Presenter Disclosure Information

Aniket Mishra, PhD

Exome Sequence Study on Extreme MRI Markers of Cerebral Small Vessel Disease

FINANCIAL DISCLOSURE:

No relevant financial relationship exists

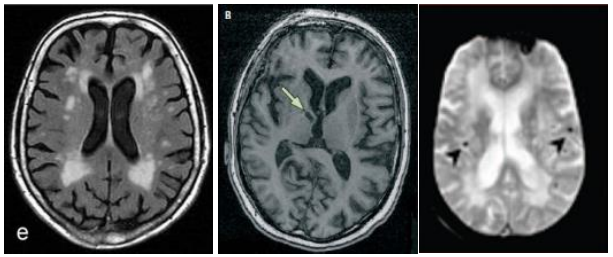
Exome Sequence Study on Extreme MRI Markers of Cerebral Small Vessel Disease

Aniket Mishra, PhD
INSERM U1219, Univ of Bordeaux, France
On behalf of the neuro-CHARGE group

université de BORDEAUX

Introduction

- Cerebral small vessel disease (SVD) is a group of pathological processes affecting **small arteries, arterioles, capillaries and small veins in the brain**
- Small vessels are **hard to image in the brain:**
- **MRI-markers** of consequences of SVD on brain parenchyma used for diagnosis:
  - White matter hyperintensity (WMH)
  - Lacunar brain infarct
  - Cerebral microbleeds



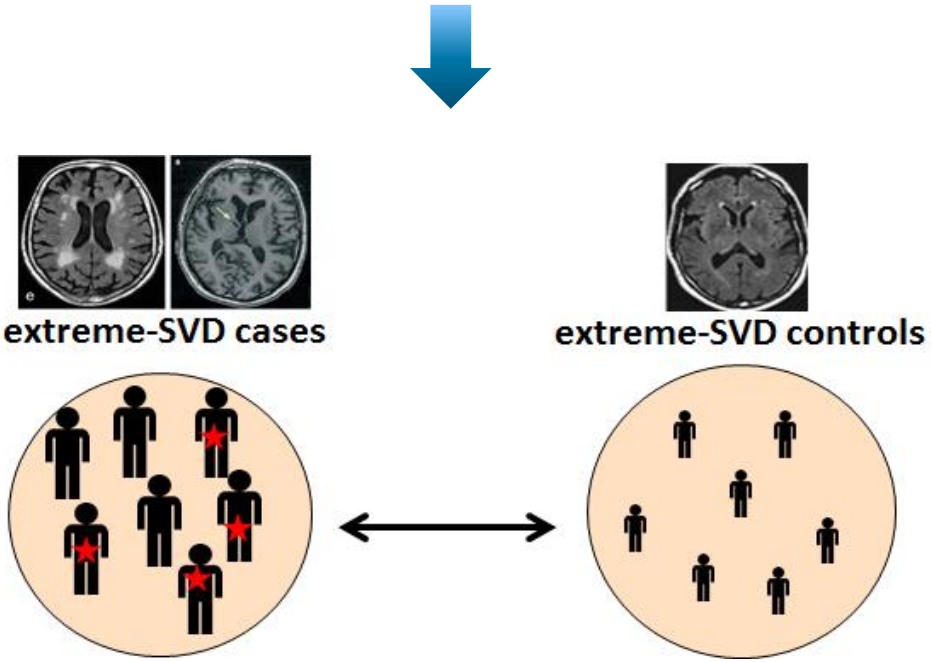
White Matter Hyperintensities (WMH) | Lacunar Brain Infarcts (Lac-BI) | Cerebral Microbleeds (CMB)

- These **MRI-markers** are associated with impairments of cognition, mood disorders and **powerful predictor of stroke, dementia and death**
- They are highly **prevalent in older community persons** and highly **heritable**: WMH volume = 0.76, lacunar brain infarct = 0.35
  - *Five WHM burden risk loci* (common variants) and *no robustly reported* risk loci for other MRI-markers of SVD

Wardlaw, *Lancet Neurol.*, 2013
Verhaaren, *Circ. Cardiovasc. Genet.*, 2015

Objective

- ➔ Identify novel genetic determinants for MRI-markers of cerebral SVD
- ✓ Using **next generation sequencing** to explore role of **rare variants**
- ✓ Taking into account **several markers simultaneously** – **avoid misclassification** due to presence of one SVD marker in the control group of another SVD marker
- ✓ And using an "**extreme phenotype**" approach – **more power** for rare variants



Peloso et al., *Eur. J. Hum. Genet.*, 2016

Study population

- 3C-Dijon: A population based cohort study, Dijon, France
  - Participants aged 65+ years

1,527 participants with brain MRI + genome-wide genotypes, (excluding individuals with brain tumor, dementia, stroke)

382 individuals in upper quartile of WMHV residual*

382 individuals in bottom quartile of WMHV residuals*

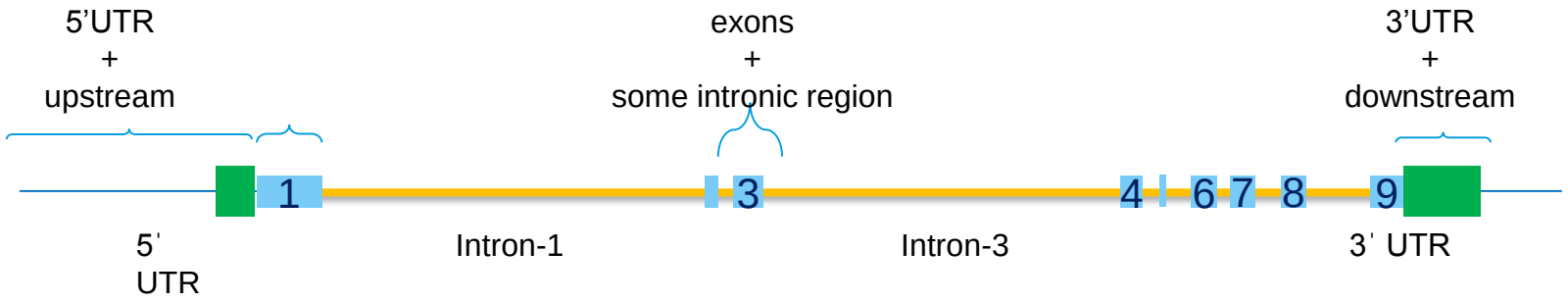
261 cases = 58 individuals with lacunar brain infarct + 203 top ranking individuals in upper quartile for WMH volume residuals*

253 controls = 253 low ranking individuals in bottom quartile for WMH volume residuals* without any brain infarct

**WMHV residuals adjusted for age, sex and intracranial volume*

3C Study Group, *Neuroepidemiology*, 2003

Whole exome sequencing (WES)



- **Whole Exome Sequencing** was carried out at the Genome Quebec Innovation Centre at McGill University, Montreal, Canada using Illumina HiSeq200 PE100.

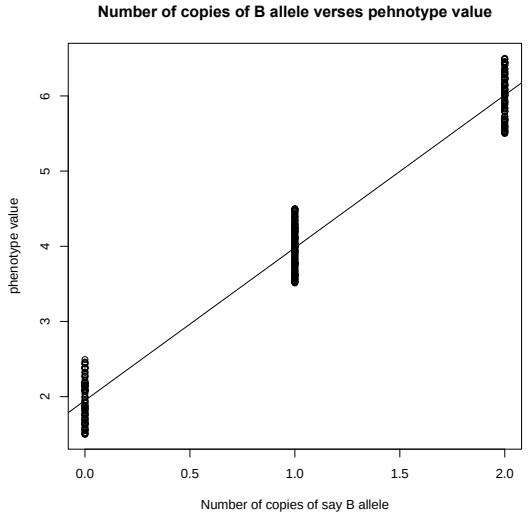
Exome capture kit	: SureSelectXT Human All Exon V5
Reads	: Paired-end sequencing (2 × 100 bp)
Seq. Instrument	: Illumina HiSeq200 PE100
Align to human RG	: Burrows-Wheeler Aligner (BWA)
Realigning and QC	: Genome Analysis Toolkit (GATK)

- Average **depth coverage: 100X**
- Genotype calling using the standard GATK pipeline

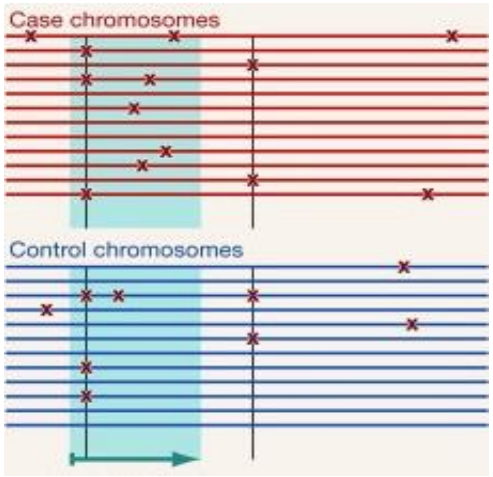
Statistical analysis outline

- **Candidate gene approach:** whether genetic variants in genes causing Mendelian cerebral SVD (CADASIL, CARASIL, COL4A1/2 syndrome, RVCL) influence the risk of extreme-SVD in older community persons
  - Genes: *NOTCH3*, *HTRA1*, *COL4A1*, *COL4A2* and *TREX1*
  - ➢ Step 1: Survey of pathogenic variants in extreme-SVD
  - ➢ Step 2: Association tests

1) <u>Single variant association test</u>



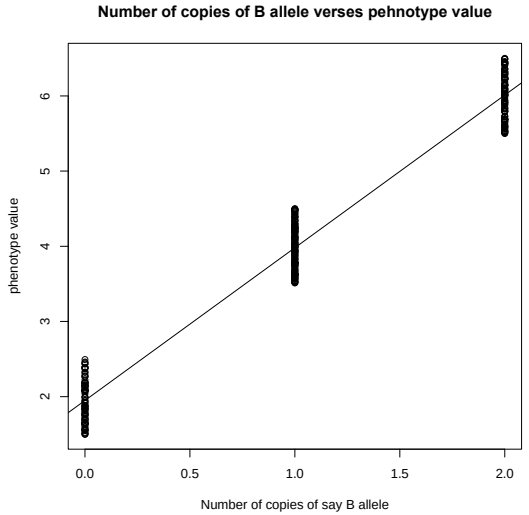
2) <u>Rare-variant gene-based test</u>



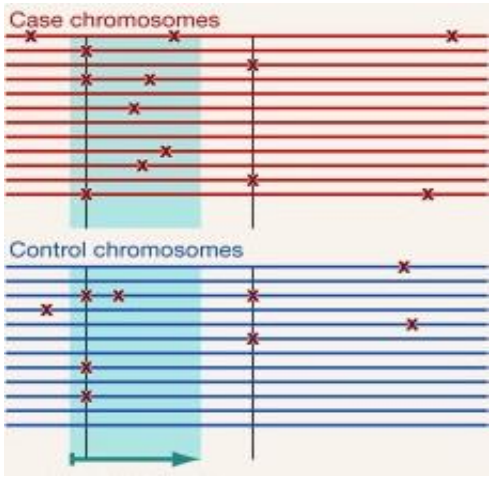
Statistical analysis outline

- **Candidate gene approach:** whether genetic variants in genes causing Mendelian cerebral SVD (CADASIL, CARASIL, COL4A1/2 syndrome, RVCL) influence the risk of extreme-SVD in older community persons
  - Genes: *NOTCH3*, *HTRA1*, *COL4A1*, *COL4A2* and *TREX1*
  - Step 1: Survey of pathogenic variants in extreme-SVD
  - Step 2: Association tests

1) Single variant association test



2) Rare-variant gene-based test



- Step 3: Replication within Cohorts for Heart and Ageing Research in Genomic Epidemiology: Atherosclerosis Risk in Communities (ARIC) study, Cardiovascular Health Study (CHS), Framingham Heart Study (FHS) and Rotterdam Study.

Results: Population characteristics

Characteristics	Cases	Controls	P-value
Participants, N	259*	253	
Age at MRI, years (mean±SD)	**73.50 ± 4.01**	**73.19 ± 4.45**	**0.4**
Women, N (%)	**150 (58.1%)**	**155 (61)**	**0.51**
WMH, ml (mean±SD)	13.18 ± 7.07	2.05 ± 0.63	<0.0001
Presence of lacunar brain infarct	58 (22.4%)	0	NA
Systolic BP, mmHg (mean±SD)	152.05 ± 22.51	147.07 ± 21.85	0.011
Hypertension, N (%)	223 (86.4%)	184(72.4)	<0.0001
Antihypertensive drug, N (%)	146 (56.6)	93(36.6)	<0.0001
Fasting plasma glucose, mmol/L (mean±SD)	5.18 ± 1.51	4.95 ± 0.67	0.026
Diabetes Mellitus, N (%)	25 (9.7)	14(5.5)	0.07
HDL, mmol/L (mean±SD)	1.64 ± 0.39	1.68 ± 0.41	0.23
LDL, mmol/L (mean±SD)	3.53 ± 0.89	3.68 ± 0.84	0.046
TG, mmol/L (mean±SD)	1.26 ± 0.56	1.15 ± 0.52	0.031
Lipid lowering drug, N (%)	96 (37.2)	71(28)	0.026
BMI, kg/m2 (mean±SD)	25.84 ± 3.92	24.86 ± 3.71	0.004
Current smoker, N (%)	22 (8.5)	8(3.1)	0.012
History of CVD at MRI, N (%)	15 (5.8)	5(2)	0.025

**Two cases filtered out*

Results: Population characteristics

Characteristics	Cases	Controls	P-value
Participants, N	259*	253	
Age at MRI, years (mean±SD)	73.50 ± 4.01	73.19 ± 4.45	0.4
Women, N (%)	150 (58.1%)	155 (61)	0.51
WMH, ml (mean±SD)	**13.18 ± 7.07**	**2.05 ± 0.63**	**<0.0001**
Presence of lacunar brain infarct	**58 (22.4%)**	**0**	**NA**
Systolic BP, mmHg (mean±SD)	152.05 ± 22.51	147.07 ± 21.85	0.011
Hypertension, N (%)	223 (86.4%)	184(72.4)	<0.0001
Antihypertensive drug, N (%)	146 (56.6)	93(36.6)	<0.0001
Fasting plasma glucose, mmol/L (mean±SD)	5.18 ± 1.51	4.95 ± 0.67	0.026
Diabetes Mellitus, N (%)	25 (9.7)	14(5.5)	0.07
HDL, mmol/L (mean±SD)	1.64 ± 0.39	1.68 ± 0.41	0.23
LDL, mmol/L (mean±SD)	3.53 ± 0.89	3.68 ± 0.84	0.046
TG, mmol/L (mean±SD)	1.26 ± 0.56	1.15 ± 0.52	0.031
Lipid lowering drug, N (%)	96 (37.2)	71(28)	0.026
BMI, kg/m2 (mean±SD)	25.84 ± 3.92	24.86 ± 3.71	0.004
Current smoker, N (%)	22 (8.5)	8(3.1)	0.012
History of CVD at MRI, N (%)	15 (5.8)	5(2)	0.025

**Two cases filtered out*

Results: Population characteristics

Characteristics	Cases	Controls	P-value
Participants, N	259*	253	
Age at MRI, years (mean±SD)	73.50 ± 4.01	73.19 ± 4.45	0.4
Women, N (%)	150 (58.1%)	155 (61)	0.51
WMH, ml (mean±SD)	13.18 ± 7.07	2.05 ± 0.63	<0.0001
Presence of lacunar brain infarct	58 (22.4%)	0	NA
Systolic BP, mmHg (mean±SD)	**152.05 ± 22.51**	**147.07 ± 21.85**	**0.011**
Hypertension, N (%)	**223 (86.4%)**	**184(72.4)**	**<0.0001**
Antihypertensive drug, N (%)	**146 (56.6)**	**93(36.6)**	**<0.0001**
Fasting plasma glucose, mmol/L (mean±SD)	5.18 ± 1.51	4.95 ± 0.67	0.026
Diabetes Mellitus, N (%)	25 (9.7)	14(5.5)	0.07
HDL, mmol/L (mean±SD)	1.64 ± 0.39	1.68 ± 0.41	0.23
LDL, mmol/L (mean±SD)	3.53 ± 0.89	3.68 ± 0.84	0.046
TG, mmol/L (mean±SD)	1.26 ± 0.56	1.15 ± 0.52	0.031
Lipid lowering drug, N (%)	96 (37.2)	71(28)	0.026
BMI, kg/m2 (mean±SD)	25.84 ± 3.92	24.86 ± 3.71	0.004
Current smoker, N (%)	22 (8.5)	8(3.1)	0.012
History of CVD at MRI, N (%)	15 (5.8)	5(2)	0.025

**Two cases filtered out*

Results of Step 1: survey of pathogenic variants

- **ClinVar / OMIM database** *(downloaded on 27th Feb 2017)*
  - **One extreme-SVD case** carries a heterozygote genotype at (NM_002775.4[**HTRA1**]:c.1108C>T [p.Arg370Ter]).
  - Recessive genotype at this locus reported as pathogenic for **CARASIL**
  - Not reported yet in dominant forms of *HTRA1*-related SVD

- **Cysteine residue altering variants** of *NOTCH3* EGFr domain, typical of **CADASIL**
  - **One extreme-SVD case** carries a heterozygote genotype at (NM_000435.2 [**NOTCH3**]:c.C2353T:p.R785C).
  - Reported previously in one Italian family with CADASIL-like phenotype

Hara et al. *N Engl J Med,* 2009
Verdura et al, *Brain,* 2015
Mosca et al. *J Mol Neurosci,* 2014

Results of Step 2: Single variant association test

- Variants within **100kb of the 5' and 3' untranslated region (UTR)** of the 5 candidate genes: ***NOTCH3, HTRA1, COL4A1, COL4A2,* and *TREX1***
  - Minor allele frequency ≥0.05 *(low power for single variant analyses on rare variants)*
  - Total 361 common variants across the 5 genes
- Additive model (genotype = number of risk alleles 0, 1 or 2)
  - ➔ *Extreme-SVD* (1/0) ~ genotype + age + sex + four principal components
  - ➔ Software: SeqMeta (v1.6.6)

Gene	Top Variant	Hg19_chr:bp	RA/ OA	RA Freq	Odds ratio	95% CI	P-value*
HTRA1	**rs2293871**	**10:124273671**	**T/C**	**0.194**	**1.92**	**(1.39-2.65)**	**8.21×10^{-5}**
COL4A1/ COL4A2	rs2275842	13:110813523	T/C	0.167	1.52	(1.09-2.11)	0.013
NOTCH3	rs1043997	19:15300136	G/A	0.052	1.69	(0.96-2.96)	0.067
TREX1	rs78159609	3:48419898	A/G	0.055	0.62	(0.36-1.08)	0.09

**Significance threshold is p-value<1.385×10^{-04} correcting for 361 common variants tested*

Results of Step 2: Single variant association test

- Variants within **100kb of the 5' and 3' untranslated region (UTR)** of the 5 candidate genes: ***NOTCH3, HTRA1, COL4A1, COL4A2,* and *TREX1***
  - Minor allele frequency ≥0.05 *(low power for single variant analyses on rare variants)*
  - Total 361 common variants across the 5 genes
- Additive model (genotype = number of risk alleles 0, 1 or 2)
  - ➔ *Extreme-SVD* (1/0) ~ genotype + age + sex + four principal components
  - ➔ Software: SeqMeta (v1.6.6)

Gene	Top Variant	Hg19_chr:bp	RA/ OA	RA Freq	Odds ratio	95% CI	P-value*
HTRA1	**rs2293871**	**10:124273671**	**T/C**	**0.194**	**1.92**	**(1.39-2.65)**	**8.21×10^{-5}**
COL4A1/ COL4A2	rs2275842	13:110813523	T/C	0.167	1.52	(1.09-2.11)	0.013
NOTCH3	rs1043997	19:15300136	G/A	0.052	1.69	(0.96-2.96)	0.067
TREX1	rs78159609	3:48419898	A/G	0.055	0.62	(0.36-1.08)	0.09

➢ **rs2293871 is in LD with two blood eQTLs: rs876790 and rs2736928; suggesting inverse relationship between risk of SVD and expression of *HTRA1* in blood.**

Westra et al. *Nat. Genet.*, 2014

Results of Step 2: Single variant association test

- Variants within **100kb of the 5' and 3' untranslated region (UTR)** of the 5 candidate genes: ***NOTCH3, HTRA1, COL4A1, COL4A2,* and *TREX1***
  - Minor allele frequency ≥0.05 *(low power for single variant analyses on rare variants)*
  - Total 361 common variants across the 5 genes
- Additive model (genotype = number of risk alleles 0, 1 or 2)
  - ➔ *Extreme-SVD* (1/0) ~ genotype + age + sex + four principal components
  - ➔ Software: SeqMeta (v1.6.6)

Gene	Top Variant	Hg19_chr:bp	RA/ OA	RA Freq	Odds ratio	95% CI	P-value*
HTRA1	**rs2293871**	**10:124273671**	**T/C**	**0.194**	**1.92**	**(1.39-2.65)**	**8.21×10^{-5}**
COL4A1/ COL4A2	rs2275842	13:110813523	T/C	0.167	1.52	(1.09-2.11)	0.013
NOTCH3	rs1043997	19:15300136	G/A	0.052	1.69	(0.96-2.96)	0.067
TREX1	rs78159609	3:48419898	A/G	0.055	0.62	(0.36-1.08)	0.09

➢ **Ongoing replication of rs2293871 in independent population cohorts of elderly: ARIC, CHS, FHS and Rotterdam studies.**

Results of Step 2: Gene-based association test

- Considered only:
  - **Protein-modifying variants** (splice-site, stop-loss, stop-gain frame-shift or missense) with **minor allele frequency (MAF) < 0.05**
  - Genes with cumulative rare allele frequency > 0.01
- Additive model
  - ➔ *extreme-SVD* (1/0) ~ genotype + age + sex + four principal components
  - ➔ Software: SeqMeta (v1.6.6)

Gene	3C-Dijon, N=512			Replication, N=908 (ARIC, CHS, FHS and RS1)			Combined p-value
	# protein-modifying alleles	Cumulative MAF	SKAT-O p-value	# protein-modifying alleles	Cumulative MAF	SKAT-O p-value	
NOTCH3	**33**	**0.103**	**1.578×10^{-2}**	**36**	**0.108**	**0.035**	**4.216×10^{-3}**
COL4A2	29	0.093	0.188	NA	NA	NA	NA
COL4A1	13	0.023	0.479	NA	NA	NA	NA

**Significance threshold is p-value<0.017 correcting multiple tests for three genes*

Summary and Significance

- We used **a novel composite phenotype to study extreme distributions of cerebral small vessel disease** in older community persons
- In ~500 older community participants, **two (0.4%) carried known mutations described in Mendelian SVD**
  - a heterozygote *HTRA1* genotype at CARASIL causing mutation
  - a heterozygote *NOTCH3* cysteine modifying variant, typical of CADASIL
- Report **novel association of common *HTRA1*** variant with extreme-SVD
  - ongoing replication and functional exploration
- Significant association of **burden in *NOTCH3* protein-modifying rare and low frequency variants** with extreme-SVD
- Overall, these findings suggest some shared mechanisms and continuum between Mendelian and multifactorial cerebral SVD. Motivation for large multi-cohort study for SVD gene-mapping.

Acknowledgments

Bordeaux Population Health Center - Inserm U1219

Vascular and Neurological Diseases: Integrative and Genetic Epidemiology **(VINTAGE)**

Stéphanie Debette, *MD, PhD* (director)
Carole Dufouil, *PhD* (co-director)

Christophe Tzourio, *MD, PhD, director Inserm U1219*
Geneviève Chêne, *MD, PhD, director Public Health dept.*
Guillaume Albaret, *MD*, *PhD candidate*
Vincent Bouteloup, *Statistician*
Marie-Gabrielle Duperron, *MD*, *PhD candidate*
Clelia Favary, *Master student*
Leslie Ferreira, *PhD candidate, visiting researcher*
Eric Frison, *MD, PhD candidate*
Chloé Galmishe, *MD*, *PhD candidate*
Allard Hauer, *MD*, *PhD candidate*
Morgane Lachaize-Gaboreau, *Senior Clin. Res. Associat*
Aniket Mishra, *PhD post-doctoral fellow*
Shanti Neff-Baro, *Master student*
Murali Sargurupremraj, *PhD post-doctoral fellow*
Sabrina Schilling, *PhD research engineer*
Aicha Soumare, *PhD research engineer*
Judith Thomas-Crusells, *PhD*, *Project Coordinator*
Marie-Helene Violleau, *Master student*

erc
JPND research
HORIZON 2020
Fondation Leducq
INVESTISSEMENTS D'AVENIR
Agence Nationale de la Recherche ANR


Myriam Fornage
Joshua Bis
Stephanie Debette
Cornelia van Duijn
Will Longstreth
Neuro-CHARGE working group
Sudha Seshadri
Arfan Ikram
Ganesh Chauhan

Acknowledgments

CHARGE consortium

COHORTS FOR HEART AND AGING RESEARCH
IN GENOMIC EPIDEMIOLOGY