Online Data Supplement

Genetic overlap between diagnostic subtypes of ischemic stroke

Supplementary methods and results

Quantitative bias analysis – subtype misclassification

For given misclassification rates, the greatest bias will arise under two extreme scenarios. First, if all misclassified LAA cases were truly SVD and vice versa (S1), a proportion of r_g will reflect *within*-subtype correlation, biasing r_g upwards. Alternatively, if all misclassified cases were neither LAA nor SVD (S2), the observed r_g will be typically, but not always, biased downwards [1]. We re-estimated r_g under S1 and S2 allowing for misclassification rates consistent with reported inter-rater reliability (Cohen's kappa) values for TOAST subtypes. Averaged across subtypes, kappa statistics generally range from 0.5 to 0.8 [2, 3]. One study reported subtype-specific values of 0.8 for LAA and 0.53 for SVD [3]. Since kappa is a correlation measure, these values were specified as assumed correlations of marker effects between discovery and target samples in the *within*-trait analyses of LAA and SVD during estimation of the within-trait variance ($vg1/h^2$) explained by profile scores. This bias-adjusted estimate of vg1 was then used to estimate the genetic correlation (r_g) between LAA and SVD using profile scores from the cross-trait analyses (Table XIV). With kappa=0.8 for LAA and 0.5 for SVD, under S1 r_g was reduced, but still significantly different from zero (r_g=0.63, 95% CI: [0.34, 0.74]). Under S2 the estimate was slightly higher (r_g=0.75, 95% CI: [0.43, 0.98]). The true misclassification scenario will likely lie between the extremes described by S1 and S2, suggesting robustness of the observed genetic correlation to likely levels of subtype misclassification.

Table I. Metastroke study details

Cohort Name	IS – broadly defined	LAA	CE	SVD	Controls
ARIC	385	31	93	63	8,803
ASGC	1,162	421	240	310	1,244
BRAINS	361	120	29	97	444
CHS	454	–	147	73	2,817
deCODE	2,391	255	399	240	26,970
FHS	171	–	48	–	4,164
GEOS	448	37	90	54	498
HPS	578	–	–	–	468
HVH	566	61	88	173	1,290
ISGS/SWISS	1,070	229	247	201	2,329
MGH/GASROS	516	95	169	38	1,202
Milano	372	74	65	25	407
Rotterdam	367	–	–	–	5,396
WTCCC2-Munich	1,174	346	330	106	797
WTCCC2-UK	2,374	498	460	474	5,175
Total	**12,389**	**2,167**	**2,405**	**1,854**	**62,004**

Notes: IS – ischaemic stroke; LAA – large artery atherosclerosis; CE – cardioembolic stroke; SVD – small vessel disease.

Table II. Linear mixed modelling: univariate analyses within subtypes

Trait	Cases	Controls	K[a]	h^2 (SE) liability scale [b]	P [c]
Broad IS	4,561	7,094	0.02	0.18 (0.025)	$1x10^{-14}$
LAA	1,207	7,094	0.004	0.19 (0.046)	$2x10^{-5}$
CE	998	7,094	0.004	0.24 (0.053)	$2x10^{-6}$
SVD	862	7,094	0.004	0.10 (0.060)	0.04

[a] Lifetime population prevalence. [b] Estimates based on 345,336 genome-wide SNPs.
[c] from a likelihood ratio test of the null hypothesis H_0: $h^2=0$.

Table III. Linear mixed modelling: bivariate analyses (between subtypes)

	Trait 1/Trait 2			
	LAA/CE	LAA/SVD	CE/SVD	Control/Control[d]
Cases	1,207/998	1,207/862	998/862	
Controls	3,547/3,547	3,547/3,547	3,547/3,547	
h^2 trait 1 (SE)[a]	0.18 (0.052)	0.17 (0.052)	0.23 (0.059)	0.00 (0.014)
h^2 trait 2 (SE)[a]	0.23 (0.059)	0.10 (0.066)	0.10 (0.066)	–
r_g (SE)[a]	0.39 (0.21)	0.96 (0.47)	0.64 (0.40)	–
SD (r_g)	0.066	0.059	0.18	–
P^b	0.024	9×10^{-4}	0.017	0.43
P^c	0.012	0.44	0.11	

[a] Mean estimate from ten permuted assignments of controls to one of the two case traits. Each of the permuted control assignments was used to estimate of h^2 (trait 1), h^2 (trait 2) and r_g for all three trait combinations. [b] P-value for H_0: $r_g=0$, H_1: $r_g>0$. [c] P-value for H_0: $r_g=1$, H_1: $r_g<1$. [d] The permuted control subsets also allowed us to benchmark the expected h^2_{SNP} by assessing genetic sharing within controls. Across ten control subsets the mean was zero, suggesting lack of inflation by factors unrelated to case-control status.

Table IV. Polygenic scoring results for broadly defined ischaemic stroke (IS)

P_T	N_{SNPs} ASGC	N_{SNPs} WTCCC2-Mun	N_{SNPs} WTCCC2-UK	*P* (score)[a]	Het I^2 [b]	*P* (het) [c]	R^2 (%) [d]	h^2 [e]
0.0001	73	93	85	**0.003**	0%	0.53	0.06	0.004
0.001	602	619	627	**3.2×10^{-6}**	0%	0.44	0.16	0.013
0.01	4,710	4,633	4,726	**7.2×10^{-5}**	42%	0.18	0.22	0.025
0.05	18,432	18,557	18,933	**1.8×10^{-7}**	0%	0.88	0.19	0.043
0.1	32,876	32,843	33,548	**7.1×10^{-7}**	0%	0.47	0.17	0.044
0.2	57,318	57,288	57,969	**2.2×10^{-7}**	0%	0.60	0.19	0.051
0.3	78,183	77,965	78,660	**8.8×10^{-8}**	0%	0.88	0.20	0.056
0.4	95,954	95,459	96,220	**1.6×10^{-8}**	0%	0.88	0.22	0.062
0.5	111,005	110,406	110,944	**1.7×10^{-8}**	0%	0.83	0.22	0.064
1	154,056	152,907	152,394	**1.1×10^{-8}**	0%	0.79	0.22	0.074

[a] From random effects meta-analysis across the three target cohorts (see Table 1, main paper) using all samples listed in Table I above (except the target cohort) in the discovery meta-analysis. [b] Percent of variation of score effects across studies due to heterogeneity, based on Cochran's Q statistic. [c] Significance of Q statistic. [d] Nagelkerke's R^2: Sample-size weighted mean of estimates from the three individual studies. [e] IS target sample variance explained by the polygenic score. h^2 was calculated using quantitative genetics theory [4] with the observed sample sizes, P_T, N_{SNPs}, *P*(score), assuming genetic correlation =1 between IS in discovery and target samples, and IS lifetime population prevalence = 0.02 [5, 6].

Table V. Polygenic scoring results for large artery atherosclerosis (LAA)

P_T	N_{SNPs} ASGC	N_{SNPs} WTCCC2-Mun	N_{SNPs} WTCCC2-UK	*P* (score) [a]	Het I^2 [b]	P (het) [c]	R^2 (%) [d]	h^2 [e]
0.0001	109	90	90	0.36	0%	0.43	0.024	0.005
0.001	689	650	672	0.24	0%	0.60	0.047	0.015
0.01	5,216	5,099	5,336	**9.2×10^{-5}**	0%	0.52	0.28	0.072
0.05*	20,158	20,258	20,794	**1.7×10^{-8}**	0%	0.64	0.48	0.128
0.1	35,765	35,815	36,481	**1.9×10^{-7}**	0%	0.94	0.40	0.128
0.2	62,393	62,099	63,150	**2.5×10^{-7}**	0%	0.99	0.41	0.138
0.3	84,922	84,770	85,603	**1.5×10^{-7}**	0%	0.97	0.41	0.150
0.4	104,363	104,249	104,803	**5.0×10^{-8}**	0%	0.85	0.43	0.163
0.5	121,026	120,984	121,283	**3.8×10^{-8}**	0%	0.90	0.45	0.172
1	170,896	170,552	169,957	**3.4×10^{-8}**	0%	0.85	0.45	0.197

[a] From random effects meta-analysis across the three target cohorts (see Table 1, main paper) using all samples listed in Table I above (except the target cohort) in the discovery meta-analysis. [b] Percent of variation of score effects across studies due to heterogeneity, based on Cochran's Q statistic. [c] Significance of Q statistic. [d] Nagelkerke's R^2: Sample-size weighted mean of estimates from the three individual studies. [e] LAA target sample variance explained by the polygenic score. h^2 was calculated using quantitative genetics theory [4] with the observed sample sizes, P_T, N_{SNPs}, *P*(score), assuming genetic correlation =1 between LAA in discovery and target samples, and LAA lifetime population prevalence = 0.004 [5, 6].

* Score showing maximum association with both LAA (above) and SVD (Table VII).

Table VI. Polygenic scoring results for cardioembolic stroke (CE)

P_T	N_{SNPs} ASGC	N_{SNPs} WTCCC2-Mun	N_{SNPs} WTCCC2-UK	*P* (score) [a]	Het I^2 [b]	*P* (het) [c]	R^2 (%) [d]	h^2 [e]
0.0001	107	101	96	0.23	60%	0.08	0.19	0.006
0.001	664	695	727	**1.7×10^{-4}**	0%	0.91	0.23	0.034
0.01	4,928	4,961	5,031	**2.1×10^{-4}**	0%	0.89	0.26	0.066
0.05	19,454	19,437	20,000	**0.009**	28%	0.25	0.30	0.059
0.1	34,326	34,543	35,109	**0.017**	29%	0.24	0.27	0.058
0.2	59,723	59,554	60,215	**0.049**	44%	0.17	0.28	0.052
0.3	80,684	80,550	81,223	0.064	43%	0.18	0.26	0.052
0.4	98,386	98,378	98,563	0.16	57%	0.10	0.27	0.041
0.5	113,435	113,474	113,287	0.13	54%	0.11	0.26	0.046
1	156,404	155,514	155,123	0.18	55%	0.11	0.24	0.046

[a] From random effects meta-analysis across the three target cohorts (see Table 1, main paper) using all samples listed in Table I above (except the target cohort) in the discovery meta-analysis. [b] Percent of variation of score effects across studies due to heterogeneity, based on Cochran's Q statistic. [c] Significance of Q statistic. [d] Nagelkerke's R^2: Sample-size weighted mean of estimates from the three individual studies. [e] CE target sample variance explained by the polygenic score. h^2 was calculated using quantitative genetics theory [4] with the observed sample sizes, P_T, N_{SNPs}, *P*(score), assuming genetic correlation =1 between CE in discovery and target samples, and CE lifetime population prevalence = 0.004 [5, 6].

Table VII. Polygenic scoring results for small vessel disease (SVD)

P_T	N_{SNPs} ASGC	N_{SNPs} WTCCC2-Mun	N_{SNPs} WTCCC2-UK	*P* (score) [a]	Het I^2 [b]	*P* (het) [c]	R^2 (%) [d]	h^2 [e]
0.0001	91	86	114	0.51	0%	0.61	0.018	0.005
0.001	677	620	728	0.95	0%	0.85	0.0019	0.001
0.01	5,114	4,861	5,449	0.92	0%	0.47	0.022	0.003
0.05	20,332	19,636	20,974	0.91	0%	0.67	0.010	0.003
0.1*	35,767	35,057	36,383	0.78	20%	0.29	0.048	0.008
0.2	61,832	60,873	62,892	0.97	38%	0.20	0.058	0.001
0.3	83,691	82,827	84,725	0.76	52%	0.13	0.069	0.010
0.4	102,456	101,697	103,063	0.90	36%	0.21	0.052	0.004
0.5	118,290	117,688	118,447	0.94	34%	0.22	0.046	0.003
1	163,946	163,305	162,329	0.87	42%	0.18	0.053	0.007

[a] From random effects meta-analysis across the three target cohorts (see Table 1, main paper) using all samples listed in Table I above (except the target cohort) in the discovery meta-analysis. [b] Percent of variation of score effects across studies due to heterogeneity, based on Cochran's Q statistic. [c] Significance of Q statistic. [d] Nagelkerke's R^2: Sample-size weighted mean of estimates from the three individual studies. [e] SVD target sample variance explained by the polygenic score. h^2 was calculated using quantitative genetics theory [4] with the observed sample sizes, P_T, N_{SNPs}, *P*(score), assuming genetic correlation =1 between SVD in discovery and target samples, and SVD lifetime population prevalence = 0.004 [5, 6].

* Score showing maximum association with LAA (Table IX).

Table VIII. Pleiotropy analyses: association of LAA-based polygenic scores with SVD

P_T	N_{SNPs} ASGC	N_{SNPs} WTCCC2-Mun	N_{SNPs} WTCCC2-UK	*P* (score)	Het I^2	*P* (het)	Nagelkerke's R^2 (%)
0.0001	109	90	90	0.35	0%	0.54	0.031
0.001	689	650	672	0.28	0%	0.50	0.029
0.01	5,216	5,099	5,336	0.15	34%	0.22	0.090
0.05	20,158	20,258	20,794	**0.032**	0%	0.59	0.082
0.1	35,765	35,815	36,481	0.060	0%	0.39	0.075
0.2	62,393	62,099	63,150	0.069	36%	0.21	0.104
0.3	84,922	84,770	85,603	0.064	13%	0.32	0.080
0.4	104,363	104,249	104,803	**0.039**	0%	0.37	0.080
0.5	121,026	120,984	121,283	**0.038**	0%	0.42	0.085
1	170,896	170,552	169,957	0.054	0%	0.46	0.069

Table IX. Pleiotropy analyses: association of SVD-based polygenic scores with LAA

P_T	N_{SNPs} ASGC	N_{SNPs} WTCCC2-Mun	N_{SNPs} WTCCC2-UK	*P* (score)	Het I^2	*P* (het)	Nagelkerke's R^2 (%)
0.0001	91	86	114	0.75	41%	0.18	0.038
0.001	677	620	728	0.51	0%	0.67	0.010
0.01	5,114	4,861	5,449	0.23	0%	0.94	0.022
0.05	20,332	19,636	20,974	**2.4×10^{-4}**	0%	0.93	0.197
0.1	35,767	35,057	36,383	**2.1×10^{-4}**	0%	0.81	0.192
0.2	61,832	60,873	62,892	**0.001**	0%	0.96	0.161
0.3	83,691	82,827	84,725	**0.001**	0%	0.86	0.150
0.4	102,456	101,697	103,063	**0.001**	0%	0.88	0.154
0.5	118,290	117,688	118,447	**0.001**	0%	0.88	0.153
1	163,946	163,305	162,329	**0.001**	0%	0.80	0.141

Table X. Pleiotropy analyses: association of LAA-based polygenic scores with CE

P_T	N_{SNPs} ASGC	N_{SNPs} WTCCC2-Mun	N_{SNPs} WTCCC2-UK	*P* (score)	Het I^2	*P* (het)	Nagelkerke's R^2 (%)
0.0001	109	90	90	0.58	73%	0.03	0.083
0.001	689	650	672	0.32	0%	0.64	0.018
0.01	5,216	5,099	5,336	0.48	0%	0.55	0.023
0.05	20,158	20,258	20,794	0.18	0%	0.95	0.029
0.1	35,765	35,815	36,481	0.27	13%	0.32	0.048
0.2	62,393	62,099	63,150	0.21	0%	0.53	0.037
0.3	84,922	84,770	85,603	0.19	0%	0.71	0.035
0.4	104,363	104,249	104,803	0.15	0%	0.80	0.032
0.5	121,026	120,984	121,283	0.13	0%	0.88	0.036
1	170,896	170,552	169,957	0.15	0%	0.97	0.038

Table XI. Pleiotropy analyses: association of CE-based polygenic scores with LAA

P_T	N_{SNPs} ASGC	N_{SNPs} WTCCC2-Mun	N_{SNPs} WTCCC2-UK	*P* (score)	Het I^2	*P* (het)	Nagelkerke's R^2 (%)
0.0001	107	101	96	0.63	0%	0.48	0.015
0.001	664	695	727	0.99	0%	0.81	0.003
0.01	4,928	4,961	5,031	0.88	0%	0.79	0.001
0.05	19,454	19,437	20,000	0.73	42%	0.18	0.040
0.1	34,326	34,543	35,109	0.83	68%	0.04	0.065
0.2	59,723	59,554	60,215	0.85	41%	0.18	0.031
0.3	80,684	80,550	81,223	0.99	26%	0.26	0.023
0.4	98,386	98,378	98,563	0.71	44%	0.17	0.038
0.5	113,435	113,474	113,287	0.57	42%	0.18	0.034
1	156,404	155,514	155,123	0.70	51%	0.13	0.042

Table XII. Pleiotropy analyses: association of CE-based polygenic scores with SVD

P_T	N_{SNPs} ASGC	N_{SNPs} WTCCC2-Mun	N_{SNPs} WTCCC2-UK	*P* (score)	Het I^2	*P* (het)	Nagelkerke's R^2 (%)
0.0001	107	101	96	0.068	0%	0.43	0.105
0.001	664	695	727	0.31	8%	0.34	0.060
0.01	4,928	4,961	5,031	0.32	0%	0.85	0.018
0.05	19,454	19,437	20,000	0.85	0%	0.48	0.021
0.1	34,326	34,543	35,109	0.62	50%	0.14	0.082
0.2	59,723	59,554	60,215	0.63	53%	0.12	0.105
0.3	80,684	80,550	81,223	0.33	40%	0.19	0.124
0.4	98,386	98,378	98,563	0.50	58%	0.09	0.170
0.5	113,435	113,474	113,287	0.45	57%	0.10	0.170
1	156,404	155,514	155,123	0.36	54%	0.11	0.176

Table XIII. Pleiotropy analyses: association of SVD-based polygenic scores with CE

P_T	N_{SNPs} ASGC	N_{SNPs} WTCCC2-Mun	N_{SNPs} WTCCC2-UK	*P* (score)	Het I^2	*P* (het)	Nagelkerke's R^2 (%)
0.0001	91	86	114	0.15	0%	0.60	0.034
0.001	677	620	728	0.74	52%	0.13	0.057
0.01	5,114	4,861	5,449	0.89	22%	0.28	0.033
0.05	20,332	19,636	20,974	0.42	0%	0.46	0.026
0.1	35,767	35,057	36,383	0.53	5%	0.35	0.031
0.2	61,832	60,873	62,892	0.57	19%	0.29	0.029
0.3	83,691	82,827	84,725	0.56	0%	0.41	0.021
0.4	102,456	101,697	103,063	0.56	0%	0.38	0.023
0.5	118,290	117,688	118,447	0.58	0%	0.43	0.020
1	163,946	163,305	162,329	0.60	0%	0.50	0.015

Table XIV. Bias analysis: genetic correlations between LAA and SVD allowing for varying levels of subtype misclassification

		Scenario 1 [a]	Scenario 2 [b]
Kappa LAA	Kappa SVD	r_g (95% CI)	r_g (95% CI)
1	1	0.72 (0.51,0.92) [c]	0.72 (0.51,0.92)
0.9	0.9	0.68 (0.51,0.86)	0.71 (0.51,0.90)
0.9	0.8	0.67 (0.51,0.84)	0.72 (0.51,0.92)
0.9	0.7	0.66 (0.50,0.81)	0.73 (0.52,0.95)
0.9	0.6	0.65 (0.43,0.79)	0.75 (0.48,0.99)
0.9	0.5	0.64 (0.36,0.77)	0.78 (0.41,0.99)
0.8	0.8	0.65 (0.51,0.80)	0.69 (0.51,0.87)
0.8	0.7	0.64 (0.47,0.78)	0.71 (0.51,0.90)
0.8	0.6	0.63 (0.40,0.76)	0.72 (0.51,0.93)
0.8	0.5	0.62 (0.33,0.74)	0.75 (0.43,0.98)
0.7	0.7	0.62 (0.44,0.74)	0.68 (0.51,0.85)
0.7	0.6	0.62 (0.37,0.73)	0.70 (0.51,0.88)
0.7	0.5	0.60 (0.31,0.71)	0.71 (0.45,0.92)
0.6	0.6	0.60 (0.35,0.69)	0.67 (0.51,0.82)
0.6	0.5	0.59 (0.29,0.68)	0.68 (0.49,0.86)
0.5	0.5	0.57 (0.26,0.65)	0.65 (0.51,0.80)

Notes: Each r_g is the average of estimates from two directional polygenic profiling analyses: i) LAA used as derivation trait and profile scores assessed for association with SVD, and; ii) SVD used as derivation trait and profile scores assessed for association with LAA.

[a] Under Scenario 1, all misclassified LAA cases were truly SVD, and vice versa. For the derivation trait, the relevant kappa was specified as the within-trait correlation of genetic effect sizes. This resulted in proportional inflation of the total fraction of trait1 variance explained by genetic effects (vg1) in the profile score. That is, when misclassification exists, the true proportion of trait1 variance explained by the score would be higher. The genetic correlation between genetic effects across traits (r_g) was estimated using the adjusted vg1 and an adjusted estimate of trait2 variance explained by the score (vg2) which was similarly inflated upwards by that trait‘s kappa value. That is, under this scenario of misclassification, the estimate of trait 2 variance explained by the score would be higher since a component of vg2 now reflects spurious correlation of genetic effects within trait 1. This proportional inflation of both vg1 and vg2 reduces the estimated genetic correlation consistent with the observed association statistics.

[b] Under Scenario 2, all misclassified cases were neither LAA nor SVD. For the derivation trait, vg1 was proportionally inflated as before. For the validation trait, vg2 was reduced to the proportion represented by the relevant kappa value. That is, under this misclassification scenario, the estimate of trait 2 variance explained by the score is lower, since the misclassified proportion of cases are of a different subtype to the derivation trait.

[c] Reported value, assuming no subtype misclassification

Table XV. Association results for SNPs reaching $P<1x10^{-5}$ in joint meta-analysis of LAA and SVD

SNP	CHR	BP	A1	Freq	OR	95% CI	*P*	Het I^2 (%)	Het *P*	Nearest Gene
rs11155944	6	154,245,875	t	0.24	0.84	(0.78, 0.89)	1.6E-07	21	0.25	*OPRM1*
rs932671	6	154,258,051	a	0.25	0.84	(0.79, 0.89)	2.3E-07	21	0.25	*OPRM1*
rs9371764	6	154,259,797	t	0.25	0.86	(0.81, 0.92)	5.1E-06	57	0.01	*OPRM1*
rs17084671	6	154,261,451	a	0.24	0.84	(0.78, 0.89)	1.3E-07	15	0.31	*OPRM1*
rs6938958	6	154,262,120	a	0.76	1.19	(1.11, 1.26)	1.8E-07	20	0.26	*OPRM1*
rs7763080	6	154,263,321	a	0.25	0.84	(0.78, 0.89)	1.7E-07	17	0.28	*OPRM1*
rs12429886	13	46,092,686	t	0.09	1.27	(1.15, 1.39)	1.4E-06	48	0.04	*LRCH1*
rs7983635	13	46,093,268	a	0.92	0.79	(0.71, 0.87)	7.2E-06	52	0.03	*LRCH1*
rs12427953	13	46,101,896	t	0.93	0.79	(0.71, 0.87)	6.0E-06	50	0.03	*LRCH1*
rs12429970	13	46,138,327	c	0.93	0.78	(0.69, 0.86)	2.3E-06	52	0.03	*LRCH1*
rs1483968	14	22,021,733	t	0.075	1.35	(1.18, 1.53)	9.4E-06	0	0.44	*TCRA*
rs8072419	17	46,982,134	a	0.20	1.16	(1.08, 1.23)	6.5E-06	0	0.80	*CA10*
rs134197	22	26,993,507	a	0.12	1.22	(1.12, 1.33)	4.8E-06	33	0.14	*TTC28*
rs6519761	22	27,431,600	a	0.86	1.20	(1.10, 1.29)	9.4E-06	0	0.92	*CHEK2*
rs1884816	22	27,436,733	t	0.84	1.20	(1.11, 1.29)	2.7E-06	0	0.95	*CHEK2*

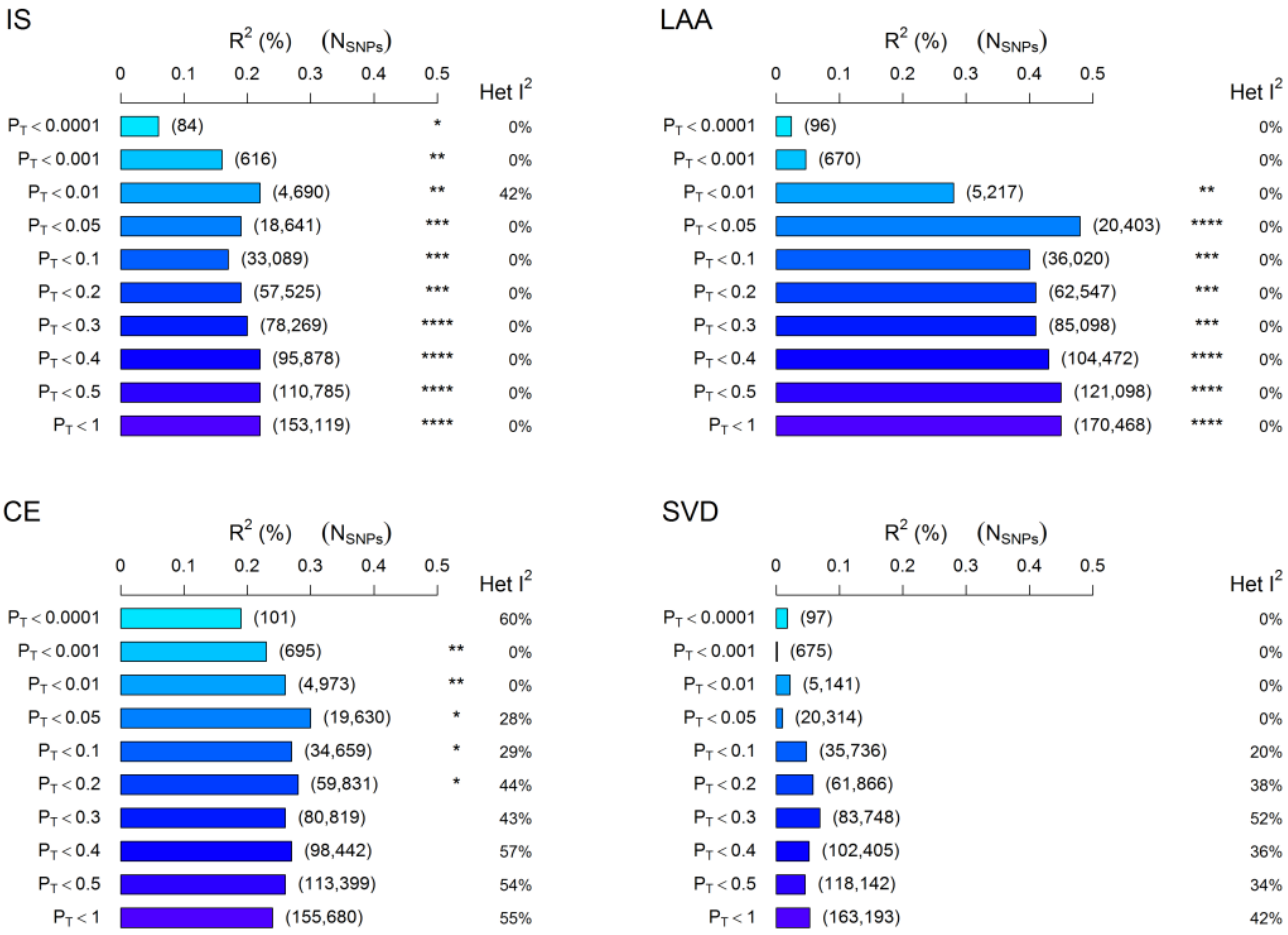

Figure I. Results of polygenic profile scoring within ischemic stroke (IS) and its three major subtypes: large artery atherosclerosis (LAA), cardioembolism (CE) and small vessel disease (SVD). P_T: discovery threshold for including SNPs in the score. R^2 (%): percentage of observed case-control variance explained by the score in target samples. N_{SNPs}: number of approximately independent SNPs included in the score. I^2: heterogeneity of score effects between target cohorts. * $P \leq 0.05$, ** $P \leq 0.001$, *** $P \leq 1x10^{-5}$, **** $P \leq 1x10^{-7}$.

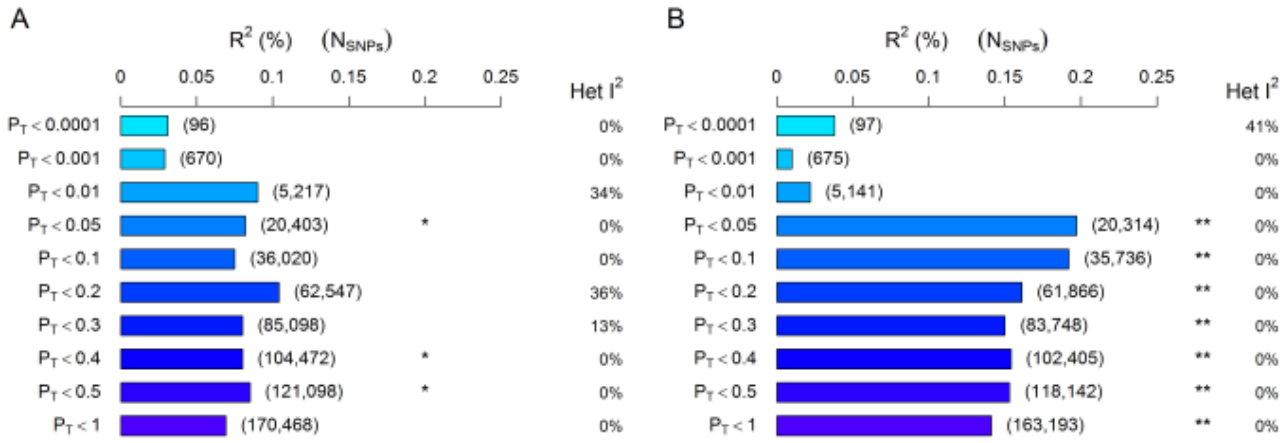

Figure II. Polygenic profile scoring between large artery atherosclerotic stroke (LAA) and small vessel disease (SVD). **A**. Scores derived from meta-analyses of LAA and tested for association with SVD. **B**. Profile scores derived from meta-analyses of SVD and tested for association with LAA. PT: threshold for including SNPs in the score. R^2 (%): estimated percentage of target sample case-control variance explained by the score using Nagelkerke's pseudo-R^2 measure. NSNPs: number of approximately independent SNPs included in the score. I^2: heterogeneity of score effects between target cohorts. * $P \leq 0.05$, ** $P \leq 0.001$

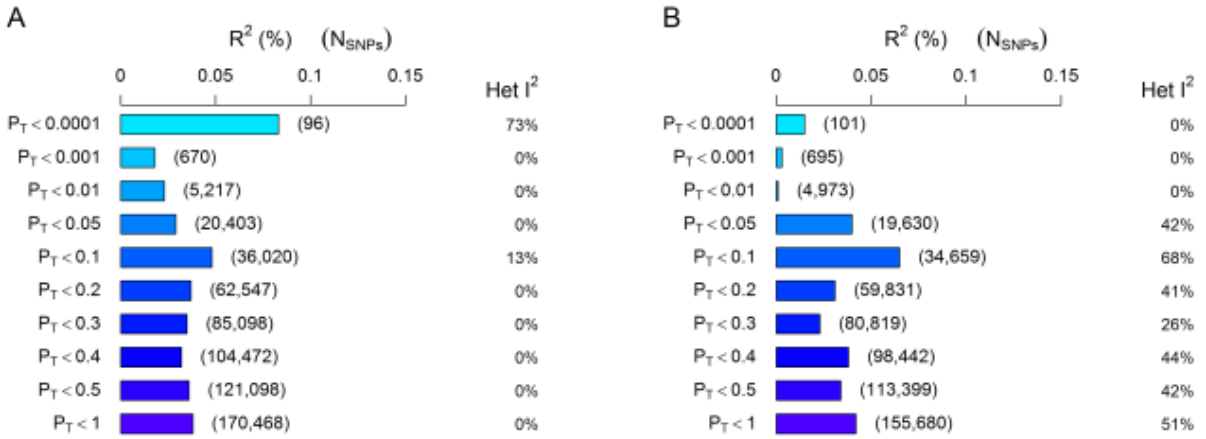

Figure III. Polygenic profile sharing results for LAA and CE. **A:** Scores derived from meta-analyses of LAA and tested for association with CE. **B:** Scores derived from meta-analyses of CE and tested for association with LAA. No score associated with the target trait at $P<0.05$.

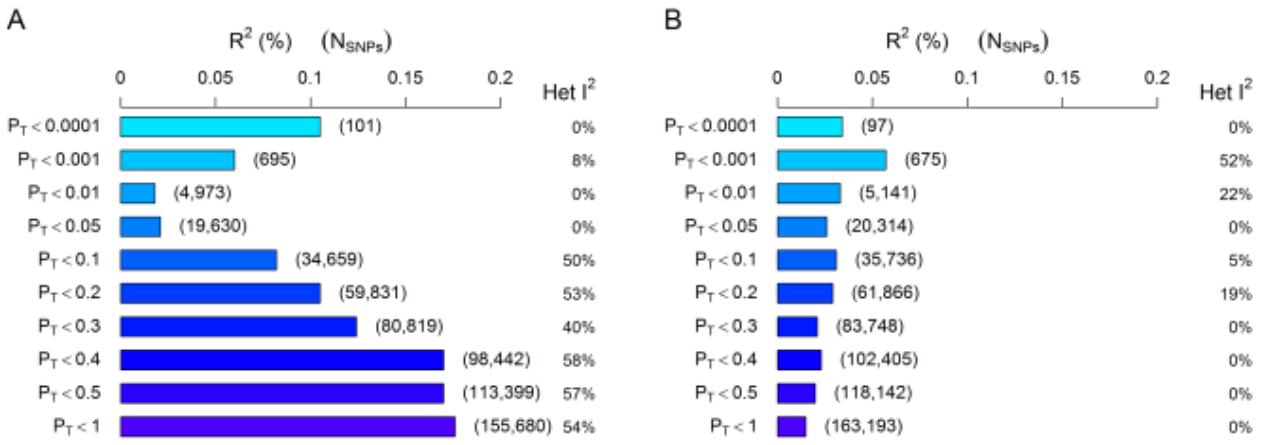

Figure IV. Polygenic profile sharing results for CE and SVD. **A:** meta-analysis performed for CE and profile scores tested for association with SVD. **B:** meta-analysis performed for SVD and derived profile scores tested for association with CE. No score associated with the target trait at $P<0.05$.

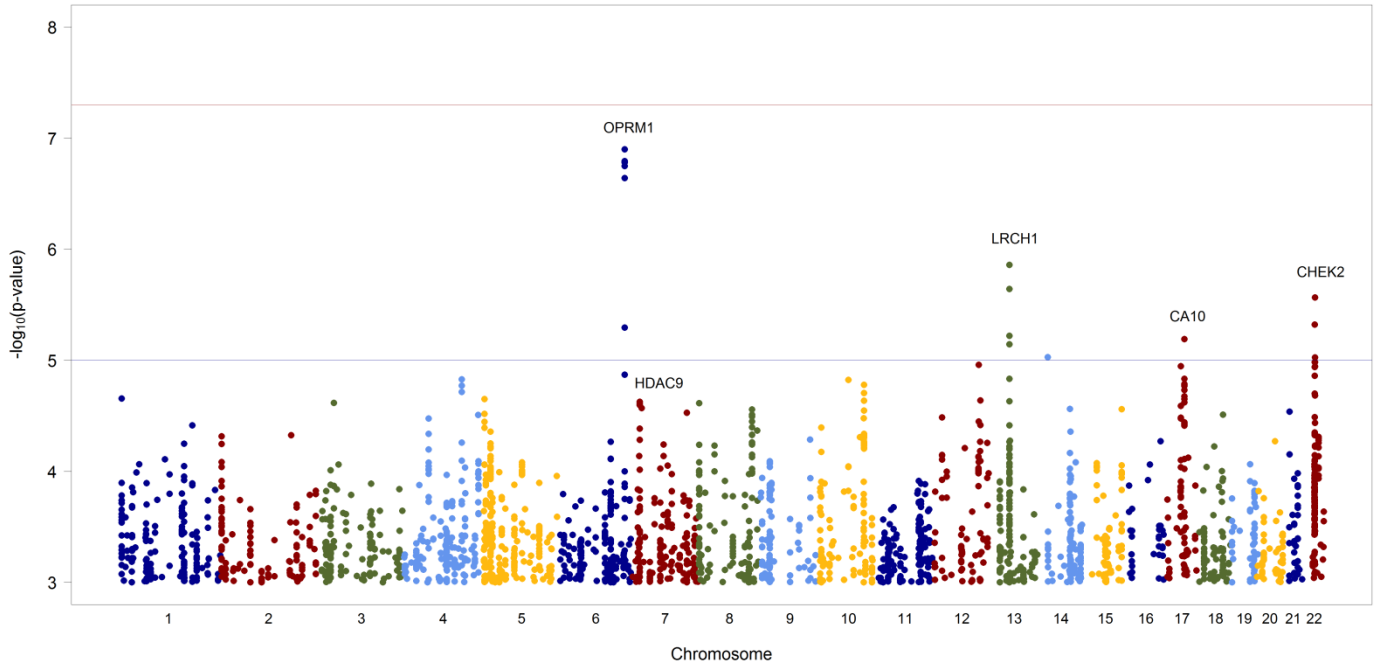

Figure V. Association results for the joint LAA-SVD meta-analysis including 4,021 LAA/SVD cases and 51,976 controls. The plot shows results for SNPs reaching $P \leq 1x10^{-3}$. The upper red line corresponds to $P=5x10^{-8}$ and the lower grey line corresponds to $P=1x10^{-5}$.

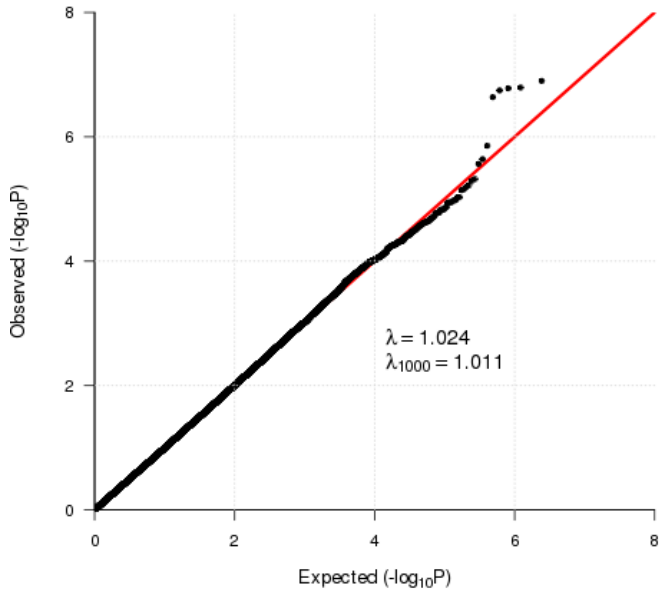

Figure VI. QQ-plot of association results from the joint meta-analysis of LAA and SVD. λ_{1000} is the genomic control factor scaled to 1000 cases and 1000 controls.

References

1. Rothman KJ, Greenland S, Lash TL. In: Rothman KJ, ed. *Modern epidemiology*. Philadelphia, PA USA: Lippincott Williams & Wilkins; 2008:138-139.
2. Goldstein LB, Jones MR, Matchar DB, Edwards LJ, Hoff J, Chilukuri V, et al. Improving the reliability of stroke subgroup classification using the trial of org 10172 in acute stroke treatment (toast) criteria. *Stroke*. 2001;32:1091-1098
3. Meschia JF, Barrett KM, Chukwudelunzu F, Brown WM, Case LD, Kissela BM, et al. Interobserver agreement in the trial of org 10172 in acute stroke treatment classification of stroke based on retrospective medical record review. *J Stroke Cerebrovasc Dis*. 2006;15:266-272
4. Dudbridge F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet*. 2013;9:e1003348
5. Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, et al. Genetic risk factors for ischaemic stroke and its subtypes (the metastroke collaboration): A meta-analysis of genome-wide association studies. *Lancet Neurol*. 2012;11:951-962
6. Zhang Y, Chapman AM, Plested M, Jackson D, Purroy F. The incidence, prevalence, and mortality of stroke in france, germany, italy, spain, the uk, and the us: A literature review. *Stroke Res Treat*. 2012;2012:436125