6.0
	1.12	8.6
	1.16	11.8
	1.20	15.7
	1.24	20.1

Power calculations assumed a coefficient of determination (r^2) for the association between the random plasma glucose genetic risk score and random plasma glucose of 0.002 and a 0.05 significance level.

HR, hazard ratio.

Table G.4. Effect allele frequencies by study area

SNP	Nearby gene	Alleles		Total	Qingdao	Harbin	Haikou	Suzhou	Liuzhou	Sichuan	Gansu	Henan	Zhejiang	Hunan
		Effect	Other											
rs10811661	<i>CDKN2A/B</i>	C	T	0.541	0.522	0.526	0.589	0.517	0.581	0.559	0.544	0.528	0.521	0.553
rs10830963	<i>MTNR1B</i>	C	G	0.427	0.420	0.418	0.436	0.425	0.450	0.429	0.423	0.414	0.427	0.441
rs11708067	<i>ADCY5</i>	A	G	0.996	0.995	0.995	0.999	0.999	0.998	0.998	0.992	0.994	0.999	0.998
rs13266634	<i>SLC30A8</i>	T	C	0.536	0.553	0.562	0.488	0.541	0.494	0.519	0.557	0.556	0.546	0.513
rs1470579	<i>IGF2BP2</i>	A	C	0.257	0.258	0.263	0.249	0.265	0.232	0.252	0.270	0.261	0.257	0.251
rs1552224	<i>CENTD2</i>	A	C	0.916	0.904	0.908	0.939	0.913	0.938	0.924	0.897	0.905	0.920	0.927
rs340874	<i>PROX1</i>	T	C	0.390	0.371	0.369	0.402	0.382	0.425	0.412	0.388	0.379	0.377	0.402
rs4607517	<i>GCK</i>	A	G	0.211	0.234	0.230	0.187	0.230	0.174	0.188	0.214	0.219	0.225	0.197
rs780094	<i>GCKR</i>	T	C	0.487	0.447	0.463	0.584	0.457	0.571	0.503	0.472	0.463	0.459	0.510
rs7901695	<i>TCF7L2</i>	C	T	0.031	0.036	0.034	0.020	0.027	0.023	0.030	0.036	0.035	0.026	0.033

P for heterogeneity across study areas <0.001 for all SNPs.

SNP, single nucleotide polymorphism.