The genetics of cerebrovascular complications in sickle cell disease

David Rees

King’s College Hospital and King’s College London

Global Sickle Cell Conference, Paris, 2022

Cerebrovascular disease in SCD

- One of the most important complications in SCD
  - 40-50% children with SCA have some evidence of cerebrovascular disease by age of 14 years
  - Increasingly recognized in adults
- Genetic predisposition towards stroke
  - Many publications identifying potential variants changing stroke risk
    - Some genome wide studies
    - Bayesian network modelling (Sebastiani et al, Nat Genetics 2005)
    - Many studies of candidate genes
  - Very few variants validated in duplicate studies
  - None in routine clinical use

Pathophysiology of ischemic cerebrovascular disease in SCD

- Vascular endothelial dysfunction
- Damage to blood vessel structure
  - Vaso-occlusion
  - Bood vessel wall ischemia
    - Stenosis
    - Aneurysm
    - Tortuosity
    - Moya Moya
- Reduced oxygen delivery to brain
  - Anemia – chronic with acute exacerbations
  - Impaired regulation of blood flow
  - Hypoxemia
- Conventional risk factors
  - Hypertension
  - Lipid abnormalities
  - Renal and cardiac disease

Cerebrovascular complications

- Overt Ischaemic Stroke (OIS)
  - Up to 11% of children with SCA without primary stroke prevention
  - 1.02% annual risk between ages of 2 to 5 years.
  - Typically preceded by progressive vasculopathy
- Haemorrhagic Stroke
  - Peak occurrence between 20-29yrs
  - Often result of aneurysm or moya-moya vasculopathy
- Silent Cerebral Infarct (SCI)
  - “Hyperintensity >3mm seen in two planes on T2 weighted MRI with no focal clinical deficit”
  - Up to 35% of children
  - Negative impact on IQ
  - 14 fold increased risk of OIS
- Vasculopathy
  - Smooth muscle hyperplasia with overlying endothelial damage and fibrosis
  - Tortuosity, stenosis, Moya Moya

Heritability of large vessel disease in SCD

- Sibling analysis
  - Ischaemic stroke occurs in families more often than predicted by chance
  - Abnormal TCD 50x more likely if sibling has abnormal TCD
- HLA genotypes associated with increased and decreased risks of OIS
- 4 or more functional α globin genes
  - 2-6 x increased risk of abnormal TCDs
  - 2 x increased risk of ischaemic stroke

- Other potential variants
  - *Associative: G6PD, TNFalpha, ANXA2, TEK, TGFBR3 & PON1*
  - *Protective: GOLGB1, ENPP1, ADCY9*

Driscoll, Blood 2003
Sebastiani, Am J Haem 2012
Kwiatkowski BJHaem 2003
Ohene-Frimpong, Blood 1998
Cox, BJHaem, 2014
Bernaudin, Blood 2008
Flanagan, Blood 2011 & 2013

Heritability of silent cerebral infarcts

- No family studies
- One GWAS as an ASH abstract
  - 570 children from SIT cohort
  - 151 cases and 343 controls after QC
    - NOM1
    - FRMD4A
    - ADAM10
- No evidence α-thal protective
- Other associative features
  - Lower steady state Hb
  - Inconsistent association with lower HbF
  - Male sex
  - Relative increase in blood pressure
  - Inconsistent association with G6PD deficiency

Kinney, Pediatrics, 1999
Bhatnagar, Blood 2009

Genome Wide Association Study

Use of a pre-existing SNP array dataset available in our research group.

- 832 patients with complete variant datasets (~16.7million SNPs)
- Establish precise categorization of cerebrovascular disease of each patient

Limitations

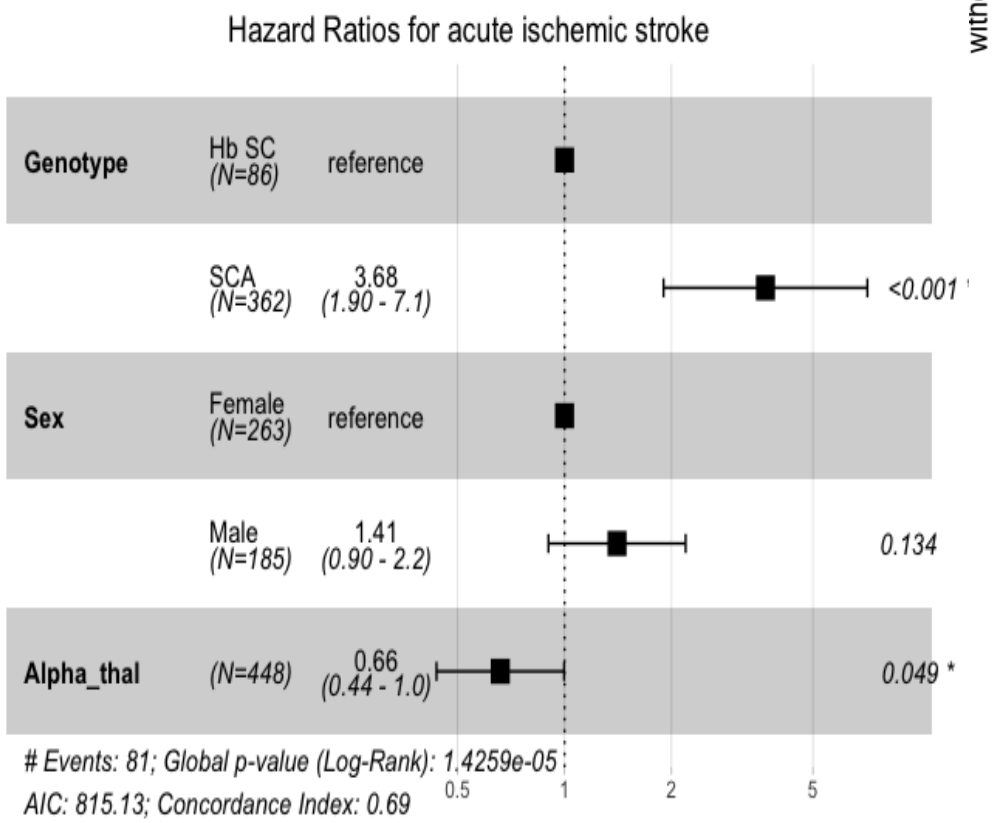
- Small cohort with respect to GWAS studies
- Variable age of incidence of cerebrovascular events

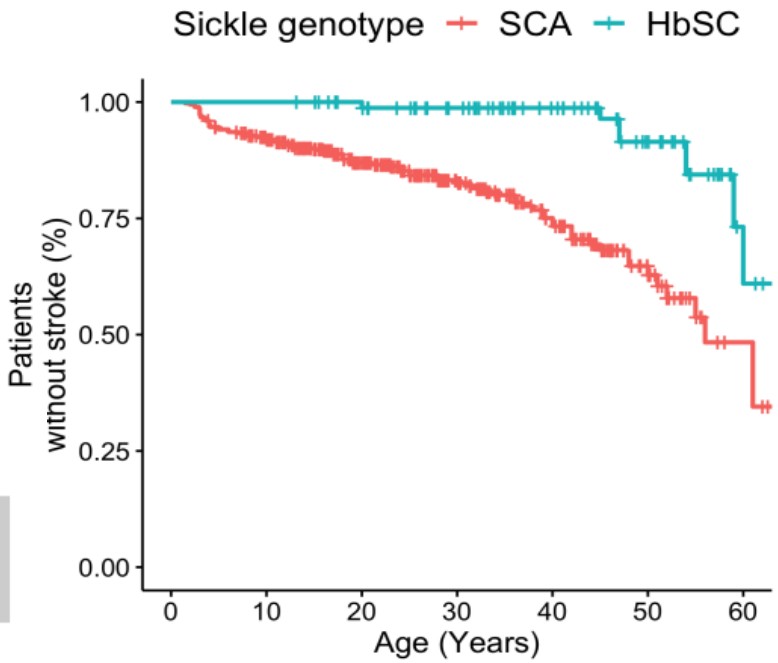
Cohort summary information

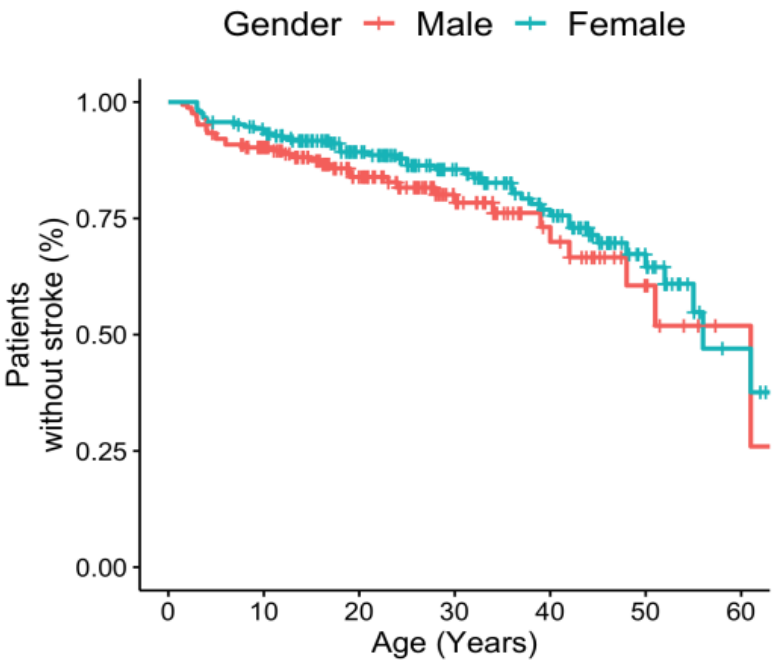
- 516 patients, average age 34.9yrs (range 10.8 to 83.6yrs)
  - 295 female, 221 male
  - 410 HbSS or HbS/B^0, 100 HbSC, 6 HbS/B^+,
- α-thal found in 198 of 509 cases (38.9%)
- G6PD in 35 of 329 cases (10.6%)

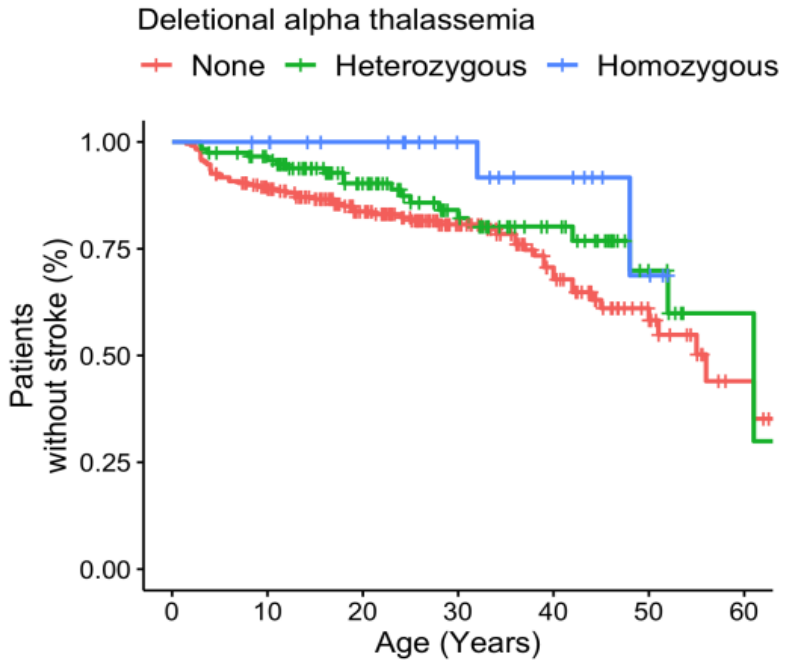
- Ischaemic Stroke in 81 (av age 28yrs, 1.5-70yrs) of 455 cases (17.5%)
- Vasculopathy in 118 (av age 20yrs, 1.5-61yrs) of 503 cases (16.6%)
- SCI in 222 (average age 31yrs, 1.8-75ys) of 415 case (53.9%)

Overt ischaemic stroke

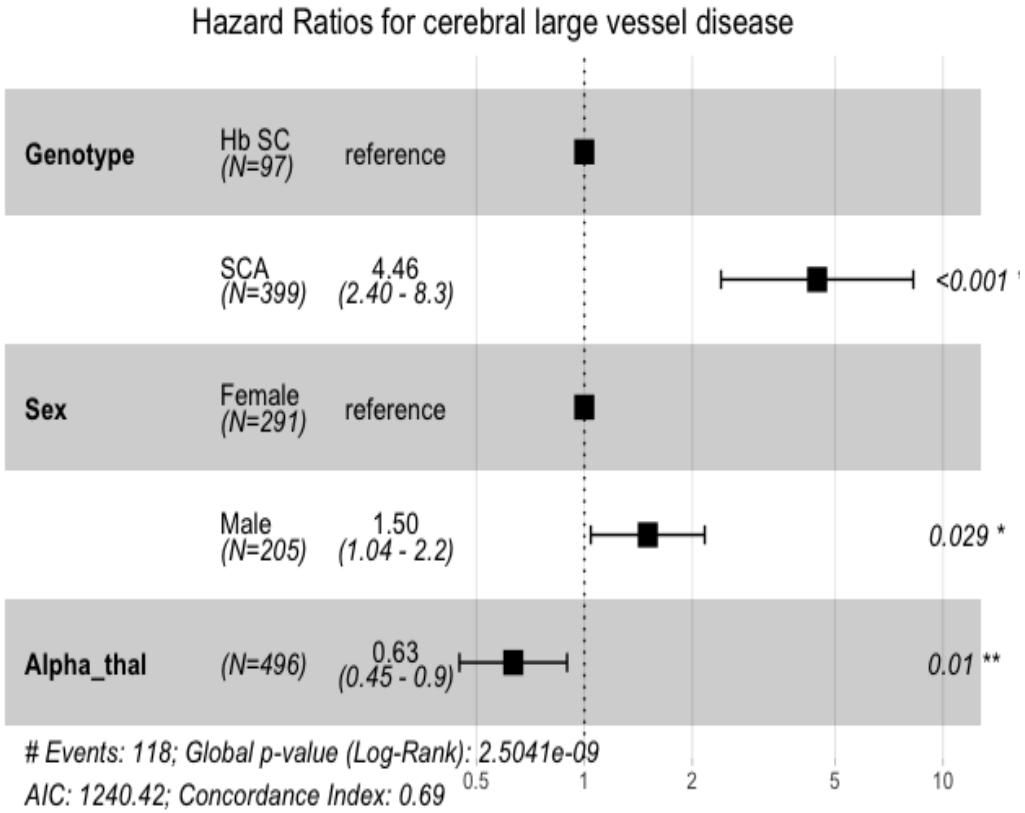


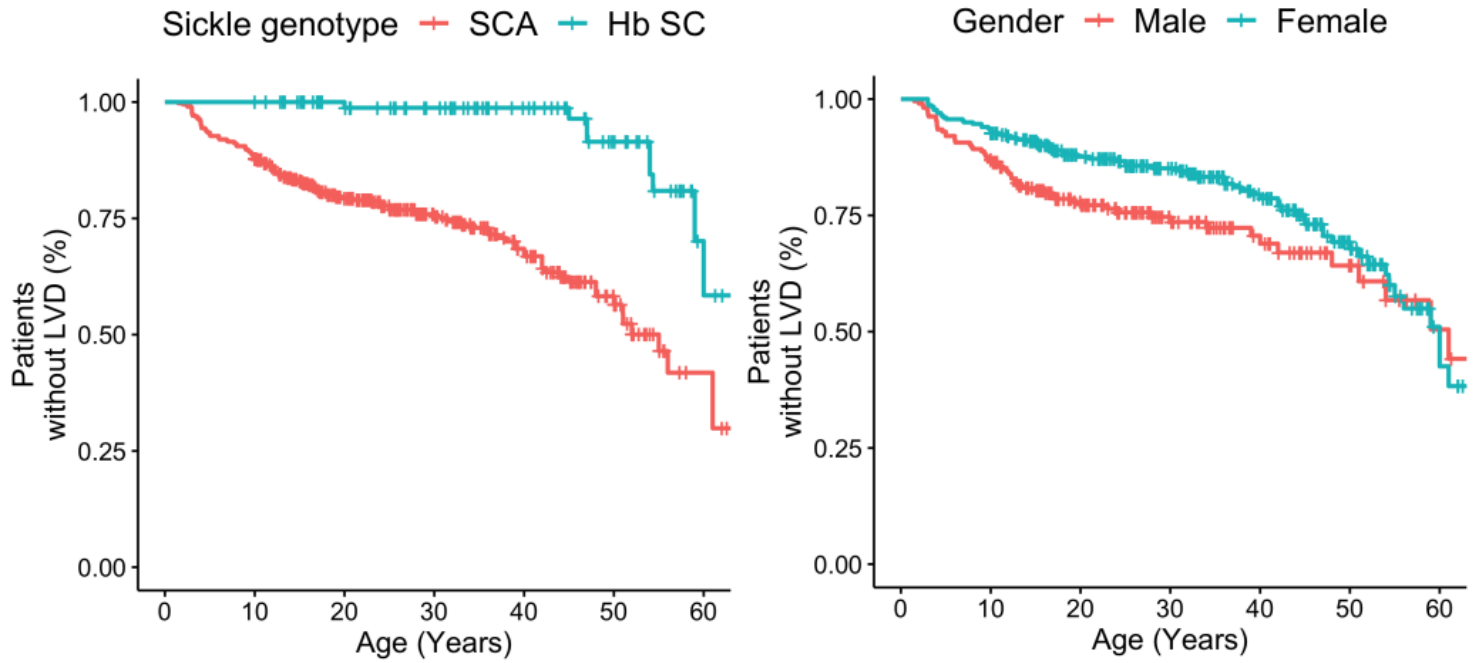


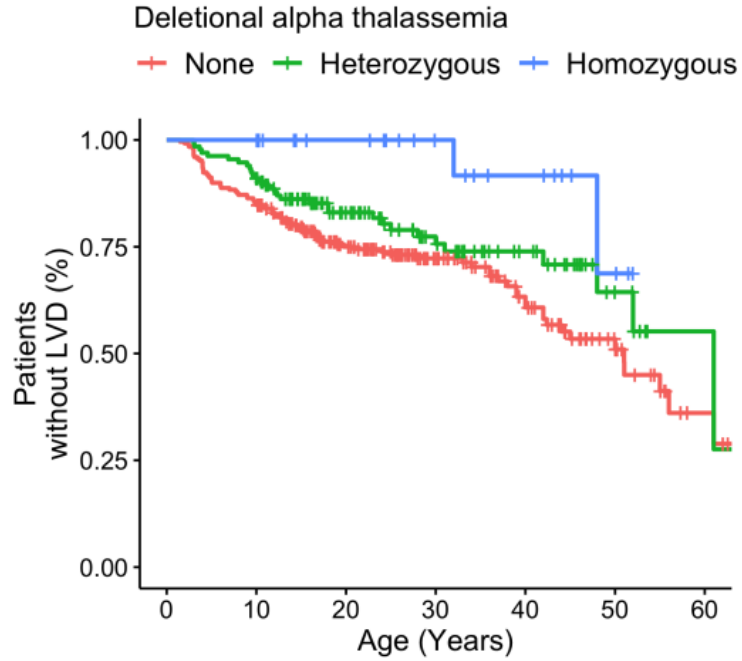




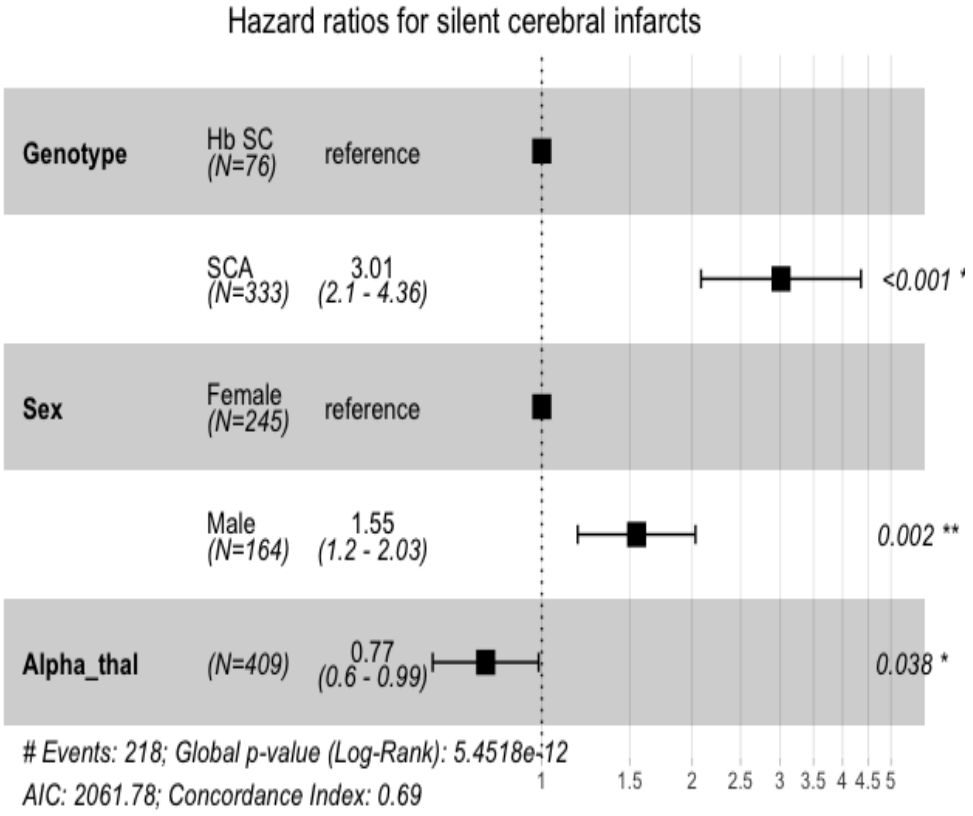
Large Vessel Disease

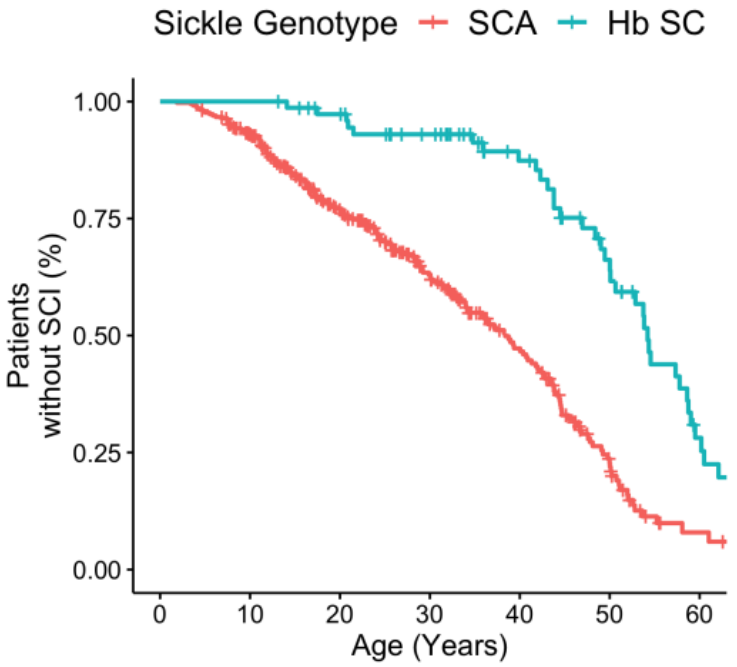


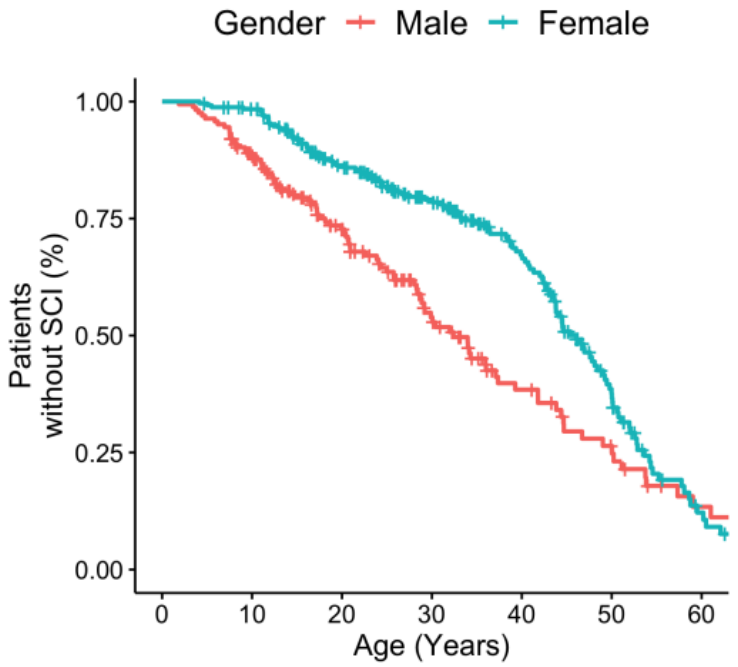


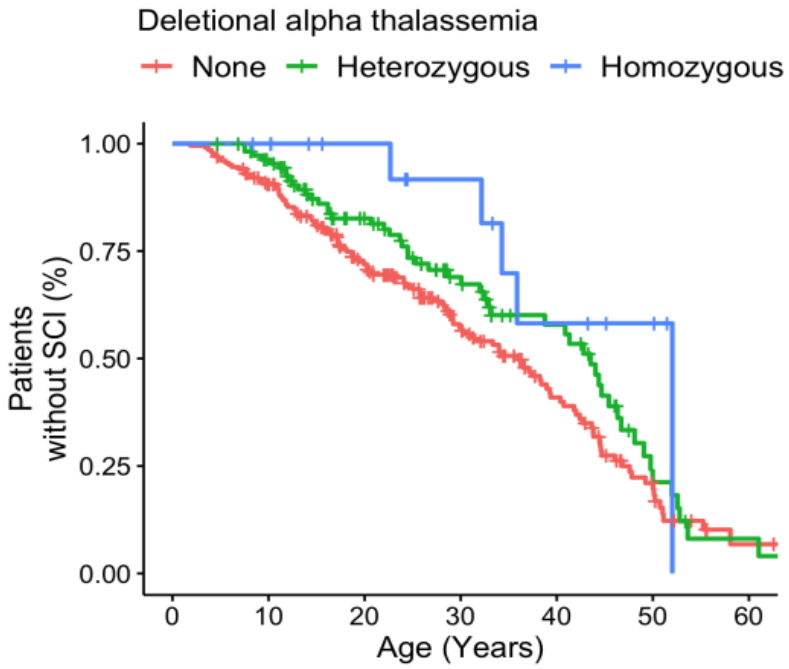


Silent cerebral infarcts









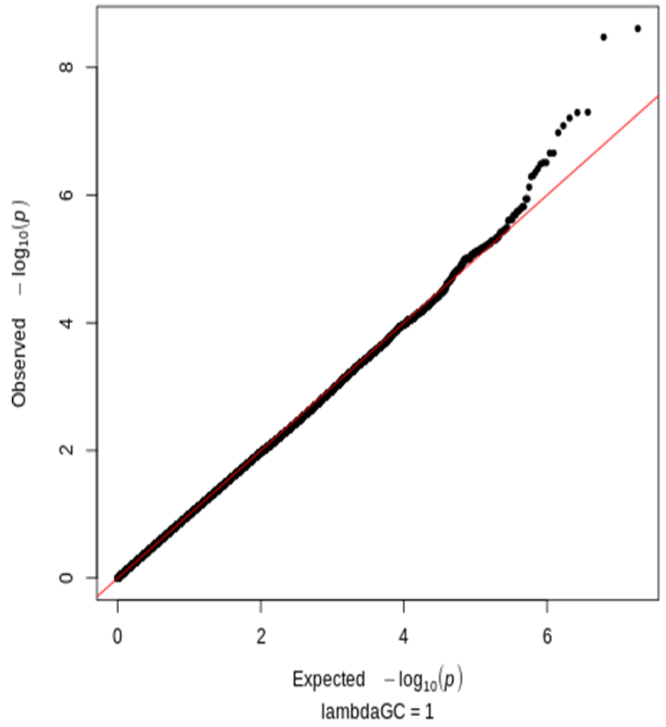
α-thal and TCD scans

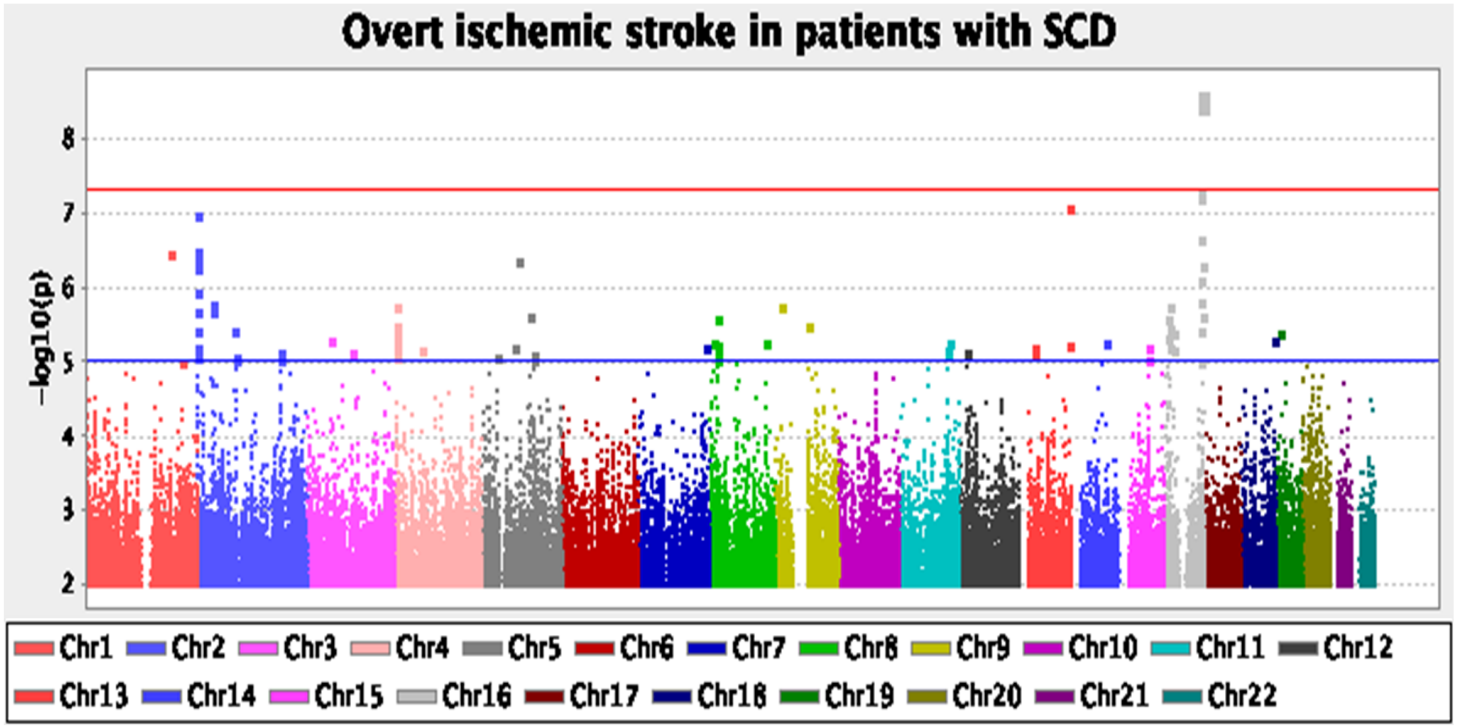
- Mean difference in TCD velocity 7.17cm/s (95%CI 4.12-10.2, $p=4.02x10^{-9}$)
- Between 4-7.5yrs mean difference was 9.58cm/s, $p=6.99x10^{-4}$

Genome wide analysis

- Linear mixed modelling
  - Genomic relatedness matrix to account for inter-relatedness
  - Age, sickle genotype, gender, and α-thal as covariates
  - Minimum minor allele frequency >0.05

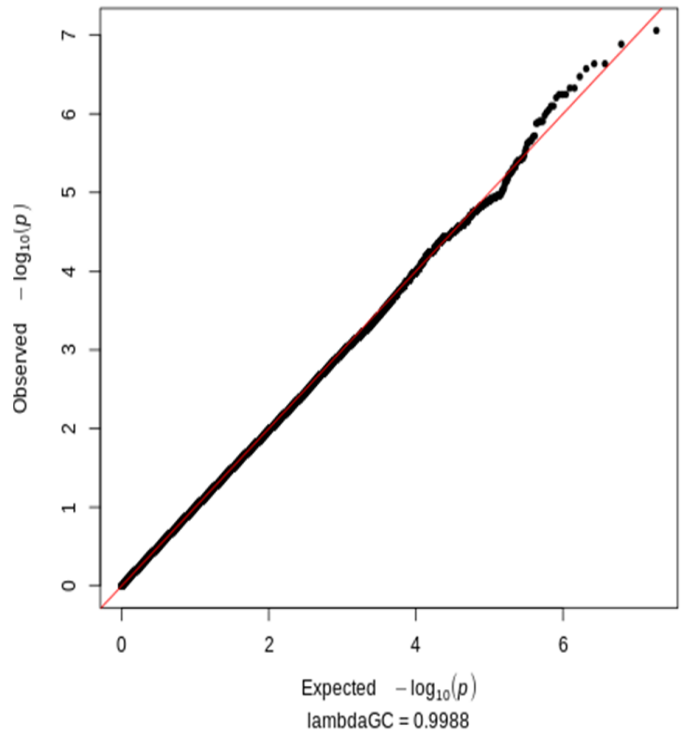
Overt Ischaemic Stroke (n=447)

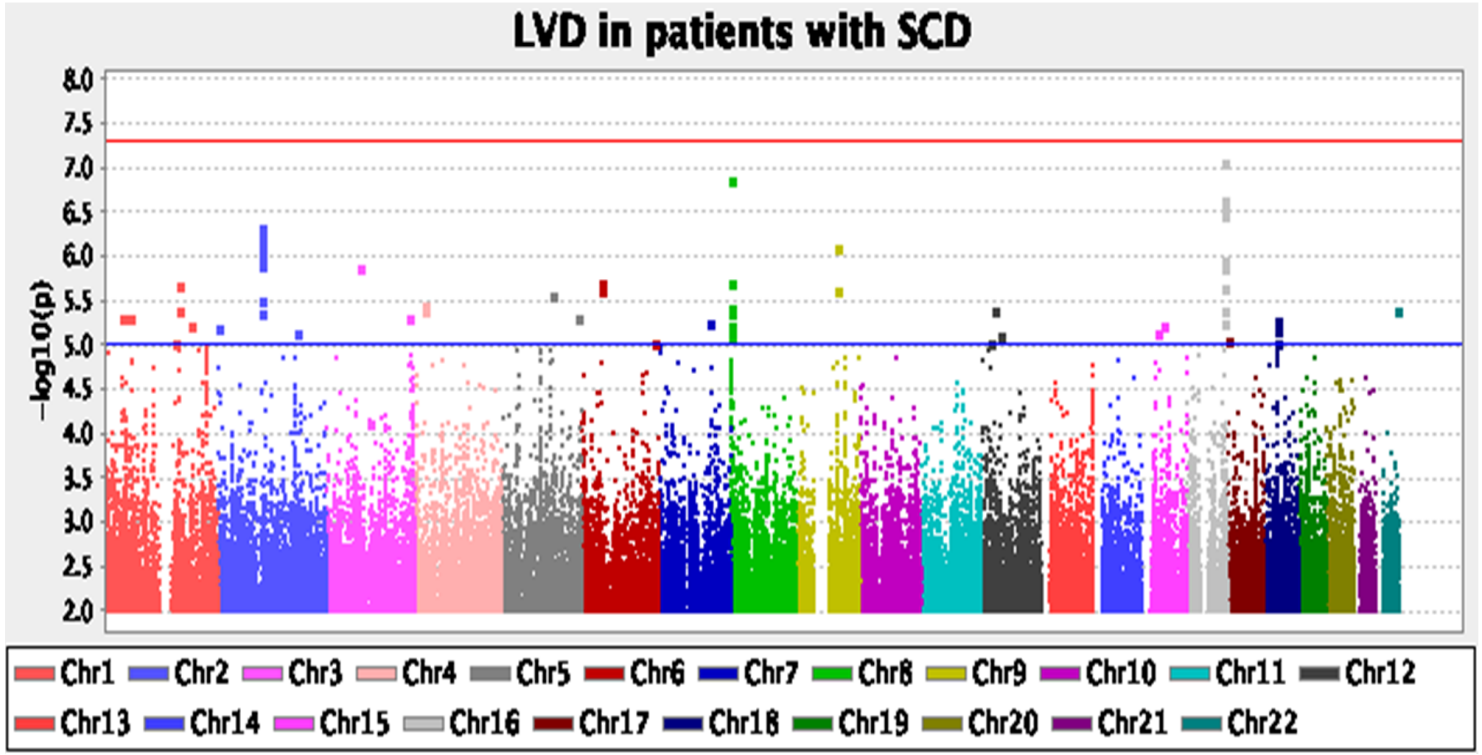




Gene	Chr	rsID	Change	Freq	OR	p	Variant location type	Gene Function
CMIP	16	rs73598466	G>A	0.068	3.81	2.45E-09	Intronic	T cell regulator
SOX1	13	rs78048662	G>A	0.058	3.36	8.16E-08	Upstream Variant	Transcriptional regulation
FAM110C	2	rs114888422	C>T	0.173	2.1	1.05E-07	TF Binding site	Microtubule organization
VCAN	5	rs73144204	T>A	0.072	2.89	4.39E-07	Upstream Variant	Extracellular matrix adhesin
CDH13	16	rs13338739	T>G	0.175	1.97	4.94E-07	Intronic	Cell adhesion in plasma membrane

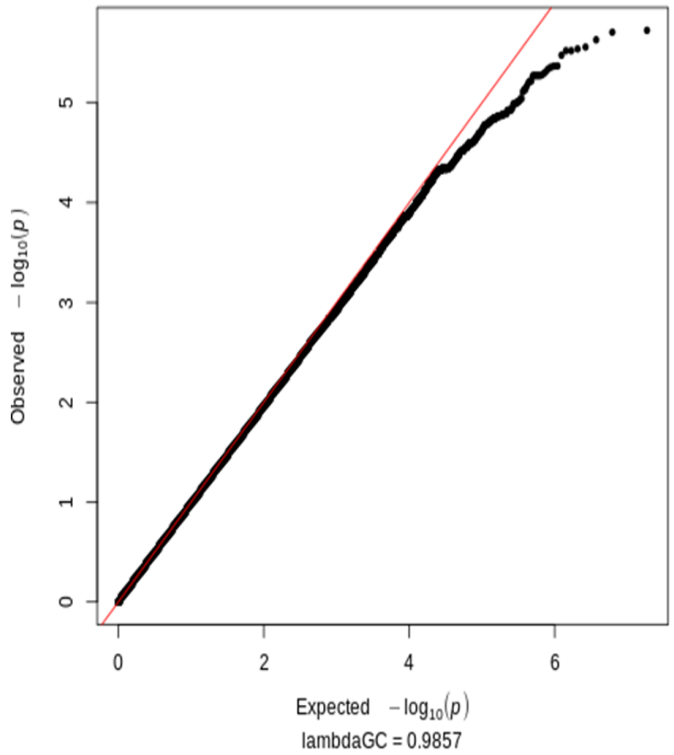
Large Vessel Disease (n=494)

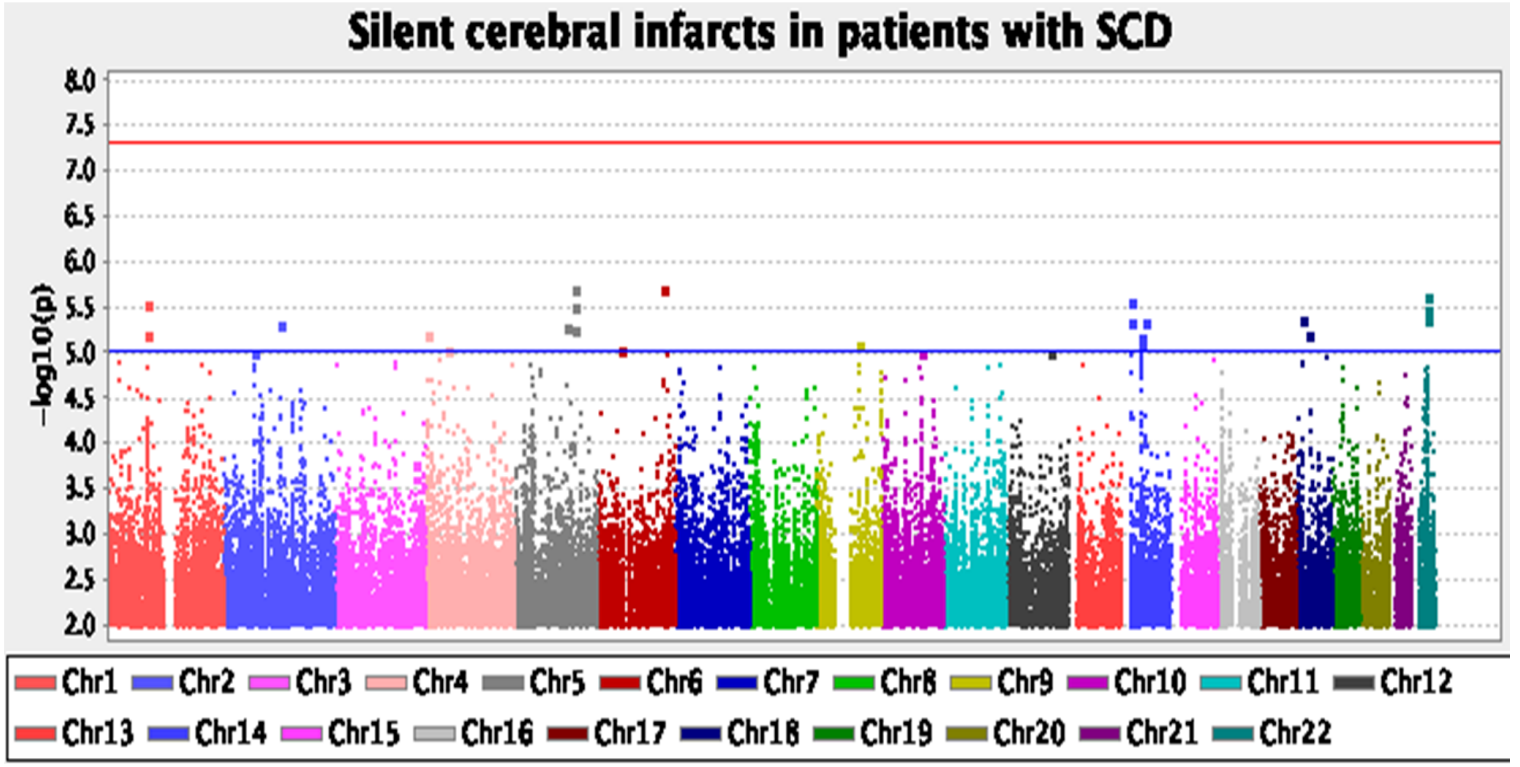




Gene	Chr	rsID	Change	Freq	OR	p	Variant location type	Gene Function
CDH13	16	rs13338739	T>G	0.175	2.13	8.72E-08	Intronic	Cell adhesion in plasma membrane
MYOM2	8	rs2010702	A>G	0.222	1.97	1.30E-07	Downstream	Component of myofibrillar M band
CMIP	16	rs73598466	G>A	0.068	3.24	2.67E-07	Intronic	T cell regulator
KIAA1211L	2	rs149912238	GT>T	0.132	2.21	4.70E-07	Intronic	No data
SHC3	8	rs9410292	A>G	0.124	2.26	7.97E-07	Intronic	Growth factor signalling in neurons

Silent Cerebral Infarcts (n=403)





Gene	Chr	rsID	Change	Freq	OR	p	Variant location type	Gene Function
CHSY3	5	rs1557759	G>A	0.222	0.44	1.88E-06	Upstream variant	Glycosyltransferase
PHACTR2	6	rs6930487	G>A	0.388	1.94	1.96E-06	Intronic	Platelet response to cytosolic Ca^{2+}
None	22	rs201658643	G>GTA	0.318	2.08	2.34E-06	Intergenic	NA
TOX4	14	rs10142478	A>C	0.067	0.25	2.76E-06	Intronic	Chromatin Binding
PRKACB	1	rs2250806	A>G	0.415	0.51	2.88E-06	Intronic	Mediates cAMP-dependent signalling

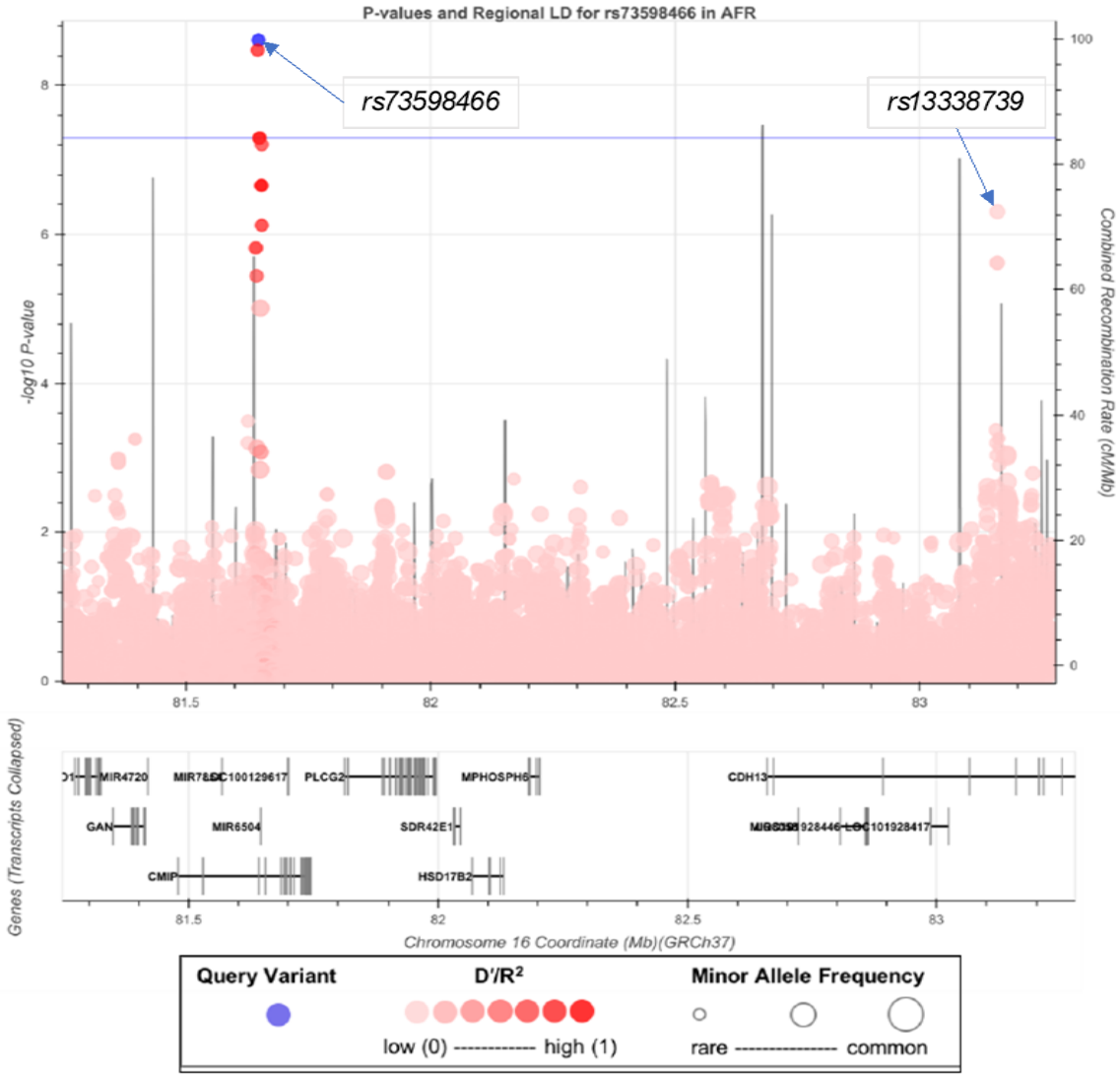
Two candidate variants on chr 16

rs73598466 in CMIP

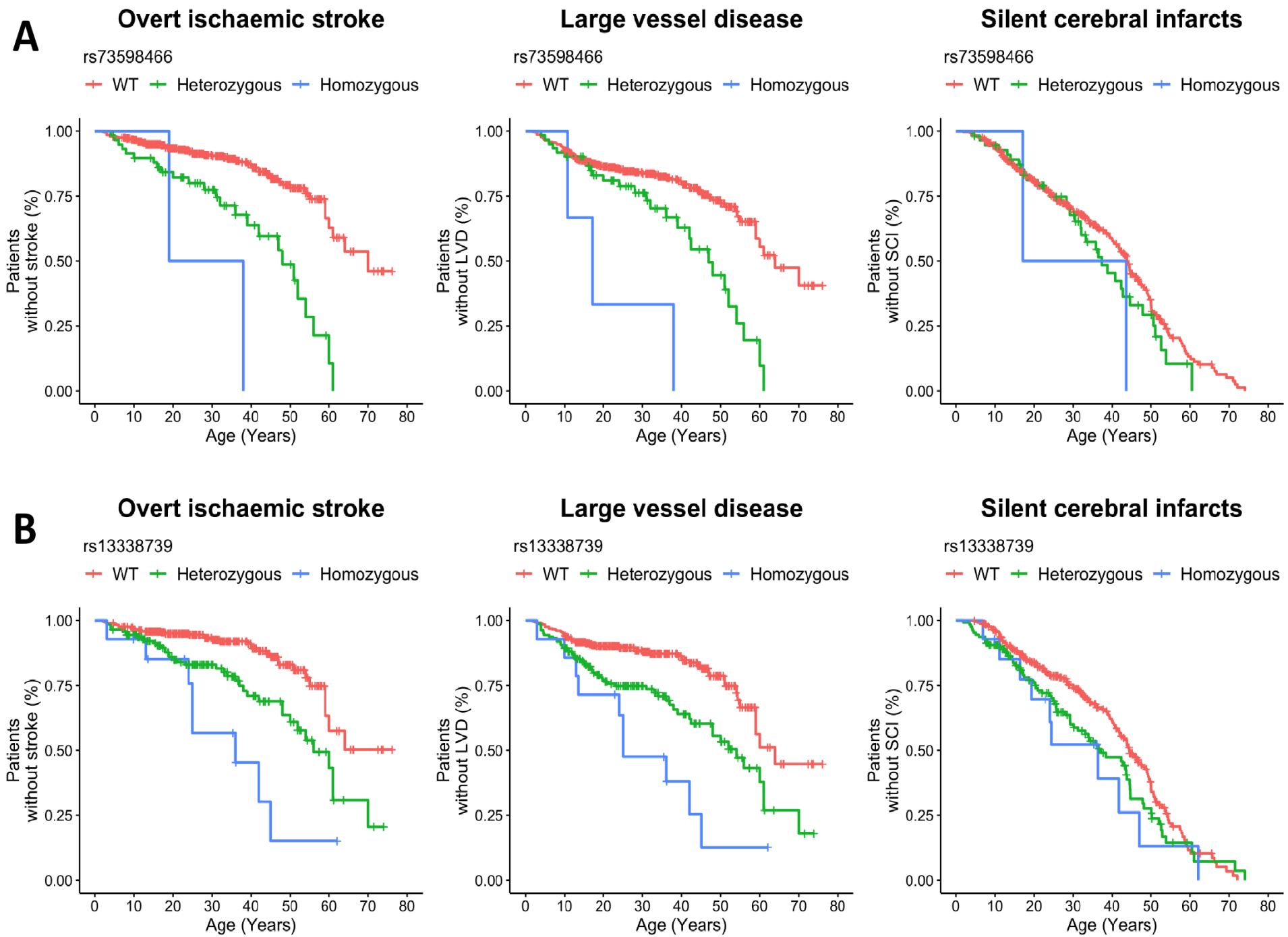
- p=2.45x10^{-9} in OIS analysis
- Gene involved in T cell signaling.
- SNPs associated with cardiovascular disease, cholesterol, and type 2 diabetes melitus

rs13338739 in CDH13

- p=8.72x10^{-8} in LVD analysis
- Gene involved in vascular endothelial response to oxidative stress



Effect of variants across the three phenotypes



Validation study

- 84 DNA samples from Henri Mondor Hospital, Paris (Bartolucci group)
  - All patients HbSS adults
  - 54 with cerebrovasculopathy, 30 without

- CMIP gene

SNP	Odds Ratio	95% CI	p
rs16955659	1.7	0.61 - 4.74	0.4785845

  - Limited sample size
- Validation not yet performed on CDH13

Evaluation of variants identified in previous studies

- 35 variants influencing OIS/LVD
  - Genes: ECE1, TGFBR3, VCAM1, SELP, AGT, GOLGB1, CSF2, ADRB2, LTC4S, BMP6, TNF, ENPP1, SOD2, PON1, BAI1, LPL, TEK. MMP3, NINJ2, ANXA2, ADCY9, IL4R, CCL2, ERG

  ➢ *BMP6*: *rs267201*.A>G (OR=1.26, p=0.039) was associated with OIS, whilst *rs449853*.C>T (OR=0.4, p=0.005) was protective against OIS
  ➢ PON1: rs662.C>T was associated with OIS in childhood (OR=2.70, p=0.007)

- 5 variants influencing SCI
  - Genes: VCAM1, ADAMTS10, NOM1, FRMD4A, CACNB2
  ➢ None showed significant association

- HbF levels, 11 known genetic markers individually and g(HbF)
  ➢ HbF% did not influence SCI (OR=0.80, 95%CI=0.51-1.09, p=0.126)
  ➢ No association of genetic markers, nor g(HbF) with any phenotype

Key findings

- α-thal offers ~1.5x risk reduction against SCI
- CVD in childhood extremely rare with homozygous α-thal
- SCIs are common in adults with HbSC
- No evidence genetic modulators of HbF influences CVD outcomes
- Male sex associated with increased risk of OIS and SCI
- No evidence G6PD deficiency influences CVD outcomes

- CMIP and CDH13 are candidate genes for further investigation
  - Both have roles in lipid metabolism
- Replication of variants in BMP6 and PON1
  - BMP6 possible role in neuroinflammation (TGFβ/SMAD/BMP pathway)
  - PON1 linked to role in atherosclerosis and OIS in general population

Whole exome sequencing in children with stroke at a very young age

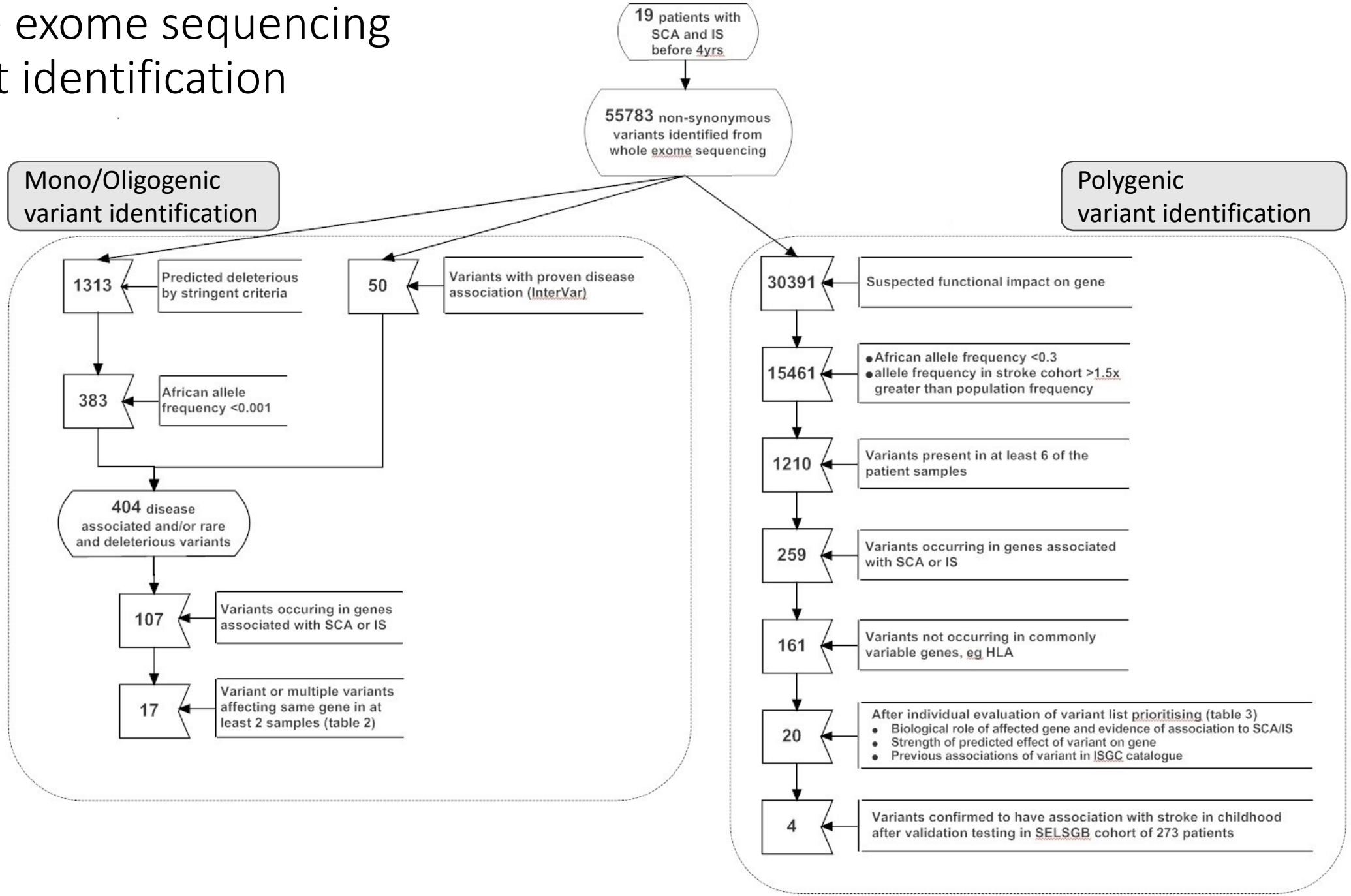
- Inclusion criteria
  - HbSS
  - MRI/A evidence of ischaemic stroke + clinical documentation of event having taken place <4yrs of age
- Exclusion criteria
  - Precipitating factor contributing to stroke
- Deletional α-thal genotyping
  - Multiplex PCR performed according to published methods
- Whole exome sequencing
- Analysis of variants using pathway analysis

Clinical features of the very young stroke cohort

- 19 patients recruited to study (from population of about 5000 children)
- Average age of stroke 3.3yrs (2.0-4.0yrs), m=10, f=9
- Most common infarctive territory was middle cerebral artery
- 7 of 17 patients had progressive CVD despite adherence to transfusion programme

- 3 (15.8%) aa/a- (expected frequency of ~38%)
- 4 male patients had G6PD
- Steady state blood results age 1 year(n=10)
- Hb 66 (54-74) g/L
  - Retic% 14.9 (5.9-19.0)%
  - HbF 12.8 (4.1-22.8)%
  - WCC 17.3 (3.1-41)x10^9/L

Whole exome sequencing variant identification



Variants identified under a monogenic/oligogenic model

Gene	Number of variants identified	Number of young stroke samples with at least one variant in the gene	Variants SNP ID or AA change	Gene function
CX3CR1	1	8	rs3732379	Leukocyte adhesion and migration
COL7A1	4	4	rs753443085 rs139416346 rs267599859 rs756703259	Extracellular matrix protein
TENM3	3	4	R1793L rs536479740 rs902566447	Regulation of neuronal development
FDFT1	2	2	rs199864384 rs748267738	Cholesterol synthesis
LRP1	2	2	E1626Q P3309L	LDL receptor
TSC2	2	2	rs876659800 G703V	Tuberous sclerosis complex protein
APOE	2	13	rs7412 rs429358	Lipoprotein and omega fatty acid metabolism
G6PD	1	7	rs1050828	NADPH synthesis

APOE variants:

- *rs7412* & *rs429358* occur in opposition
- define alleles APOE ε2 and ε4 respectively,
  - ε3 is defined by the absence of both
- Strongly linked with dyslipidemia and stroke in general population

Additional lipidaemic genes identified

- LRP1 (LDL-receptor 1)
- FDFT1 – cholesterol biosynthesis

Candidate genes identified using polygenic inheritance model

GRCh37 coordinates	SNP ID	Gene	AA Change	MAF in young stroke cohort (gnomAD african MAF)	Odds Ratio (95% CI) determined using SELSGB cohort	p-value
1:183091058	rs2230157	LAMC1	E731K	0.18 (0.10)	0.62 (0.28-1.4)	0.2340
2:169824983	rs11568364	ABCB11	M677V	0.26 (0.17)	0.95 (0.51-1.75)	0.8571
2:216235046	rs6728999	FN1	R2275Q	0.21 (0.12)	1.43 (0.7-2.93)	0.3390
3:39307256	rs3732379	CX3CR1	V249I	0.26 (0.13)	1.15 (0.58-2.28)	0.6864
4:88903774	rs7435825	SPP1	S210N	0.26 (0.15)	0.75 (0.38-1.47)	0.3962
6:31639979	rs9267532	LY6G5B	R176C	0.18 (0.10)	0.94 (0.44-2)	0.8758
10:54531226	rs1800451	MBL2	G57E	0.37 (0.23)	1.46 (0.87-2.46)	0.1554
11:71191817	rs7121106	NADSYN1	P297L	0.26 (0.14)	0.93 (0.5-1.71)	0.8143
11:82564294	rs2229437	PRCP	E133D	0.18 (0.12)	0.98 (0.44-2.19)	0.9676
12:6061675	rs78353028	VWF	T2666M	0.16 (0.07)	1.52 (0.6-3.84)	0.3886
12:6145649	rs57950734	VWF	H817Q	0.21 (0.12)	0.66 (0.31-1.41)	0.2612
13:113773159	rs6046	F7	R413Q	0.21 (0.12)	1.65 (0.86-3.14)	0.1412
15:39874396	rs41515347	THBS1	S24A	0.18 (0.11)	0.83 (0.38-1.82)	0.6455
17:26096597	rs2297518	NOS2	S608L	0.21 (0.12)	**2.25 (1.21-4.19)**	**0.0139**
17:40451844	rs2230123	STAT5A	V209A	0.24 (0.07)	**2.6 (1.3-5.2)**	**0.0090**
19:11489049	rs35423344	EPOR	P380A	0.32 (0.20)	1.35 (0.81-2.24)	0.2587
19:15272218	rs114447350	NOTCH3	P2074L	0.18 (0.09)	0.92 (0.42-2)	0.8346
19:45411941	rs429358	APOE	C156R	0.34 (0.22)	**1.95 (1.13-3.34)**	**0.0168**
X:153764217	rs1050828	G6PD	V68M	0.29 (0.12)	0.83 (0.26-2.52)	0.7410

rs2297518.G>A, NOS2

- OR 2.25, 95%CI 1.21-4.19, p=0.0139

rs2230123.T>C, STAT5a

- OR 2.6, 95%CI 1.3-5.2, p=0.0090

rs429358.T>C, APOE

- OR 1.95, 95%CI 1.13-3.34, p=0.0169

Validation of findings in population with strokes at an older age

- Validation of APOE findings in cohort of 832 patients with SCD and stroke
  - ε2: OR 0.97 95%CI 0.49-1.94, p=0.940
  - ε3: OR 0.59, 95%CI 0.36-0.97, p=0.038
  - ε4: OR 1.95, 95% CI 1.13-3.34, p=0.016
- ε2ε4 or ε4ε4
  - OR 4.35 (95% CI 1.85-10.0, p=0.0011)

APOE and stroke risk

- APOE alleles ε2ε4 or ε4ε4 confer 4.35x increased risk of ischemic stroke in children with SCA
  - ~10% of sickle population will carry this genotype
  - Alleles known to influence many aspects of human biology
    - Plasma cholesterol levels
    - Omega-3-fatty acid (docosahexanoic acid) metabolism
    - Neuroinflammation (Alzheimers)
    - Associated with stroke risk in general population by large scale GWAS
  - Body of evidence to suggest biologically plausible effect in SCA

Conclusions – Genetic associations

- *APOE* genotypes ε2/ε4 or ε4/ε4 confer a significantly increased risk of overt stroke occurring in childhood
  - OR=4.35, 95% CI 1.85-10.0, p=0.0011
- Increased risk of stroke also associated with variants in
  - *NOS2*
    - OR=2.25, 95%CI 1.21-4.19, p=0.0139
  - *STAT5a*
    - OR=2.60, 95%CI 1.30-5.20, p=0.009
  - *CMIP*
    - OR_{OIS}=3.81, p=2.45x10^{-9})
  - *CDH13*
    - OR_{LVD}=2.17, p=8.72x10^{-08}
- Validation of previously reported genetic associations
  - *BMP6*
  - *PON1*

Conclusions – Known genetic modifiers

- α-thal protective against all forms of CVD, including SCI
  - Effect only seen in HbSS cohort
  - homozygous alpha thalassaemic appear to be especially protected
- No evidence for influence of HbF levels either by clinical measurement or genetic markers
- Increased risk of CVD with male sex
- No evidence for influence of G6PD deficiency
- Need for larger, multinational studies

Acknowledgements

- King’s College London
  - **John Brewin**
  - John Strouboulis
  - Stephan Menzel
  - Sara El Hoss
  - Amina Nielsen Nardo Marino
  - Sanjay Tewari
- Cambridge University
  - John Gibson
- Robert Debre Hospital
  - Valentine Brousse
- Hopital Henri Mondor
  - Pablo Bartolucci