Supplementary table 1. Risk factors and stroke etiology according to sex, gender and age-group.

	All (n=437)	Men (n=273)	Women (n=164)	p-value	18-34 (n=65)	35-44 (n=103)	45-54 (n=269)	p-value
Age, median (IQR)	47 (39-51)	48 (42-52)	45 (37-50.5)	0.001				
NIHSS 0, median (IQR)	3 (2-6)	4 (2-7)	3 (1-5)	0.05	2 (1-5)	3 (1-5.5)	4 (2-7.5)	0.004
NIHSS 7, median (IQR)	1 (0-3)	1 (0-3)	1 (0-2)	0.81	0 (0-1)	0.5 (0-2.5)	1 (0-3)	0.0023
Well-documented risk factors, n (%)								
Hypertension	226 (51.7)	153 (56.0)	73 (44.5)	0.02	10 (15.4)	45 (43.7)	171 (63.6)	<0.001
Diabetes	54 (12.4)	37 (13.6)	17 (10.4)	0.33	0 (0)	9 (8.7)	45 (16.7)	0.001
Obesity	137 (33.2)	82 (32.2)	55 (34.8)	0.58	11 (17.5)	29 (30.5)	97 (38.0)	0.007
Dyslipidemia	278 (64.4)	185 (68.5)	93 (57.4)	0.02	25 (38.5)	67 (67.7)	186 (69.4)	<0.001
Cigarette smoking	239 (56.4)	172 (65.2)	67 (41.9)	<0.001	30 (46.9)	53 (54.1)	156 (59.5)	0.16
Ischemic heart disease	29 (6.6)	22 (8.1)	7 (4.3)	0.12	0 (0)	8 (7.8)	21 (7.8)	0.03
Atrial fibrillation	33 (7.6)	29 (10.6)	4 (2.4)	0.002	1 (1.5)	4 (3.9)	28 (10.4)	0.01

Heart failure	27 (6.2)	19 (7)	8 (4.9)	0.38	1 (1.5)	5 (4.9)	21 (7.8)	0.15
Transient ischemic attack	34 (7.8)	20 (7.3)	14 (8.5)	0.63	4 (6.2)	7 (6.8)	23 (8.6)	0.94
Peripheral artery disease	7 (1.6)	7 (2.6)	0 (0)	0.05	0 (0)	1 (1)	6 (2.2)	0.64
Hormone replacement therapy*	3 (1.8)	0 (0)	3 (1.8)	NA	1 (3)	0 (0)	2 (2.4)	0.50
Less well-documented risk factors, n (%)								
Hematological disease	3 (0.7)	2 (0.7)	1 (0.6)	0.99	1 (1.5)	0 (0)	2 (0.7)	0.50
Antiphospholipid syndrome	6 (1.4)	0 (0)	6 (3.7)		1 (1.5)	1 (1)	4 (1.5)	1
Recent infection	34 (7.8)	17 (6.2)	17 (10.4)	0.12	9 (13.8)	9 (8.7)	16 (5.9)	0.09
Patent foramen ovale	52 (11.9)	26 (9.5)	26 (15.9)	0.05	17 (26.2)	14 (13.6)	21 (7.8)	<0.001
Migraine	28 (6.4)	7 (2.6)	21 (12.8)	<0.001	7 (10.8)	13 (12.6)	8 (3)	0.001
Obstructive sleep apnea	12 (12.7)	7 (2.6)	5 (3)	0.77	0 (0)	3 (2.9)	9 (3.3)	0.44
HIV	5 (1.1)	4 (1.5)	1 (0.6)	0.66	0 (0)	5 (4.9)	0 (0)	0.001
Oral contraceptives *	17 (10.3)	NA	17 (10.3)	NA	8 (24.2)	7 (4.8)	2 (2.4)	<0.001
Heavy drinking	68 (15.6)	61 (22.3)	7 (4.3)	<0.001	5 (7.7)	8 (7.8)	55 (20.4)	0.002
Recent heavy drinking	47 (10.8)	39 (14.3)	8 (4.9)	0.002	7 (10.8)	9 (8.7)	31 (11.5)	0.74

Low birth weight	11 (2.5)	4 (1.5)	7 (4.3)	0.11	2 (3.1)	2 (1.9)	7 (2.6)	0.83
Drug abuse	2 (0.5)	2 (0.7)	0 (0)	0.53	2 (3.1)	0 (0)	0 (0)	0.02
Stroke family history	135 (30.8)	88 (32.2)	47 (28.6)	0.24	12 (18.4)	31 (30)	92 (34.2)	0.02
Stroke etiology								
TOAST, n (%)								
Large-artery atherosclerosis	81 (18.5)	58 (21.2)	23 (14)	0.19	3 (4.6)	13 (12.6)	65 (24.2)	0.008
Cardioembolism	82 (18.8)	53 (19.4)	29 (17.7)		17 (26.2)	19 (18.4)	46 (17.1)	
Small-vessel disease	49 (11.2)	29 (10.6)	20 (12.2)		7 (10.8)	11 (10.7)	31 (11.5)	
Other determined etiology	40 (9.2)	20 (7.3)	20 (12.2)		10 (15.4)	10(9.7)	20 (7.4)	
Undetermined etiology	185 (42.3)	113 (41.4)	72 (43.9)		28 (43.1)	50 (48.5)	107 (39.8)	
A. Incomplete evaluation	24 (5.5)	16 (5.9)	8 (4.9)		2 (3.1)	9 (8.7)	13 (4.8)	
B. Extensive evaluation	150 (34.3)	88 (32.2)	62 (37.8)		24 (36.9)	41 (39.8)	85 (31.6)	
C. Multiple etiologies	11 (2.5)	9 (3.3)	2 (1.2)		2 (3.1)	0 (0)	9 (3.3)	
CCS, n (%)								
Large-artery atherosclerosis	83 (19)	62 (22.7)	21 (12.8)	0.02	3 (4.6)	10 (9.7)	70 (26)	<0.001
Cardioembolism	97 (22.2)	66 (24.2)	31 (18.9)		17 (26.2)	21 (20.4)	59 (21.9)	

Small-vessel disease	51 (11.7)	30 (11)	21 (12.8)	9 (13.8)	11 (10.7)	31 (11.5)
Other determined etiology	36 (8.2)	17 (6.2)	19 (11.6)	9 (13.8)	10 (9.7)	17 (6.3)
Undetermined etiology	170 (38.9)	98 (35.9)	72 (43.9)	27 (41.5)	51 (49.5)	92 (34.2)
A. Incomplete evaluation	18 (4.1)	12 (4.4)	6 (3.7)	2 (3.1)	7 (6.8)	9 (3.3)
B. Cryptogenic embolism	23 (5.3)	15 (5.5)	8 (4.9)	3 (4.6)	10 (9.7)	10 (3.7)
C. Other cryptogenic	125 (28.6)	68 (24.9)	57 (34.8)	22 (33.8)	33 (32)	70 (26)
D. Unclassified	4 (0.9)	3 (1.1)	1 (0.6)	0 (0)	1 (1)	3 (1.1)

*Percentage calculated for women;

NIHSS 0 (National Institutes of Health Stroke Scale on admission); NIHSS 7 (National Institutes of Health Stroke Scale at 7 days); IQR (interquartile rate); NA (not applicable); HIV (human immunodeficiency virus); TOAST (Trial of Org 10172 in Acute Stroke Treatment); CCS (The Causative Classification of ischemic stroke)

Supplementary Table 2. Sources of cardioembolism and cases with ohter determined etiology (OE) according to TOAST classification.

Cardioembolism, n (%)	82 (18.8)
High risk	
Atrial fibrillation	30 (36.6)
Cardiomyopathy	6 (7.3)
Myocardial infarction	3 (3.7)
Akinetic left ventricular segment	3 (3.7)

Atrial myxoma	2 (2.4)
Other	5 (6.1)
Medium risk	
Patent foramen ovale	33 (40.3)
Other determined etiology, n (%)	40 (9.2)
Dissection	18 (45.0)
Active cancer	6 (15.0)
Antiphospholipid syndrome	5 (12.5)
Moyamoya disease	3 (7.5)
Other	8 (20.0)

Supplemental material

Supplemental methods.

The diagnostic workup consisted of brain imaging by computed tomography and/or magnetic resonance imaging; vascular imaging of extra- and intracranial arteries (by ultrasonography, computed tomography, magnetic resonance-angiography or catheter angiography) and cardiac evaluation by electrocardiography, transthoracic echocardiography, transesophageal echocardiography and 24-hour cardiac rhythm Holter monitoring. In addition, profound set of blood tests were analysed (including markers for systemic autoimmune diseases, coagulation factors, metabolic and cardiac markers etc). All patients underwent imaging with computed tomography and/or magnetic resonance imaging, intra- and extracranial arteries were studied in 98% of patients. Cardiac investigations were done as follows: transthoracic echocardiography in 409 (94%), cardiac rhythm Holter monitoring in 380 (87%) and transesophageal echocardiography in 276 (63%) patients.

A total of 350 patients (80%) were screened for Fabry disease (all negative).

Classification of stroke etiology.

The TOAST classification consists of 5 different categories: large-artery atherosclerosis (LAA), small-artery occlusion (SAO), cardioembolism (CE), stroke of other determined etiology (OE) and stroke of undetermined etiology (UND). The CE subtype is further divided into high and low risk causes and the UND subtype comprises cases with several causes, with unknown cause despite thorough investigations and unknown cause due to incomplete investigations. The investigations were defined as incomplete if any of the abovementioned diagnostic procedures were not done, only transesophageal echocardiography was not considered obligatory. In addition, if the etiology was obvious with the investigations done, but some of the abovementioned

investigations were not done, the stroke was classified accordingly (rather than UND) in the TOAST system. The CCS classification has also 5 different categories (LAA, SAO, CE, OE and UND). The UND subtype is divided into cryptogenic embolism, other cryptogenic, incomplete investigation and unclassified. The web-based algorithm was used to classify patients according to CCS system (http://ccs.mgh.harvard.edu). In difficult cases, the etiology was decided by at least two study members via discussions.

Definitions of stroke risk factors.

The definition for hypertension was as follows: >140 mmHg systolic blood pressure and/or >90 mmHg diastolic blood pressure before stroke or 7 days after stroke or if on antihypertensive treatment. The criteria for dyslipidemia were serum total cholesterol ≥5.0 mmol/L, low-density lipoprotein cholesterol ≥3.0 mmol/L, high-density lipoprotein cholesterol <1.0 mmol/L, or previous cholesterol-lowering therapy. Diabetes mellitus was diagnosed on the basis of one of the following criteria: fasting plasma glucose ≥7.0 mmol/L, two-hour post-glucose challenge value ≥11.1 mmol/L, glycated hemoglobin ≥ 6.5%, or the patient was taking anti-diabetic medication. Obesity was defined as the body mass index ≥30kg/m_2. Cigarette smoking was considered as risk factor if the patient was an active smoker Ischemic heart disease, cardiac insufficiency and periferal artery disease were risk factors if diagnosed before or during the acute stroke. Atrial fibrillation was a risk factor if diagnosed before or during the acute stroke. Hormone replacement therapy was a risk factor if the therapy was used at stroke onset. Oral contraceptives were considered as a risk factor if were used at stroke onset. Hematological diseases and coagulation disorders were active leuckemia, aneemia, polycytemia or antiphospholipid syndrome. Factor V Leiden and prothrombin gene mutations were not included as stroke risk factors and were classified as UND etiology. Recent infection was considered as risk factor if any infection was diagnosed in two weeks before stroke or at the time of admission. Heavy drinking was defined as constant use of alcohol reported by the patient or a relative. Recent heavy drinking was defined as intake of >40g alcohol within the 24h prior to stroke. Low birth weight was patient-reported and defined as birth-weight <3000g. Patent foramen ovale (PFO) was diagnosed by transesophageal echocardiography (TEE). We considered PFO as the etiology of stroke if there were no other causes evident. Sleep apnea and migraine were considered as stroke risk factors when diagnosed prior to stroke. Family history of stroke was positive if ischemic or hemorrhagic stroke had been diagnosed in patients' 1st degree relative.