Appendix A: CLSA analysis methods

Data source and sample overview

We used data from baseline (2011-2015), first follow-up (2015-2018), and second follow-up (2018-2021) of the Canadian Longitudinal Study on Aging (CLSA) Comprehensive cohort. CLSA is an ongoing nationwide cohort study, with 51,338 participants from 10 Canadian provinces aged 45-85 years old at the time of recruitment (2011-2015). Participants provided data on demographic, biological, medical, psychosocial, economic, lifestyle, and behavioural factors through either a computer-assisted telephone survey or an in-person interview. Of the 51,338 participants, a subset of 30,097 (Comprehensive cohort) living within 25 to 50 km of one of eleven data collection sites located in seven provinces were invited to complete detailed physical assessments. From this cohort, we assessed a sample of 26,093 participants who had blood-biomarker data at baseline and stroke or transient ischaemic attack (TIA) data at the second follow-up. A detailed description of CLSA recruitment and data collection procedures has been previously described elsewhere [1, 2].

Blood-based biomarkers

We used data on eight blood biomarkers assessed at the CLSA baseline visit (2011-2015), including HbA1c, total cholesterol, high-density lipoprotein (HDL-C), low-density lipoprotein (LDL-C), triglycerides, interleukin-6 (IL-6), tumour necrosis factor (TNF), and high-sensitivity C-reactive protein (hs-CRP). All biomarker values used in analyses therefore reflect exposure status at baseline, prior to stroke/TIA ascertainment at follow-up 2.

Biological specimens, including blood samples, were collected at participants' visits to a CLSA data collection site by trained phlebotomists following standardised protocols. Blood was collected into ethylenediaminetetraacetic acid (EDTA) tubes for plasma and serum separator tubes as appropriate to each assay. Plasma and serum were separated by centrifugation and stored at $-80^{\circ}C$ until analysis. HbA1c was measured using high-performance liquid chromatography (HPLC); lipid fractions (total cholesterol, HDL-C, LDL-C, triglycerides) were measured by enzymatic colorimetric methods; IL-6 and TNF were measured by electrochemiluminescence immunoassay; and hs-CRP was measured by high-sensitivity immunoturbidimetric assay. Further details of the CLSA biospecimen collection and laboratory procedures are described elsewhere [1].

Incident stroke/TIA

Participants reporting a physician-diagnosed stroke or TIA at baseline were excluded from the analytical sample. Incident stroke/TIA was ascertained at the first and second follow-up using self-report. Stroke was defined by the question: "Has a doctor ever told you that you have experienced a stroke or cerebrovascular accident? (CVA)?" Likewise, TIA was defined by the question: "Has a doctor ever told you that you have experienced a mini-stroke or TIA?" Participants who responded yes to either question at follow-up 1 or follow-up 2 (but not at baseline) were classified as incident stroke/TIA cases. Stroke and TIA were combined into a single outcome due to the relatively small number of events in the analytical sample, which precluded reliable subtype-specific analyses.

Covariates

Sociodemographic and clinical covariates were selected based on established associations with stroke risk in the literature. Selected covariates included age, biological sex, educational level, annual household income, race/ethnicity, residence (i.e., urban vs. rural), body mass index (BMI), heart disease, and diabetes. Due to the small representation of individual non-white racial/ethnic groups in the CLSA cohort ($<6\%$), race/ethnicity was categorised as white versus non-white to ensure adequate cell sizes and robust statistical analyses.

Statistical analysis

Descriptive statistics were used to summarise characteristics of the study samples stratified by stroke/TIA status. Proportions and means with standard deviations (SDs) using sampling (inflation) weights were used to summarise continuous and categorical variables, respectively. Baseline participant characteristics and biomarker concentrations were compared between incident stroke/TIA cases and non-cases using Wilcoxon rank-sum test for complex survey samples (continuous variables) and the chi-squared test with Rao & Scott's second-order correction (categorical variables). All descriptive analyses incorporated CLSA sampling weights.

Sampling weights provided by the CLSA were used. The CLSA calculates these weights to account for the study's sampling design and for non-response, and calibrates them to national census-based population estimates so that weighted estimates reflect the corresponding Canadian population aged 45-85 years. Two types of weights are available. Inflation weights, which sum to the size of the target population, were applied to the descriptive statistics, whereas analytic weights, which are proportional to the inflation weights but rescaled to the sample size, were applied to the logistic regression models.

Binomial logistic regression models using analytic weights were used to assess the association between each biomarker level and 6-year incident stroke/TIA (binary outcome). A crude (unadjusted) model was first estimated for each biomarker. An adjusted model was then estimated, controlling for age, sex, race/ethnicity, heart disease, stroke, diabetes, alcohol consumption, smoking status, physical activity, BMI, income, education, and urban/rural residence. All biomarkers were modelled as continuous variables, with units as reported in Table 1 and Appendix B. Sex-stratified analyses were conducted by repeating these models separately for males and females. Results are reported as odds ratios (OR) with 95% confidence interval (CI) estimates. Analyses were conducted using R (R Core Team, 2022).

References

1. Raina PS, Wolfson C, Kirkland SA, Griffith LE, Oremus M, Patterson C, et al. The Canadian Longitudinal Study on Aging (CLSA). Canadian Journal on Aging/La

Revue canadienne du vieillissement 2009 Sept.;28:221–9. https://doi.org/10.1017/S0714980809990055

2. Raina P, Wolfson C, Kirkland S, Griffith LE, Balion C, Cossette B, et al. Cohort Profile: The Canadian Longitudinal Study on Aging (CLSA). International Journal of Epidemiology 2019 Dec.;48:1752–1753j. https://doi.org/10.1093/ije/dyz173

Appendix B: CLSA results tables

Table B1: Sociodemographic characteristics of sample cohort (N=26,093); Unweighted count (Weighted prevalence)

Characteristic	Overall, N=26,156	Stroke/TIA, N=822	No stroke/TIA, N=25,334	p-value[1]
Age	61 (54, 69)	70 (62,78)	61 (54, 69)	**<0.001**
Sex				0.013
Female	13,366 (53%)	385 (48%)	12,981 (53%)	
Male	12,790 (47%)	437 (52%)	12,353 (47%)	
Race				0.300
White	25,065 (94%)	794 (93%)	24,271 (94%)	
Non-white	1,091 (5.8%)	28 (6.7%)	1,063 (5.8%)	
Education				**<0.001**
Less than secondary school graduation	1,244 (15%)	67 (27%)	1,177 (15%)	
Secondary school graduation, no post-secondary education	2,350 (11%)	86 (11%)	2,264 (11%)	
Some post-secondary education	1,879 (9.1%)	<80 (<15%)	1,804 (9.0%)	
Post-secondary degree/diploma	20,644 (64%)	592 (51%)	20,052 (65%)	
Missing	39 (0.1%)	<10 (<0.1%)	37 (0.1%)	
Marital status				**<0.001**
Single	2,272 (8.5%)	59 (13%)	2,213 (8.4%)	
Married/Common-law relationship	18,331 (75%)	534 (66%)	17,797 (76%)	
Widowed/Separated/Divorced	<5,550 (16%)	229 (22%)	<5,320 (16%)	
Missing	<10 (<0.1%)	-	<10 (<0.1%)	
Urban/rural classification				0.100
Rural	2,136 (5.2%)	55 (4.8%)	2,080 (5.3%)	
Urban	23,705 (94%)	751 (94%)	22,954 (94%)	
Missing	315 (1.0%)	15 (1.4%)	300 (1.0%)	
Household income				**<0.001**
< $50,000	6,332 (25%)	290 (43%)	6,042 (24%)	
$50,000-$100,000	13,791 (52%)	408 (39%)	13,383 (52%)	
≥$100,000	4,464 (17%)	74 (10%)	4,390 (18%)	
Missing	1,569 (5.9%)	50 (6.9%)	1,519 (5.9%)	
Co-resident status				**<0.001**
Living along	5,602 (15%)	235 (22%)	3,367 (15%)	
Living with 1 co-resident	12,719 (44%)	445 (55%)	12,274 (44%)	
Living with 2+ co-residents	7,820 (40%)	142 (23%)	7,678 (41%)	
Missing	15 (<0.1%)	-	15 (<0.1%)	

HbA1c: glycated haemoglobin, TIA: transient ischaemic attack.
[1]p-values presented to 3 decimal places. Wilcoxon rank-sum test for complex survey samples; chi-squared test with Rao & Scott's second-order correction. Bold values indicate p<0.05.

Table B2: Clinical characteristics of sample cohort (N = 26,093); unweighted count (weighted prevalence)

Characteristic	Overall, N=26,156	Stroke/TIA, N=822	No stroke/TIA, N=25,334	p-value[1]
Chronic heart disease				**<0.001**
Yes	2,653 (9.4%)	184 (25%)	2,469 (8.9%)	
No	23,369 (90%)	624 (73%)	22,745 (91%)	
Missing	134 (0.4%)	14 (1.0%)	120 (0.4%)	
Diabetes				0.1
Yes	4,291 (16%)	184 (20%)	4,107 (16%)	
No	21,780 (84%)	<630 (79%)	21,151 (84%)	
Missing	85 (0.5%)	<10 (0.4%)	76 (0.5%)	
Angina				**<0.001**
Yes	993 (3.4%)	66 (10%)	927 (3.2%)	
No	25,042 (96%)	741 (88%)	24,301 (96%)	
Missing	121 (0.6%)	15 (1.8%)	106 (0.5%)	
Myocardial infarction				**<0.001**
Yes	1,052 (3.5%)	76 (11%)	976 (3.3%)	
No	25,004 (96%)	730 (88%)	24,274 (96%)	
Missing	100 (0.4%)	16 (1.4%)	84 (0.4%)	
Alzheimer's/Dementia				0.5
Yes	28 (<0.1%)	0 (0%)	28 (0.1%)	
No	26,068 (<100%)	<820 (<100%)	25,254 (<100%)	
Missing	60 (0.2%)	<10 (<0.5%)	52 (0.2%)	
Parkinson's Disease				0.5
Yes	92 (0.3%)	<10 (<0.5%)	89 (0.3%)	
No	26,013 (99%)	<820 (<100%)	25,202 (99%)	
Missing	51 (0.2%)	<10 (<0.5%)	43 (0.2%)	
Memory problems				**0.010**
Yes	333 (1.5%)	26 (3.5%)	307 (1.4%)	
No	25,751 (98%)	786 (96%)	24,965 (98%)	
Missing	72 (0.3%)	10 (0.6%)	62 (0.3%)	
Mood disorder				**<0.001**
Yes	4,366 (18%)	<180 (<30%)	4,187 (17%)	
No	21,703 (82%)	634 (70%)	21,069 (82%)	
Missing	87 (0.4%)	<10 (<0.7%)	78 (0.4%)	
Clinical depression				**0.003**
Yes	4,179 (16%)	166 (26%)	4,013 (16%)	
No	21,866 (83%)	646 (73%)	21,220 (84%)	
Missing	111 (0.4%)	10 (0.6%)	101 (0.4%)	
Self-rated mental health				**<0.001**
Excellent/Very good/Good	24,843 (95%)	749 (87%)	24,094 (95%)	
Fair/Poor	1,297 (5.3%)	73 (13%)	1,224 (5.1%)	
Missing	16 (<0.1%)	0 (0%)	16 (<0.1%)	
Functional impairment				**<0.001**
No functional impairment	23,931 (92%)	686 (86%)	23,245 (93%)	
Mild/moderate/severe functional impairment	2,142 (6.9%)	132 (13%)	2,010 (6.7%)	
Total functional impairment	10 (<0.1%)	<10 (<0.5%)	<10 (<0.5%)	
Missing	73 (0.6%)	<10 (<0.5%)	71 (0.6%)	

TIA: transient ischaemic attack.

[1] p-values presented to 3 decimal places. Wilcoxon rank-sum test for complex survey samples; chi-squared test with Rao & Scott's second-order correction. Bold values indicate p<0.05.

Table B3: Binomial logistic regression models demonstrating crude and adjusted association between blood biomarkers and incident stroke or transient ischaemic attack; weighted

	Unadjusted OR [95% CI	Adjusted OR [95% CI]
HbA1c (per unit)	1.26 [1.15, 1.37]	1.17 [1.02, 1.34]
Total cholesterol (mmol/L)	0.88 [0.79, 0.99]	1.06 [0.94, 1.19]
Low-density lipoprotein (mmol/L)	0.83 [0.73, 0.95]	1.03 [0.89, 1.19]
High-density lipoprotein (mmol/L)	0.83 [0.63, 1.10]	0.91 [0.67, 1.24]
Triglycerides (mmol/L)	1.10 [1.01, 1.20]	1.11 [1.02, 1.21]
Interleukin-6 (pg/mL)	1.17 [1.09, 1.25]	1.07 [0.97, 1.17]
TNF (pg/mL)	1.21 [1.06, 1.38]	1.20 [1.05, 1.38]
hs-CRP (pg/mL)	1.01 [0.99, 1.03]	1.01 [0.99, 1.03]

CI: confidence interval, HbA1c: glycated haemoglobin, hs-CRP: high-sensitivity C-reactive protein, OR: odds ratio, TNF: tumour necrosis factor.
All biomarkers modelled as continuous variables per unit increase. Adjusted model controls for age, sex, race/ethnicity, heart disease, diabetes, alcohol consumption, smoking status, physical activity, body mass index, household income, education, and urban/rural residence. Rows ordered by biomarker category: metabolic markers (HbA1c, cholesterol fractions, triglycerides) followed by inflammatory markers (interleukin-6, TNF, hs-CRP).

Table B4: Sex-stratified binomial logistic regression models demonstrating crude and adjusted associations between blood biomarkers and incident stroke or transient ischaemic attack; weighted

	Males		Females	
	Unadjusted OR [95% CI]	Adjusted OR [95% CI]	Unadjusted OR [95% CI]	Adjusted OR [95% CI]
HbA1c	1.29 [1.16, 1.44]	1.22 [1.04, 1.42]	1.19 [1.03, 1.39]	1.07 [0.81, 1.40]
Total cholesterol	0.85 [0.73, 0.99]	1.06 [0.90, 1.25]	0.94 [0.80, 1.11]	1.05 [0.89, 1.25]
Low-density lipoprotein	0.76 [0.65, 0.90]	1.01 [0.83, 1.23]	0.92 [0.74, 1.13]	1.05 [0.85, 1.30]
High-density lipoprotein	0.77 [0.52, 1.14]	0.79 [0.51, 1.23]	0.93 [0.62, 1.40]	0.99 [0.65, 1.52]
Triglycerides	1.08 [0.97, 1.19]	1.11 [0.99, 1.24]	1.12 [0.97, 1.28]	1.11 [0.96, 1.28]
Interleukin-6	1.19 [1.08, 1.32]	1.10 [0.95, 1.29]	1.14 [1.04, 1.25]	1.04 [0.94, 1.16]
TNF	1.31 [1.12, 1.53]	1.30 [1.07, 1.58]	1.11 [0.88, 1.39]	1.13 [0.92, 1.39]
hs-CRP	1.02 [1.00, 1.04]	1.02 [0.99, 1.04]	0.99 [0.96, 1.02]	0.99 [0.96, 1.02]

CI: confidence interval, HbA1c: glycated haemoglobin, hs-CRP: high-sensitivity C-reactive protein, OR: odds ratio, TNF: tumour necrosis factor.
All biomarkers modelled as continuous variables per unit increase. Adjusted model controls for age, sex, race/ethnicity, heart disease, diabetes, alcohol consumption, smoking status, physical activity, body mass index, household income, education, and urban/rural residence. Rows ordered by biomarker category: metabolic markers (HbA1c, cholesterol fractions, triglycerides) followed by inflammatory markers (interleukin-6, TNF, hs-CRP).

Appendix C: results from UK biobank studies

Table C1: Multivariable-adjusted hazard ratios (HRs) [95% CI] for the association of hs-CRP and incident ischaemic stroke and transient ischaemic attack

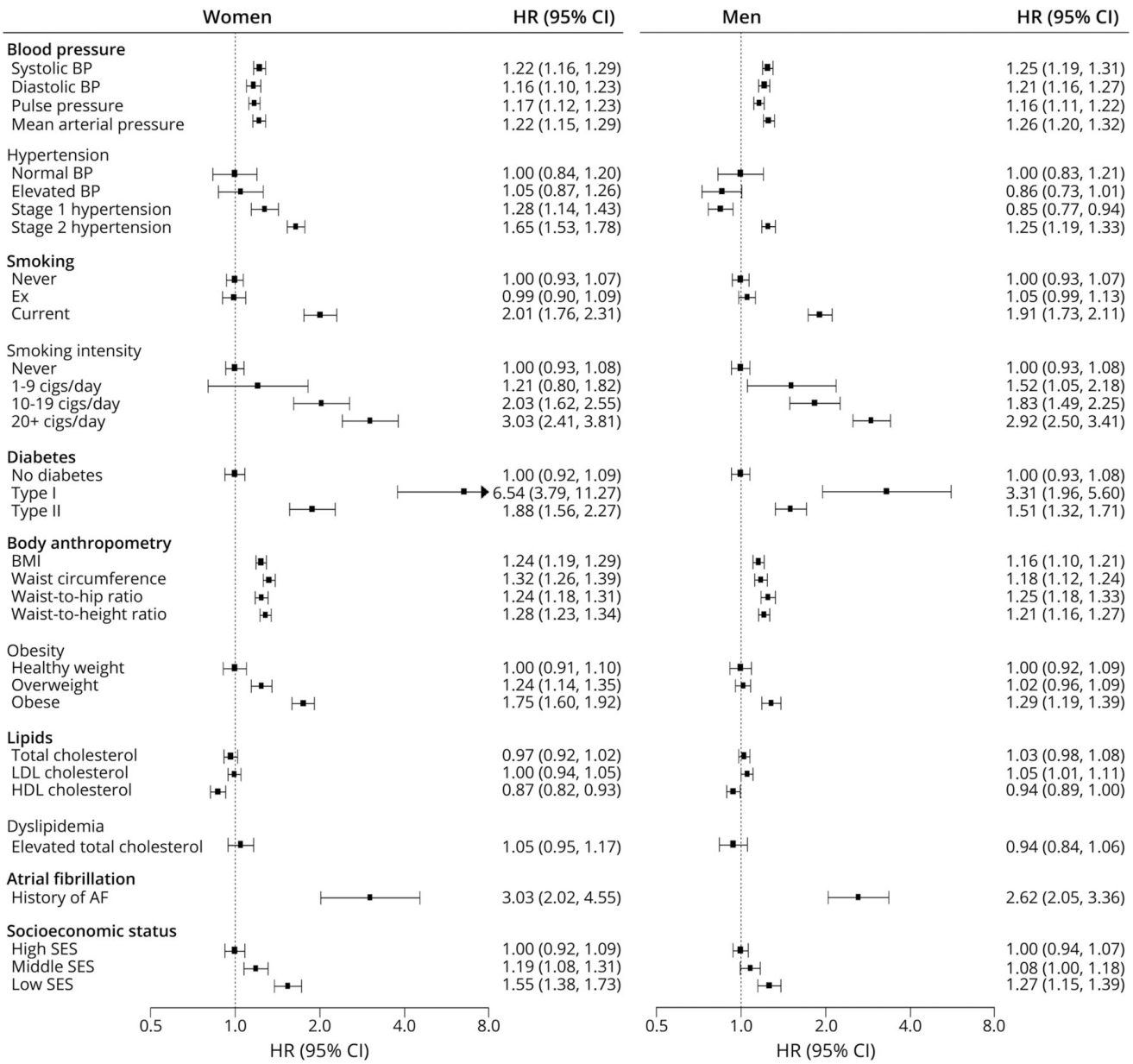
hs-CRP concentration		Incident retinal artery occlusion (n=136)		Incident ischaemic stroke (n=3206)		Incident transient ischaemic attack (n = 1489)	
		Model 1 HR (95% CI)	Model 2 HR (95% CI)	Model 1 HR (95% CI)	Model 2 HR (95% CI)	Model 1 HR (95% CI)	Model 2 HR (95% CI)
Per 10 mg/L		**1.32 (1.03 to 1.69)**	**1.43 (1.01 to 1.76)**	**1.35 (1.29 to 1.42)**	**1.24 (1.17 to 1.33)**	**1.25 (1.15 to 1.36)**	**1.12 (1.01 to 1.25)**
Category	Low	Reference	Reference	Reference	Reference	Reference	Reference
	Median	**1.89 (1.20 to 2.96)**	1.52 (0.91 to 2.52)	**1.41 (1.28 to 1.55)**	1.09 (0.98 to 1.21)	**1.40 (1.22 to 1.60)**	1.15 (0.99 to 1.34)
	High	**1.80 (1.14 to 2.83)**	1.32 (0.77 to 2.29)	**2.01 (1.84 to 2.20)**	**1.48 (1.33 to 1.64)**	**1.81 (1.59 to 2.06)**	**1.40 (1.20 to 1.64)**

Model 1 is the crude model; model 2 is adjusted for age, gender, ethnicity, education, Townsend deprivation index, physical activity, body mass index, cigarette smoking, alcohol drinking, family history of stroke, diabetes mellitus, hypertension and hyperlipidaemia. Bolded items are statistically significant ($p<0.05$).

hs-CRP, high-sensitivity C-reactive protein.

Note: From *High-sensitivity C-reactive protein and risk of retinal artery occlusion and ischaemic stroke: a cross- cohort study*, by Wang et al., 2025, Br J Ophthalmol, 0, p. 5 (doi:10.1136/bjo-2023-325044). Copyright © 2025 by BMJ Group.

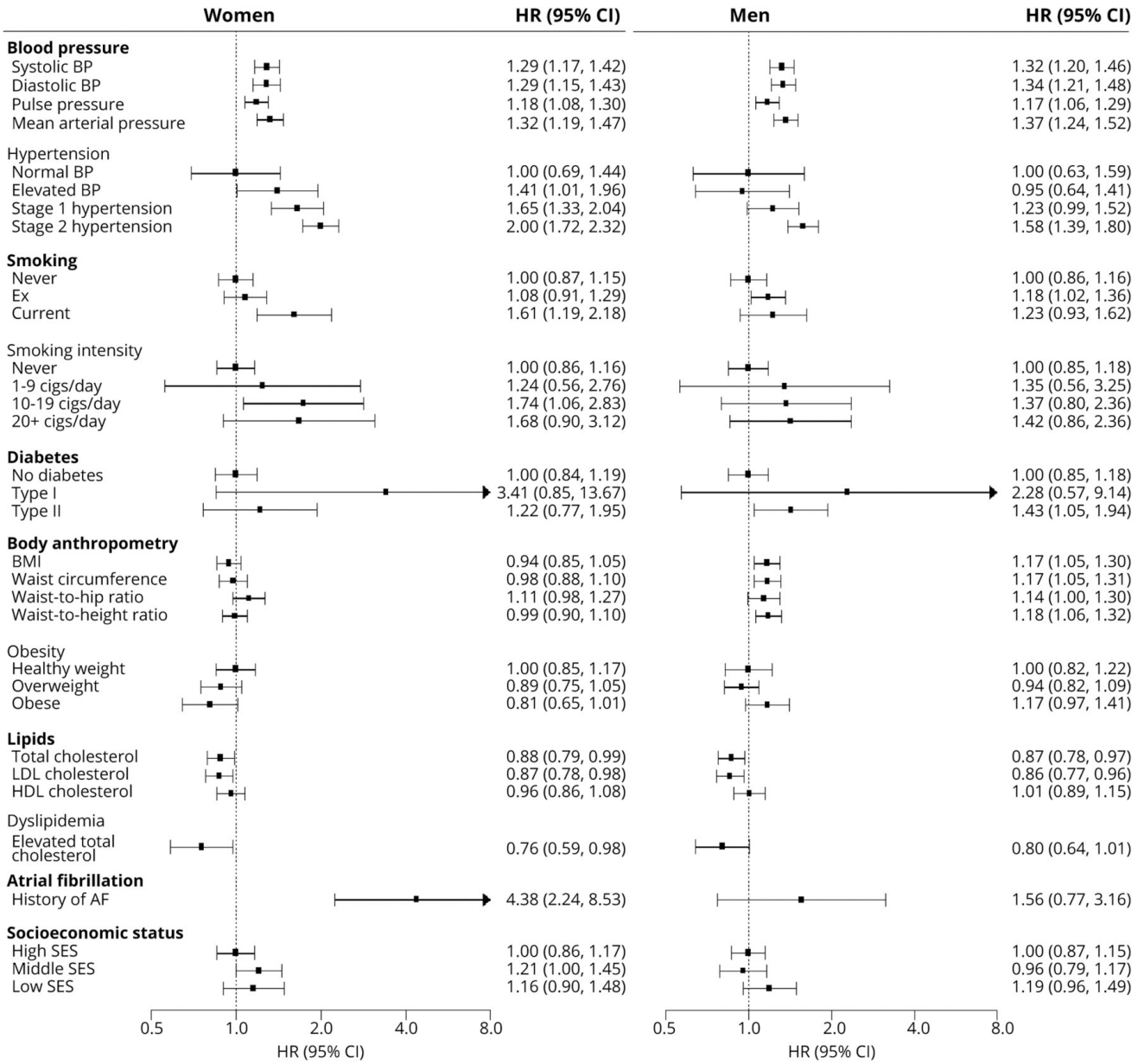
Figure C1: Adjusted HRs for incident ischaemic stroke associated with risk factors, by sex



Horizontal lines indicate the corresponding 95% confidence intervals (CIs) around the hazard ratios (HRs). AF = atrial fibrillation; BMI = body mass index; BP = blood pressure; HDL = high-density lipoprotein; LDL = low-density lipoprotein; SES = socioeconomic status.

Note: From *Sex differences in the association between major risk factors and the risk of stroke in the UK Biobank cohort study*, by Peters et al., 2020, Neurology, 95(20), p. e2719 (doi: 10.1212/WNL.0000000000010982). Copyright © 2020 by Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology.

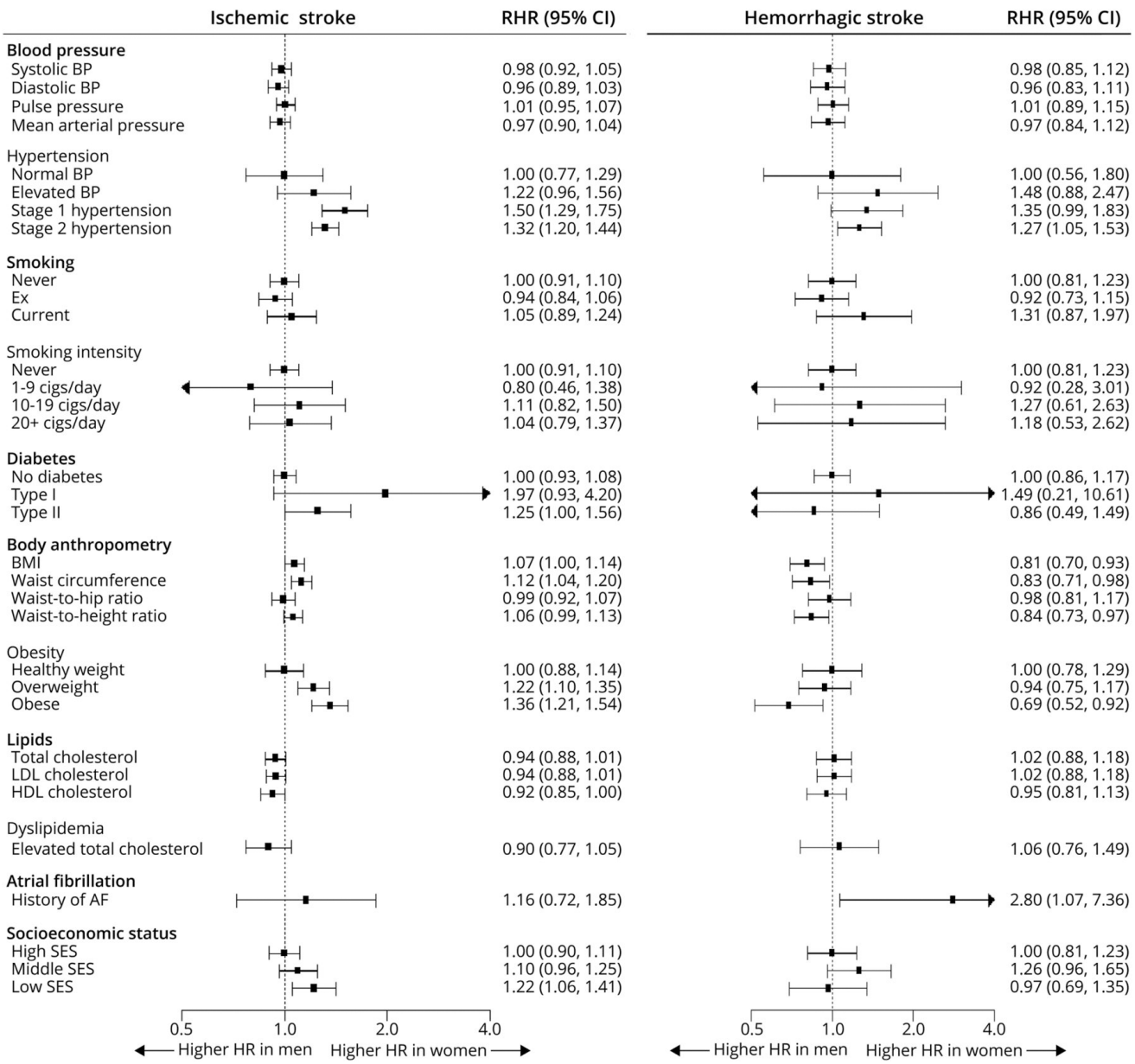
Figure C2: Adjusted HRs for incident haemorrhagic stroke associated with risk factors, by sex



Horizontal lines indicate the corresponding 95% confidence intervals (CIs) around the hazard ratios (HRs). AF = atrial fibrillation; BMI = body mass index; BP = blood pressure; HDL = high-density lipoprotein; LDL = low-density lipoprotein; SES = socioeconomic status.

Note: From *Sex differences in the association between major risk factors and the risk of stroke in the UK Biobank cohort study*, by Peters et al., 2020, Neurology, 95(20), p. e2720 (doi: 10.1212/WNL.0000000000010982). Copyright © 2020 by Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology.

Figure C3: Adjusted women-to-men RHRs for ischaemic and haemorrhagic stroke associated with risk factors



Horizontal lines indicate the corresponding 95% confidence intervals (CIs) around the ratio of hazard ratios (RHRs). AF = atrial fibrillation; BMI = body mass index; BP = blood pressure; HDL = high-density lipoprotein; LDL = low-density lipoprotein; SES = socioeconomic status.

Note: From *Sex differences in the association between major risk factors and the risk of stroke in the UK Biobank cohort study*, by Peters et al., 2020, Neurology, 95(20), p. e2719 (doi: 10.1212/WNL.0000000000010982). Copyright © 2021 by Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology.