



**HAL**  
open science

# Impact des variants génétiques à risque de la maladie des petites artères cérébrales sur la microstructure de la substance blanche et les fonctions exécutives chez le sujet jeune. À partir des données de l'étude i-Share

Constance Bordes

## ► To cite this version:

Constance Bordes. Impact des variants génétiques à risque de la maladie des petites artères cérébrales sur la microstructure de la substance blanche et les fonctions exécutives chez le sujet jeune. À partir des données de l'étude i-Share. Santé publique et épidémiologie. 2019. dumas-02435123

**HAL Id: dumas-02435123**

**<https://dumas.ccsd.cnrs.fr/dumas-02435123>**

Submitted on 10 Jan 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

2<sup>ème</sup> année Master en Sciences, Technologies, Santé  
Mention Santé publique - Parcours épidémiologie  
ISPED – Université de Bordeaux

Promotion 2018 – 2019

---

**Impact des variants génétiques à risque de la maladie des  
petites artères cérébrales sur la microstructure de la  
substance blanche et les fonctions exécutives  
chez le sujet jeune**

*À partir des données de l'étude i-Share*

---

*Stage réalisé du 04/02/2019 au 21/06/2019*

**Équipe VINTAGE – Épidémiologie intégrative et génétique du vieillissement  
cérébral et vasculaire.**

**INSERM – U1219 – « Bordeaux Population Health »**

146 rue Léo Saignat  
33076 Bordeaux cedex

**Soutenance le 24/06/2019**

**Par Constance BORDES**

Née le 6 mai 1994

Sous la responsabilité de Stéphanie Debette, PUPH Epidémiologie, Docteur en médecine,  
Neurologue au CMRR de Bordeaux, directrice de l'équipe VINTAGE, INSERM U1219

Adresse postale :

Université de Bordeaux

Institut de Santé Publique d'Epidémiologie et de Développement

146 rue Léo Saignat

CS 61292

33076 Bordeaux cedex

**FICHE D'AUTORISATION D'UTILISATION DU MEMOIRE  
à joindre au mémoire de stage dûment complété et signé**

**Diplôme :**

Année universitaire 2018/2019

**Titre du mémoire :** Impact des variants à risques de la maladie des petites artères cérébrales sur la microstructure de la substance blanche et les fonctions exécutives chez le sujet jeune (données issues de la cohorte i-share)

<p><b>Le maître de stage (NOM, Prénom, Fonction, Organisme)</b> DEBETTE, Stéphanie, Directrice de Stage, INSERM U1219</p>	
<p><b>J'autorise l'utilisation de ce travail universitaire à des fins pédagogiques (mise à disposition du pdf du mémoire à l'ISPED)</b></p> <p><input checked="" type="checkbox"/> oui <input type="checkbox"/> non</p>	<p><b>En cas de sélection par le jury de soutenance de ce travail universitaire, j'en autorise la diffusion en ligne</b> <i>Le travail universitaire sera signalé par la direction de la documentation dans le catalogue Babord+ ainsi que sur DUMAS, la plateforme nationale d'archivage des mémoires en ligne (hébergée par HAL-CCSD), consultable librement sur Internet et indexée par les moteurs de recherche.</i></p> <p><input type="checkbox"/> oui immédiatement <input checked="" type="checkbox"/> oui à compter du : 01/07/2019 <input type="checkbox"/> non</p>
<p><b>Date</b> 10/06/2019</p>	<p><b>Signature :</b> Stephanie Debette</p> <p><small>Signature numérique de Stéphanie Debette DN : cn=Stéphanie Debette, o=University of Bordeaux, ou=Bordeaux Population Health (BPH) Center-Inserm U1219, email=Stephanie.debette@u-bordeaux.fr, c=FR Date : 2019.06.10 11:38:19 +02'00'</small></p>

<p><b>L'étudiant(e) (NOM, Prénom)</b> BORDES, Constance</p>	
<p><b>J'autorise l'utilisation de ce travail universitaire à des fins pédagogiques (mise à disposition du pdf du mémoire à l'ISPED)</b></p> <p><input checked="" type="checkbox"/> oui <input type="checkbox"/> non</p>	<p><b>En cas de sélection par le jury de soutenance de ce travail universitaire, j'en autorise la diffusion en ligne</b> <i>Le travail universitaire sera signalé par la direction de la documentation dans le catalogue Babord+ ainsi que sur DUMAS, la plateforme nationale d'archivage des mémoires en ligne (hébergée par HAL-CCSD), consultable librement sur Internet et indexée par les moteurs de recherche.</i></p> <p><input type="checkbox"/> oui immédiatement <input checked="" type="checkbox"/> oui à compter du : 01/07/2019 <input type="checkbox"/> non</p>
<p><b>Date</b> 10/06/2019</p>	<p><b>Signature :</b> Bordes</p> <p><small>Signature numérique de Bordes Date : 2019.06.10 11:38:19 +02'00'</small></p>

## REMERCIEMENT

Je tiens à remercier M. Christophe Tzourio, Directeur du centre de recherche Bordeaux Population Health, Inserm U1219, à l'Université de Bordeaux et investigateur principal de la cohorte i-Share, pour m'avoir accueillie au sein de cette structure afin de réaliser mon stage de Master 2.

Je tiens à remercier tout particulièrement Stéphanie Debette, ma directrice de stage pour m'avoir accordé une seconde fois sa confiance, supervisée mon stage, mais aussi pour le partage de ses connaissances, de son savoir-faire, de sa gentillesse et de son temps, tout au long de ces deux années de master.

Je remercie également toute l'équipe VINATGE, et tout particulièrement Morgane Lachaize, Aïcha Soumaré, Aniket Mishra et Murali Sargurupremraj pour leur disponibilité lors de mes diverses sollicitations.

Je remercie Pr. Bernard Mazoyer (Investigateur Principal MRi-Share), Pr. Gregory Beaudet et Pr. Laure Zago ainsi que toute l'équipe du Groupe d'imagerie Neurofonctionnelle (GIN), supervisant le projet MRi-Share, pour m'avoir fourni les données d'imagerie cérébrale.

Je remercie l'équipe de l'étude i-Share pour m'avoir fourni les données des participants à i-Share.

Merci aux responsables pédagogiques et aux enseignants de ce Master, qui ont su prodiguer un enseignement de qualité.

Merci à mes « co-masters » pour ces deux agréables années de masters et tout particulièrement Quentin, Clara et Anaïs pour nos longs échanges et notre entraide lors de ces pauses déjeuners.

# TABLE DES MATIERES

<b>REMERCIEMENT</b> .....	<b>1</b>
<b>LISTE DES TABLEAUX</b> .....	<b>5</b>
<b>LISTE DES FIGURES</b> .....	<b>5</b>
<b>LISTE DES ABREVIATIONS</b> .....	<b>6</b>
<b>INTRODUCTION</b> .....	<b>7</b>
I.    DEFINITION DES TERMES .....	8
1. <i>Rappel anatomique</i> .....	8
2. <i>IRM de tenseur de diffusion (DTI)</i> .....	9
3. <i>Variants génétiques communs ou polymorphismes génétiques</i> .....	11
II.    LA MATURATION CEREBRALE .....	12
III.    LA MALADIE DES PETITES ARTERES CEREBRALES (MPAC).....	13
1. <i>Etiologie et caractéristiques cliniques</i> .....	13
2. <i>Les marqueurs IRM de la MPAC</i> .....	15
IV.    HYPOTHESES ET OBJECTIFS.....	16
<b>MATÉRIELS ET MÉTHODES</b> .....	<b>18</b>
I.    SCHEMA ET POPULATION D'ETUDE .....	18
1. <i>Cohorte i-Share</i> .....	18
2. <i>Deux études ancillaires : MRi-Share et Bio-Share</i> .....	18
3. <i>Projet SEGWAY</i> .....	19
4. <i>Population d'étude</i> .....	19
II.    RECUEIL DES DONNEES .....	19
1. <i>Données d'imagerie</i> .....	19
2. <i>Données génétiques</i> .....	20
3. <i>Données cognitives</i> .....	21
4. <i>Autres données</i> .....	22
III.    VARIABLES ETUDIEES .....	22
1. <i>Variables d'intérêts</i> .....	22
2. <i>Expositions d'intérêts</i> .....	24
3. <i>Covariables</i> .....	25
IV.    STRATEGIE D'ANALYSE .....	26
1. <i>Analyse descriptive de la population d'étude</i> .....	26
2. <i>Analyses principales</i> .....	26
3. <i>Analyses complémentaires</i> .....	28

<b>RÉSULTATS</b> .....	<b>29</b>
I.  POPULATION D'ETUDE.....	29
II. DESCRIPTION DE LA POPULATION.....	29
III.  RESULTATS DES ANALYSES D'ASSOCIATION .....	34
1. <i>Association du PSMD avec le wGRSres d'HSB</i> .....	34
2. <i>Association du temps de réaction au test de Stroop avec le wGRSres d'HSB et avec le PSMD</i> .....	35
3. <i>Analyses complémentaires</i> .....	37
<b>DISCUSSION</b> .....	<b>40</b>
I.  RESUME DES RESULTATS.....	40
II. LIEN AVEC LITTERATURE .....	41
III.  FORCES .....	43
IV.  LIMITES ET PERSPECTIVES.....	44
<b>CONCLUSION</b> .....	<b>46</b>
<b>BIBLIOGRAPHIE</b> .....	<b>47</b>
<b>ANNEXES</b> .....	<b>52</b>

## LISTE DES TABLEAUX

Tableau 1. Description de la population d'étude, N = 1 692. Cohorte i-Share, SEGWAY. Bordeaux, France. 2019. ....	31
Tableau 2. Associations du PSMD avec des facteurs sociodémographiques et de risques vasculaires de la MPAC. Cohorte i-Share, SEGWAY. Bordeaux, France. 2019. ....	33
Tableau 3. Association entre le wGRSres d'HSB et le PSMD, N = 1 692. Cohorte i-Share. SEGWAY. Bordeaux, France. 2019.....	34
Tableau 4. Association entre le wGRSres d'HSB, le PSMD et le test de Stroop, N = 1283. Cohorte i-Share. SEGWAY. Bordeaux, France. 2019.....	36
Tableau 5. Association du wGRSres d'HSB et PSMD (N = 854) ainsi que du wGRSres d'HSB et PSMD sur le temps de réaction du test de Stroop (N =741) sans et avec un ajustement sur les facteurs de risques vasculaires. Cohorte i-Share. SEGWAY. Bordeaux, France. 2019. ...	37
Tableau 6. Description des caractéristiques des sujets sans découvertes fortuites (n=1656) et des sujets avec découvertes fortuites (n=36). Cohorte i-Share. SEGWAY. Bordeaux, France. 2019. ....	38
Tableau 7. Association du wGRSres d'HSB et PSMD (N = 1 652) ainsi que du PSMD et du test de Stroop (N = 1 252) des participants sans DF. Cohorte i-Share. SEGWAY. Bordeaux, France. 2019. ....	39

## LISTE DES FIGURES

Figure 1. Représentation schématique descriptif d'un neurone.....	8
Figure 2 : Répartition de la substance blanche et de la substance grise au sein du système nerveux central (9). ....	9
Figure 3. Représentation schématique du tenseur [B : ellipsoïde défini par ses vecteurs ( $\epsilon_1, \epsilon_2, \epsilon_3$ ) et valeurs propres ( $\lambda_1, \lambda_2, \lambda_3$ )] qui modélise la diffusion anisotropique des molécules d'eau au sein des fibres alignées (A). (Guye M et al. Epilepsies. 2008) .....	11
Figure 4. Principaux marqueurs IRM de lésions cérébrales vasculaires silencieuses, principalement pour une maladie des petits vaisseaux. ....	14
Figure 5. Marqueurs IRM de la MPAC : phénotypes intermédiaires d'AVC, de démence et de déclin cognitif. ....	16
Figure 6. Procédure de la quantification du marqueur PSMD. (Baykara et al. Ann Neurol. 2016).....	23
Figure 7. Diagramme de flux de la population d'étude. Cohorte i-Share, SEGWAY. Bordeaux. 2019. ....	30

## LISTE DES ABREVIATIONS

ADN : Acide Désoxyribonucléique

AVC : Accident Vasculaire Cérébral

DTI : Imagerie par Tenseur de Diffusion (*Diffusion Tensor Imaging*)

FA : Anisotropie Fractionnelle (*fractional anisotropy*)

FDR : Facteurs De Risques

GRS : Score de Risque Génétique (*genetic risk score*)

GWAS : étude d'association pangénomique (*genome-wide association study*)

HSB : Hypersignaux de la substance blanche

IMC : Indice de Masse Corporelle

IP : Investigateur Principal

IRM : Imagerie par Résonance Magnétique

i-Share : internet-based Students' Health Research Enterprise

LD : déséquilibre de liaison (*linkage disequilibrium*)

MAF : fréquence de l'allèle mineur (*minor allele frequency*)

MD : diffusivité moyenne (*mean diffusivity*)

MPAC : Maladie des petites artères cérébrales

PSMD : peak width of skeletonized mean diffusivity

SB : substance blanche

SEGWAY : Study of Environmental and Genome Wide predictors of early structural brain Alterations in Young students

SNC : Système nerveux central

SNP : *Single Nucleotide Polymorphisms* ou Polymorphisme Mononucléotidique



## INTRODUCTION

L'augmentation de l'espérance de vie et du vieillissement de la population, observé mondialement, ont conduit à une augmentation des troubles cognitifs et des démences, pathologies sans traitement curatif existant à ce jour, entraînant à une augmentation importante au recours aux soins de santé (1). Selon l'Organisation Mondiale de la Santé, le nombre total de personnes ayant une démence devrait atteindre 82 millions en 2030 et 152 millions d'ici 2050 (2). La démence et le déclin cognitif précoce représentent donc un poids sociétal majeur en santé publique (3). Ralentir le déclin cognitif et retarder l'apparition des démences de seulement quelques années pourraient avoir un impact majeur sur sa prévalence au niveau populationnel. Cependant, les mécanismes sous-jacents et les facteurs de risques de ces maladies neurologiques du sujet âgé sont mal connus. De nombreuses données suggèrent que les processus physiopathologiques conduisant à la démence débutent de nombreuses années, voire décennies, avant la survenue de symptômes cliniques (4,5). Mieux comprendre ces déterminants précoces est essentiel d'une part pour améliorer notre compréhension des mécanismes physiopathologiques de la démence et d'autre part pour permettre d'identifier précocement les individus à haut risque dans la population, afin de les cibler préférentiellement pour des stratégies préventives. Dans ce cadre, nous nous intéressons aux déterminants des modifications cérébrales structurelles précoces qui pourraient être associées plus tardivement à un risque accru de démence, soit par un continuum tout au long de la vie des processus pathologiques qui pourraient déjà avoir un impact très précoce, soit parce que certains facteurs influençant la **maturation cérébrale** à un jeune âge pourraient être associés avec des processus de sénescence et/ou neurodégénératifs chez la personne âgée (4–6). De nombreuses études ont montré que certaines modifications structurelles cérébrales, telles qu'une atrophie cérébrale globale ou focalisée, notamment au niveau des hippocampes, ou des marqueurs structurels de maladie des petites artères cérébrales (MPAC), notamment les hypersignaux de la substance blanche (HSB), étaient associées fortement à un risque accru de démence en population générale (7,8). Ces modifications structurelles cérébrales sont multifactorielles et des variants génétiques communs ont été identifiés comme étant associés à un volume hippocampique moindre et une charge plus importante d'HSB chez la personne âgée. Afin de mieux appréhender les mécanismes très précoces de ces modifications structurelles cérébrales, nous nous intéressons à l'impact de ces variants génétiques sur des modifications structurelles cérébrales précoces chez l'adulte jeune en **imagerie par résonance magnétique (IRM) en tenseur de diffusion (DTI)** et l'association de ces dernières avec les performances cognitives. Nous nous intéressons ici spécifiquement à l'association de déterminants génétiques de la charge en HSB avec un marqueur précoce d'intégrité de la **microstructure de la substance**

**blanche** cérébrale en IRM-DTI, le PSMD (peak-width of skeletonized mean diffusivity), et à l'impact de ces modifications structurelles précoces sur les performances cognitives dans la cohorte d'étudiants i-Share.

## I. Définition des termes

Après cette brève introduction, il me semble nécessaire de continuer par quelques explications et rappels importants pour la compréhension de ce mémoire. Nous commençons par quelques rappels anatomiques du cerveau, suivit par une brève explication d'IRM-DTI et cette partie se termine par des notions génétiques importantes pour ce mémoire.

### 1. Rappel anatomique

Le système nerveux correspond à l'ensemble des organes qui commande l'organisme et qui doit assurer la coordination, la régulation et le contrôle des viscères à l'intérieur de l'organisme et vis-à-vis du milieu extérieur. Le système nerveux est subdivisé plusieurs systèmes dont le système nerveux central (SNC) (9). Le SNC est le centre de l'élaboration, de l'intégration et de la coordination des influx nerveux. Il est aussi appelé névraxe et est composé de deux éléments.

- l'encéphale, constitué du cerveau, du tronc cérébral et du cervelet et qui est situé dans la boîte crânienne,
- la moelle spinale ou épinière, qui est contenue dans le canal rachidien.

La cellule de base du tissu nerveux, qui compose le SNC, est le neurone (**Figure 1**). Ce dernier est constitué par un corps cellulaire d'où partent des prolongements de deux types, les dendrites et l'axone, qui permettent la communication entre les neurones et la propagation de l'influx nerveux. Les dendrites, de courte taille, conduisent l'influx nerveux vers le corps cellulaire, alors que l'axone, toujours unique, parfois très long, conduit l'influx nerveux à partir du corps cellulaire jusqu'à ses cibles.

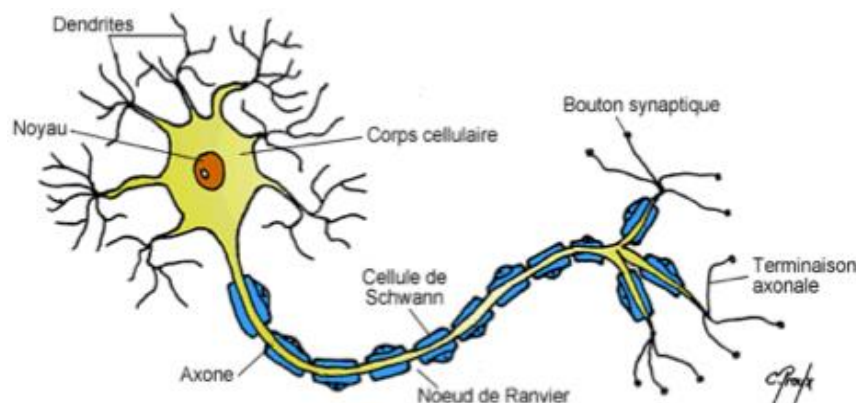


Figure 1. Représentation schématique descriptif d'un neurone.

Source : <https://sciences-cognitives.fr>

Le SNC est composé de deux structures d'aspect différent (**Figure 2**) :

- la substance grise qui est une agglomération des corps cellulaires des neurones et responsable du traitement des informations.
- la substance blanche qui est composée de l'ensemble des axones, qui relient les corps cellulaires.

La substance blanche constitue les voies de conduction nerveuse et est responsable de la propagation des informations dans le système nerveux. Les axones remplissent la majorité du volume du cerveau et sont regroupés en groupe de fibres (ou faisceaux) nerveuses. La plupart des axones sont entourés par une épaisse gaine composée d'une substance grasseuse, la myéline qui donne cet aspect blanc, qui les isole électriquement et augmente considérablement la vitesse de propagation de l'influx nerveux (9). C'est au niveau de la microstructure de la substance blanche que l'on peut observer l'architecture axonale, l'étendue de la myélinisation des axones ou bien la densité des fibres axonales. Des changements au sein de la microstructure peuvent amener à des changements macrostructuraux de la substance blanche.

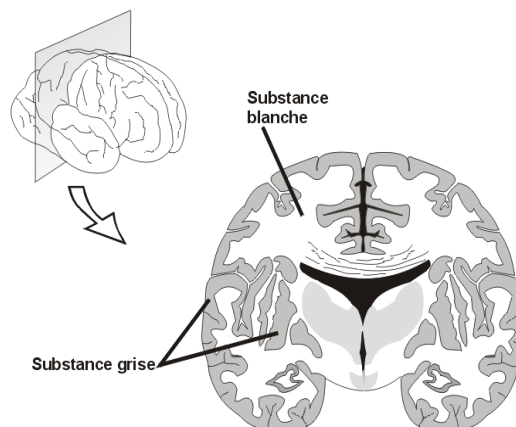


Figure 2 : Répartition de la substance blanche et de la substance grise au sein du système nerveux central (9).

## 2. IRM de tenseur de diffusion (IRM-DTI)

L'IRM-DTI est une technique d'IRM basée sur la diffusion des molécules d'eau au sein d'un milieu. Cette technique permet de quantifier les altérations de la microstructure des tissus, qui peuvent être invisibles sur les IRM conventionnelles, de façon non invasive et *in vivo* (10).

### a) *La diffusion moléculaire : mouvement brownien*

La diffusion des molécules d'eau fait référence à un mouvement aléatoire des molécules, aussi appelée le mouvement Brownien, qui résulte d'une énergie thermique portée par ces molécules. Ces mouvement obéissent à une loi statistique établie en 1905 par Einstein (11) : dans un milieu libre, pendant un intervalle de temps donné, le déplacement des molécules obéit à une distribution gaussienne tridimensionnelle. Les molécules se déplacent

aléatoirement dans un espace sur une distance qui est statistiquement décrite par le coefficient de diffusion qui ne dépend que de la taille (masse) des molécules, de la température, et de la viscosité du milieu (10). Le coefficient de diffusion correspond à la vitesse de répartition des particules par unité de temps et de surface et s'exprime en m<sup>2</sup>/s (12,13).

#### *b) Diffusion Anisotropie*

Dans un milieu sans restriction, la diffusion des molécules d'eau est qualifiée d'isotropique, c'est-à-dire identique dans toutes les directions. Dans le cerveau qui est un tissu hétérogène, composé de multiples structures, la diffusion de l'eau n'est pas identique dans toutes les directions, on parle alors de diffusion anisotropique (12,13). La substance blanche est donc qualifiée de milieu anisotropique du fait de l'organisation en faisceaux des fibres parallèles plus ou moins myélinisées constituées par les axones. Cette organisation en faisceaux représente un obstacle majeur pour la diffusion des molécules d'eau, avec une diffusion facilitée parallèlement aux faisceaux et bien plus difficile perpendiculairement.

#### *c) Tenseur de Diffusion*

Du fait de cette anisotropie, la diffusion ne peut plus être caractérisée par un seul coefficient de diffusion, mais requiert une description tensorielle. Le tenseur de diffusion  $D$  est une matrice caractérisant la mobilité moléculaire dans chaque direction ( $D_{xx}$ ,  $D_{yy}$ ,  $D_{zz}$ ) (12,14,15). Une propriété importante du tenseur est qu'il est diagonalisé. Cette diagonalisation du tenseur est nécessaire afin de déterminer les directions de diffusion :

$$\hat{D} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}$$

Ce tenseur de diffusion est représenté sous forme d'un ellipsoïde (**Figure 3**), défini par trois valeurs propres ( $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$ ), placées sur la diagonale, correspondant au coefficient de diffusion le long de ces trois axes ; et par trois vecteurs propres ( $\varepsilon_1$ ,  $\varepsilon_2$ ,  $\varepsilon_3$ ) représentant les axes de l'ellipsoïde et permettant de déterminer la direction principale de la diffusion (15,16). L'axe le plus long de l'ellipsoïde représente la direction principale de diffusion alors que le plus petit représente celui de moindre diffusion. Si la diffusion est isotropique, on obtient une sphère. Au contraire, si la diffusion est anisotropique, comme c'est le cas dans la matière blanche, on obtient une forme ellipsoïde (13–15).

Les images du DTI peuvent être analysées à l'aide de deux principaux paramètres : l'anisotropie fractionnelle (fractional anisotropy [FA]) et la diffusivité moyenne (mean

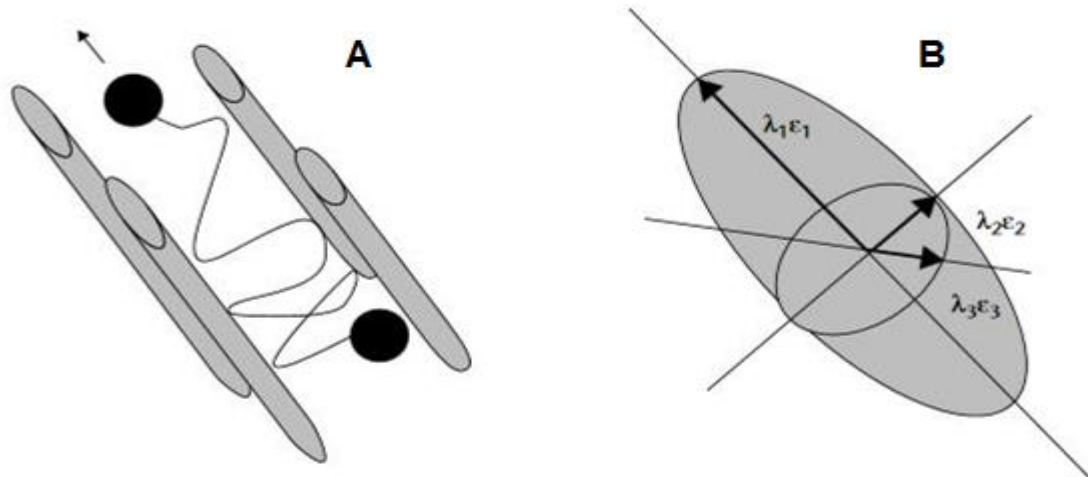


Figure 3. Représentation schématique du tenseur [B : ellipsoïde défini par ses vecteurs ( $\epsilon_1, \epsilon_2, \epsilon_3$ ) et valeurs propres ( $\lambda_1, \lambda_2, \lambda_3$ )] qui modélise la diffusion anisotrope des molécules d'eau au sein des fibres alignées (A). (Guye M et al. *Epilepsies*. 2008)

diffusivity [MD]. La FA est une mesure qui permet de déterminer le degré d'anisotropie traduisant la présence de structures orientées et leur degré de cohérence, c'est-à-dire à quel point la forme du tenseur de diffusion va dévier de la forme sphérique. Si les fibres nerveuses sont fortement orientées dans une direction alors la FA s'approche de 1 (milieu anisotrope). La FA est un index pour mesurer l'intégrité de la substance blanche et est sensible aux changements de myélinisation cérébrale liés à l'âge (6,13). La MD caractérise la mobilité moléculaire globale dans le tissu et mesure donc le déplacement global moyen des molécules et la présence d'obstacles à la diffusion (15). Si les fibres de substance blanche sont altérées, la diffusivité moyenne va alors augmenter. Un nouveau marqueur, issu d'une squelettisation des images du DTI et d'une analyse par histogramme, a récemment été développé. Il s'agit du PSMD. Ce dernier semble être un marqueur plus robuste et sensible que les autres paramètres du DTI aux changements de la microstructure de la substance blanche et ce, même dans une population âgée et malade (17).

### 3. Variants génétiques communs ou polymorphismes génétiques

Depuis quelques années, l'un des apports de la génétique dans la compréhension physiopathologique et mécanistique des maladies est l'identification de facteurs de risque génétique. Par l'intermédiaire d'études dites pangénomique (GWAS : *genome-wide association study*), les variants génétiques les plus fréquemment associés à une maladie ou à un phénotype donné sont déterminés en comparant la fréquence de centaines de milliers de ces derniers (distribués sur l'ensemble des chromosomes) entre un groupe de témoins et un groupe de cas atteints de la maladie ou porteurs du phénotype en question. Dans ce contexte, les variants génétiques les plus souvent utilisés sont les polymorphismes mononucléotidiques (*Single Nucleotide Polymorphisms*, SNP) qui sont définis comme une variation de nucléotides

à une position donnée dans la séquence d'ADN qui se produisent à une fréquence élevée (>1%) dans le génome humain (18). En général, ils ont individuellement un impact faible sur les systèmes biologiques, néanmoins ils peuvent avoir des conséquences fonctionnelles, rarement en entraînant des modifications des acides aminés et de la structure de la protéine encodée, et plus fréquemment en modulant le niveau d'expression des gènes. Bien que l'impact individuel des SNPs soit faible, collectivement, lorsqu'ils sont regroupés sous forme de score de risque génétique (GRS), leur impact peut être important. La construction d'un score de risque génétique permet de caractériser le génotype à risque de plusieurs SNPs. Ainsi, il a récemment été montré que des scores de risque polygéniques élevés de certaines maladies complexes telle que la coronaropathie sont associés, dans la population, à une augmentation du risque de la maladie du même ordre de grandeur que des mutations causales dans des maladies génétiques rares à transmission mendéliennes (19).

La plupart des SNPs présentent deux allèles, c'est-à-dire qu'il existe deux possibilités de variations nucléotidiques à cet emplacement, au sein d'une population. Ainsi, les personnes ayant l'allèle dit « à risque » du variant (allèle associé à la maladie ou au phénotype donné) seront significativement plus nombreuses à présenter ce phénotype que les personnes ayant l'autre allèle du variant génétique. Une propriété importante liée aux SNPs est le déséquilibre de liaison (LD : *linkage disequilibrium*) qui décrit le degré de transmission de l'allèle d'un SNP ou sa corrélation avec celui d'un autre SNP au sein d'une population. Il permet de décrire les changements des variations génétiques au sein d'une population au fil du temps. Le LD est généralement représenté par la mesure de corrélation  $r^2$  : un  $r^2$  élevé indique que l'allèle d'un SNP1 est souvent hérité avec celui d'un SNP2 ; ceci implique que les deux SNPs en question apportent des informations similaires et qu'il suffit donc de génotyper l'un des deux pour capturer la variation allélique. Le LD, à ce niveau, peut donc être exploité pour éviter les informations redondantes et ainsi, optimiser les études génétiques (18).

## II. La maturation cérébrale

Depuis la naissance, le cerveau est façonné par des processus qui permettent d'aboutir à une structure complexe et qu'on appelle la maturation cérébrale. Cette dernière regroupe l'ensemble des processus qui participent au développement des cellules cérébrales avec une interaction entre des facteurs génétiques et des facteurs environnementaux par notamment le mécanisme de myélinisation du cerveau humain (qui se poursuit au début de la vie adulte) (13,20). Plus tard dans la vie vient la sénescence, c'est-à-dire le vieillissement cérébral, comprenant des modifications de la substance grise et blanche, notamment avec une perte neuronale ou une démyélinisation des axones. Du fait que les faisceaux de substance blanche relient les aires de substance grise permettant ainsi l'échange d'informations entre les régions du cerveau, l'intégrité de ces substances est donc primordial pour garantir une bonne

communication entre les régions (13). Des récentes études de neuroimagerie ont montré que les taux de maturation cérébrale liés à l'âge et le déclin des faisceaux de la substance blanche sont significativement corrélés. En IRM-DTI, les valeurs de la FA atteignent leur maximum vers l'âge de 32 ans en moyenne (entre 23 et 39 ans) (6), la période entre 20 et 30 ans, qui correspond à l'âge de la population que nous étudions, est donc une période charnière pour la maturation cérébrale. Néanmoins, d'un point de vue épidémiologique et physiopathologique, la signification sur le long terme de modifications subtiles de la structure cérébrale chez l'adulte jeune est mal connue.

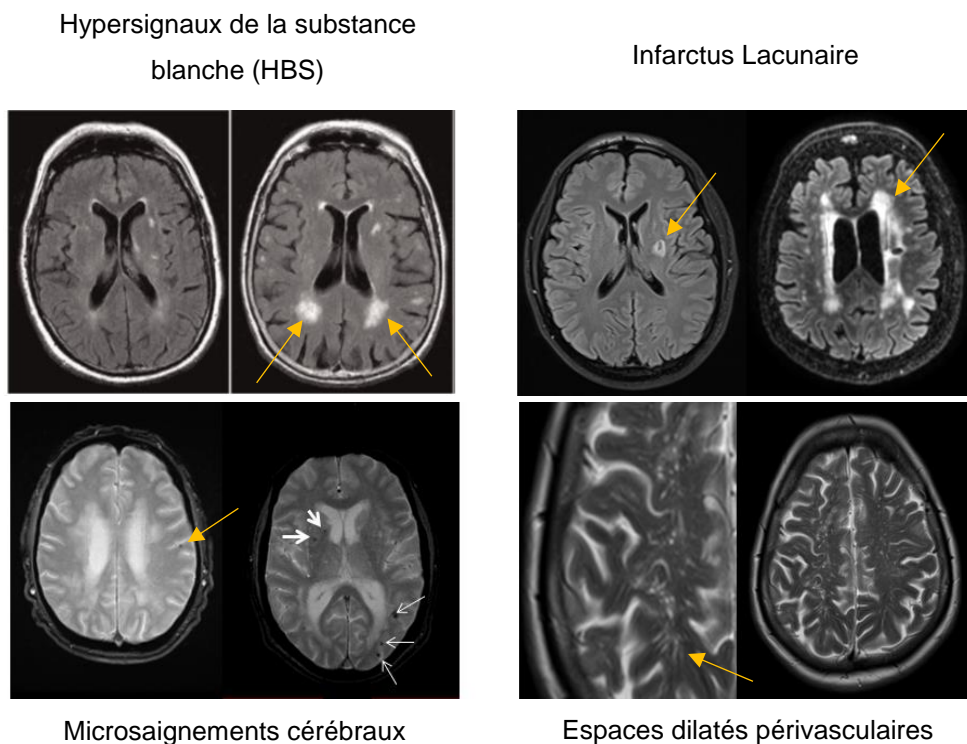
Des nombreuses études indiquent que des événements et des expositions en début de vie constituent des facteurs déterminants dans le développement de maladies neurologiques apparaissant plus tardivement dans la vie. Beaucoup d'arguments suggèrent qu'un contexte de prédisposition multifactorielle (facteurs génétiques et environnementaux) pourrait influencer d'une part la maturation cérébrale et d'autre part des stades précoces de processus neurodégénératifs et de vieillissement cérébrovasculaires (4–6). En effet, les facteurs qui modulent la maturation cérébrale pourraient avoir une influence dans les processus de la neurodégénérescence précoce. D'autres études ont montré que les processus pathologiques conduisant à la démence, commencent longtemps, probablement des décennies, avant le diagnostic clinique (21,22). En outre, le vieillissement vasculaire cérébral et notamment la maladie des petites artères cérébrales, commencerait bien avant que les symptômes cliniques ne deviennent apparents (23,24). Acquérir une connaissance approfondie des altérations structurelles cérébrales et de leurs déterminants chez des adultes jeunes permettrait de mieux comprendre les mécanismes de ces modifications ainsi que l'évolution des voies biologiques impliquées dans le vieillissement cérébral structurel et cognitif tout au long de la vie.

### **III. La maladie des petites artères cérébrales (MPAC)**

#### **1. Etiologie et caractéristiques cliniques**

La maladie de petites artères cérébrales (MPAC) est une maladie affectant les vaisseaux cérébraux associée à des altérations dans la microstructure de la substance blanche du cerveau. La MPAC regroupe des processus pathologiques qui affectent les petits vaisseaux du cerveau, incluant les petites artères et artérioles, en altérant leur structure ou leur fonction (25). Cependant, les mécanismes sur lesquels reposent ces lésions sont peu connus (26). Une des hypothèses pathogènes serait qu'un rétrécissement de la lumière des vaisseaux mènerait à un état d'hypoperfusion chronique de la substance blanche, résultant finalement d'une dégénération des fibres myélinisées. D'autres mécanismes tels que les lésions de la barrière hémato-encéphalique ou l'apoptose des oligodendrocytes pourraient être

impliqués (23). Ces mécanismes divers induiraient des lésions parenchymateuses visibles en IRM, telles que des hypersignaux de la substance blanche (HSB), des infarctus lacunaires, des micro-saignements cérébraux ou des espaces dilatés péri vasculaires (espaces de Virchow-Robin) (17,27–29) (**Figure 4**). Ces lésions conduiraient à une interruption partielle des connexions cortico-sous-corticales et cortico-corticales qui pourraient contribuer aux troubles cognitifs. En outre, les facteurs génétiques semblent aussi jouer un rôle significatif avec une héritabilité, c'est-à-dire la variance expliquée par des facteurs génétiques, allant de 55 % à 80 % (23).



*Figure 4. Principaux marqueurs IRM de lésions cérébrales vasculaires silencieuses, principalement pour une maladie des petits vaisseaux.*

La MPAC constitue une des principales causes d'accidents vasculaires cérébraux (AVC), ischémique et hémorragique, qui représentent, au niveau mondial, la première cause d'handicap neurologique acquis chez l'adulte, la deuxième cause de mortalité et de démence (1). En France, on dénombre plus de 130 000 AVC chaque année (30). Le plus souvent toutefois la MPAC est dite « silencieuse » ou « occulte », sans AVC associé. Elle est extrêmement fréquente dans la population générale âgée : sur la base d'études de neuroimagerie en population générale (31,32), on estime que plus de 4 millions de personnes âgées de 65 ans ou plus en France présentent une MPAC modérée à sévère. La MPAC est aussi en lien avec l'apparition insidieuse, lentement progressive, de troubles cognitifs (25,28). Il s'agit de la principale contribution vasculaire au déclin cognitif et à la démence dans la population (7,17,27,28). En effet, il est maintenant avéré que l'immense majorité des cas de



déclin cognitif et de démence dans la population ont une origine mixte, vasculaire et neurodégénérative (33). Le déclin cognitif peut être défini comme une baisse progressive des fonctions cognitives au cours du temps. C'est un processus qui s'observe classiquement dans le vieillissement, mais qui peut revêtir un caractère pathologique lorsqu'il conduit à une altération précoce des fonctions cognitives. Le déclin cognitif lié à la MPAC est typiquement caractérisé en premier lieu par une déficience importante de la vitesse de traitement des informations et de la fonction exécutive (34,35). La vitesse de traitement des informations est une capacité cognitive qui pourrait être définie comme le temps nécessaire à une personne pour accomplir une tâche mentale. Elle est liée à la rapidité avec laquelle une personne peut comprendre et réagir aux informations qu'elle reçoit, qu'elle soit visuelle (lettres et chiffres), auditive (langage) ou en mouvement. En d'autres termes, la vitesse de traitement est le temps écoulé entre la réception et la réponse à un stimulus. Les fonctions exécutives sont des fonctions cognitives qui sont en lien avec les processus attentionnels et la mémoire de travail et impliquent un ensemble de mécanismes élémentaires (sélection, inhibition, activation) qui gouvernent la sélection et l'exécution des processus cognitifs (mémoire, langage) (36).

## 2. Les marqueurs IRM de la MPAC

Les principaux marqueurs en IRM cérébrale de la MPAC sont les HSB, les infarctus lacunaires, les microsaignements et les espaces dilatés périvasculaires. Des études en population générale ont montré que ces marqueurs « silencieux » de la MPAC en IRM cérébrale sont très fréquents (7,31,37). Les HSB sont de loin les plus fréquents, détectables à des degrés divers chez la quasi-totalité des personnes âgées de 65 ou plus. Ils désignent des anomalies de la substance blanche, apparaissant comme des zones d'hyperintensité sur certaines séquences IRM. Des études ont montré que les HSB sont fortement associés au déclin cognitif, au risque de démence, d'AVC et de décès. En outre, une corrélation a été mise en évidence entre la progression des HSB et le déclin dans les performances cognitives (25,38). Toutefois, bien que détectables assez fréquemment également chez des sujets plus jeunes, notamment des adultes d'âge moyen, les HSB sont plus rares chez l'adulte très jeune chez qui ils peuvent refléter d'autres processus pathologiques que la MPAC, notamment des maladies inflammatoires telles que la sclérose en plaque.

Par rapport aux marqueurs de MPAC suscités, détectés sur des séquences morphologiques classiques en IRM (T1, T2, T2\*), IRM-DTI est une technique plus sensible aux altérations subtiles de la structure cérébrale au niveau de la substance blanche (35,39,40). Les marqueurs du DTI, que sont la FA et la MD, sont plus fortement corrélés avec la cognition que le volume d'HSB chez la personne âgée, avec une fonction exécutive et une vitesse de traitement de l'information moins performantes (34,35,41,42). Par ailleurs, ils sont mesurables très précocement. Un nouveau marqueur du DTI intéressant a été récemment développé, le

PSMD (peak width of skeletonized mean diffusivity) (17). Des preuves récentes indiquent que le PSMD serait un marqueur avec une plus grande sensibilité que les mesures classiques de FA et MD et que les HSB pour évaluer les lésions causées par la MPAC. Il serait aussi plus fortement lié au déclin cognitif chez la personne âgée (17). Des données récentes non publiées, générées dans le cadre de l'étude i-Share et du consortium JPND-BRIDGET par l'équipe de neuroimagerie du Professeur Bernard Mazoyer, suggère que le PSMD augmente significativement avec l'âge et ce dès 20 ans, ce qui n'est pas le cas des autres marqueurs en DTI dont la courbe de progression s'inverse vers l'âge de 35 ans (Beaudet et al., en préparation). Le marqueur PSMD semble donc intéressant pour explorer l'impact très précoce de variants génétiques à risque de MPAC de la personne âgée (**Figure 5**).

Etant mesurables de façon non invasive et quantitative dans de grands échantillons, le PSMD constitue un phénotype intermédiaire important et intéressant pour mettre en évidence des mécanismes biologiques et des voies physiopathologiques sous-jacents de la MPAC.

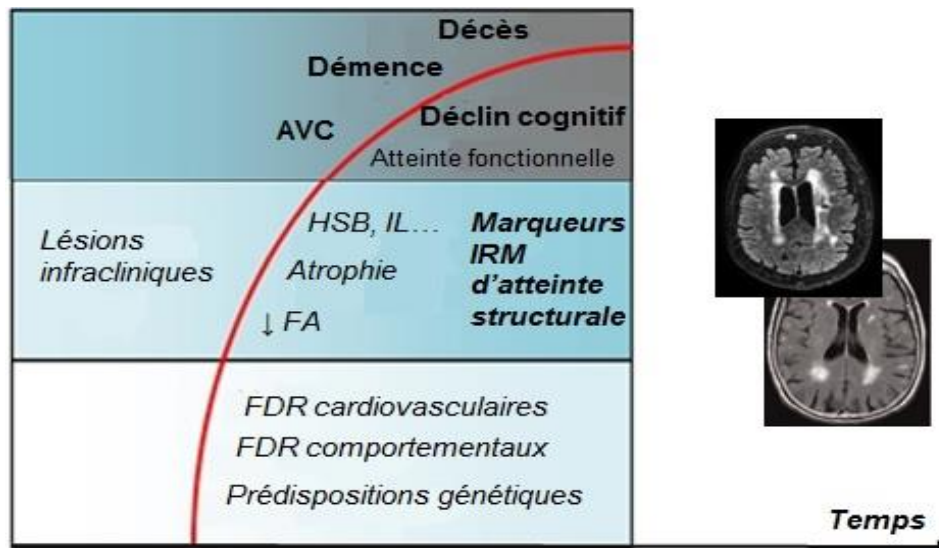


Figure 5. Marqueurs IRM de la MPAC : phénotypes intermédiaires d'AVC, de démence et de déclin cognitif.

(HSB : Hypersignaux de la substance blanche, IL : Infarctus lacunaires ; FA : Anisotropie fractionnelle, FDR : Facteurs de risques)

#### IV. Hypothèses et Objectifs

Des altérations précoces de la microstructure de la substance blanche semblent déjà être présentes à un âge relativement jeune. En effet, plusieurs études ont montré que vers 30-40 ans des variations des marqueurs du DTI de la microstructure de la substance blanche étaient déjà présentes entre les individus. La 3<sup>ème</sup> génération de la Framingham Heart Study a montré que des modifications structurelles cérébrales subtiles étaient déjà observables chez les adultes d'une trentaine d'années en association avec l'augmentation des niveaux de pression artérielle (43). Les modifications des paramètres du DTI semblent refléter la phase

initiale des modifications de la substance blanche et les HSB seraient la phase tardive du même phénomène physiopathologique commun (39,44–47). En émettant l'hypothèse qu'on se situe sur le même phénomène biologique, on peut s'attendre à ce que les gènes, qui sont là depuis la naissance et qui prédisposent à développer des HSB vers 50-70 ans, soient déjà associés à une variabilité des marqueurs en DTI d'intégrité de la microstructure de la substance blanche, et notamment du PSMD. De plus, la MPAC étant associée à une altération des performances cognitives (notamment de la vitesse de traitement de l'information et des fonctions exécutives) chez le sujet âgé, le PSMD et les variants génétiques à risque d'HSB pourraient déjà être associés à des variations de la vitesse de traitement de l'information chez l'adulte jeune.

Nous émettons donc l'hypothèse que les facteurs de risque génétiques de MPAC (notamment d'HSB) chez la personne âgée ont déjà un impact ultra précoce sur les marqueurs de la microstructure de la substance blanche (notamment PSMD) chez l'adulte jeune et que ces modifications structurelles ont déjà un impact sur certaines performances cognitives.

Les objectifs de ce travail consistent à étudier, dans la cohorte d'étudiants i-Share :

- (i) l'association des variants à risque d'HSB, sous forme de score de risque et individuellement, avec le PSMD pour évaluer si les gènes qui prédisposent au volume de HSB chez la personne âgée, prédisposent aussi à des altérations de la microstructure de la substance blanche chez l'adulte jeune
- (ii) l'impact du score de risque génétique d'HSB et du PSMD sur la vitesse de traitement de l'information chez l'adulte jeune.

# MATÉRIELS ET MÉTHODES

## I. Schéma et population d'étude

### 1. Cohorte i-Share

Les données sont issues de la cohorte i-Share (investigateur principal (IP) Pr. Christophe Tzourio) qui est une cohorte prospective dynamique dont le recrutement a débuté en février 2013. Le but principal de l'étude i-Share est de décrire l'état de santé de la population étudiante en recueillant des informations sur un large éventail de thématiques de santé : la migraine, les infections sexuellement transmissibles, la neuroimagerie, la santé mentale, les conduites à risque et accidents. Cette étude a un volet sur la recherche biomédicale qui vise à décrire les mécanismes physiopathologiques ou psychopathologiques de certaines maladies, ainsi que l'impact de facteurs d'exposition sur le risque d'apparition de maladies chroniques sévères (48). La participation à l'étude i-Share est basée sur le volontariat. Toute personne majeure (âgée d'au moins 18 ans), étudiant dans l'enseignement supérieur (université, école, classe préparatoire ...), et parlant français pouvait participer. L'inclusion s'est fait sur le site internet « <http://www.i-share.fr/> » par un questionnaire « Qui es-tu l'étudiant ? ». Des étudiants de l'Université de Bordeaux, de l'Université de Versailles – Saint-Quentin-en-Yvelines et de multiples autres universités principalement françaises, répondant aux critères d'inclusion, pouvaient participer à i-Share. Plus de 20 000 étudiants ont été inclus à ce jour.

### 2. Deux études ancillaires : MRi-Share et Bio-Share

La cohorte i-Share comporte deux études ancillaires, Bio-Share et MRi-Share, qui forment une biobanque et une banque de neuroimagerie composées d'environ de 2 000 « sharers » volontaires, tous issus de la cohorte i-Share. Ces deux projets ont pour but de mieux comprendre certains facteurs de risque de maladies pouvant impacter la vie quotidienne des étudiants, leurs études et leur vie future, de mieux comprendre le fonctionnement cérébral, ainsi que de prévenir la survenue de maladies, dont celles pouvant survenir à un âge avancé.

#### a) *Bio-Share*

L'étude Bio-Share (IP Pr. Stéphanie Debette) consiste en la collection d'échantillons sanguins chez les étudiants inclus dans la cohorte i-Share pour la réalisation de dosages biologiques (glycémie à jeun, cholestérol total, HDL-cholestérol, LDL-cholestérol, triglycérides) et d'analyses génétiques. De plus, des échantillons d'ADN ou encore d'ARN sont conservés. L'objectif principal de bio-Share est de constituer une biobanque pour i-Share qui servira de ressource unique pour alimenter des projets de recherche variés.

### *b) MRi-Share*

Les objectifs de l'étude MRi-Share (IP Pr. Bernard Mazoyer) sont (i) de caractériser la variabilité morphologique et fonctionnelle du cerveau chez le jeune adulte en constituant une base de données d'images IRM morphologiques et fonctionnelles acquises chez 2 000 étudiant(e)s inclus dans la cohorte i-Share, (ii) de décrire leur architecture cérébrale anatomo-fonctionnelle (ii) ainsi que de caractériser la connectivité cérébrale avec ses asymétries en rapport avec les compétences cognitives des sujets. De plus, les images d'IRM cérébrales permettront notamment d'étudier la fréquence des lésions structurelles cérébrales (exemple : volume de la substance blanche) chez les étudiants et de constituer une banque de données d'images IRM accessible pour des études futures. L'étude MRi-Share implique la réalisation d'IRM cérébrales et de tests de la fonction cognitive, plus poussés que ceux effectués dans i-Share.

### 3. Projet SEGWAY

Ce stage s'insère dans le projet SEGWAY (Study of Environmental and Genome Wide predictors of early structural brain Alterations in Young students, IP Pr. Stéphanie Debette) financé par la Commission Européenne (ERC) qui utilise les données de MRi-Share et Bio-Share. L'objectif de SEGWAY est d'une part d'identifier les déterminants génétiques d'altérations structurelles cérébrales précoces en IRM chez l'adulte jeune, leur signification fonctionnelle et leur impact sur les performances cognitives et d'autre part d'explorer les déterminants génétiques partagés entre des modifications structurelles cérébrales précoces chez l'adulte jeune et tardives en population générale âgée. Il s'agit d'appréhender avec une approche « vie entière » les maladies neurologiques du sujet âgé et d'explorer les liens entre la maturation du cerveau et les pathologies cérébrales liées à l'âge. L'effet modificateur de facteurs de risque vasculaires sur les facteurs de risque génétiques en association avec ces marqueurs en IRM est également étudié.

### 4. Population d'étude

Tous les participants au projet SEGWAY ont été inclus dans cette étude.

## **II. Recueil des données**

### 1. Données d'imagerie

Les IRM cérébrales des participants de l'étude MRi-Share ont été acquises et traitées selon un protocole standardisé et interprétées dans le même centre de lecture, dans le département de neuroimagerie (Groupe d'Imagerie Neurofonctionnelle [GIN], UMR5293 CEA-CNRS-UB, Pr. Bernard Mazoyer), à Bordeaux. Les acquisitions IRM ont été effectuées sur

une machine Siemens 3T Prisma avec tous les participants positionnés de manière standardisée. Plusieurs séquences d'IRM ont été acquises. Une première imagerie anatomique (IRMa) a été réalisée avec l'acquisition d'une séquence, dite d'imagerie pondérée en T1, qui permet d'obtenir un bon contraste entre la substance blanche et la substance grise. Une deuxième imagerie anatomique, dite pondérée en T2, a été acquise avec la même résolution que l'imagerie T1, sans (séquence T2-BRAVO) ou avec compensation du signal du liquide céphalo-rachidien (séquence FLAIR-3D). L'intérêt de ces deux acquisitions est de pouvoir mesurer la myélinisation intracorticale (par le rapport T2/T1) et de mieux évaluer la charge en hypersignaux de la substance blanche, biomarqueurs de l'atteinte des petits vaisseaux. Une troisième imagerie dite de susceptibilité magnétique (SWI, susceptibility weighted imaging) a été acquise qui est utile en particulier pour l'évaluation de la présence de microhémorragies. Enfin, une imagerie de diffusion a été réalisée afin d'obtenir une image 3D des différents faisceaux de fibres blanches. Cette acquisition permet d'estimer des marqueurs de la microstructure de la substance blanche tels que l'anisotropie fractionnelle (FA) traduisant l'intégrité locale des faisceaux de fibres, la diffusivité moyenne (MD) et de quantifier le PSMD, un nouveau marqueur d'intégrité de la substance blanche décrit ci-dessus.

Sur l'IRM cérébrale, des anomalies peuvent être détectées fortuitement lors de la pratique de l'IRM cérébrale, ce sont des découvertes fortuites (DF). Il existe deux types d'anomalie, les DF adressées qui sont des anomalies présentant un possible caractère de gravité avec une prise en charge médicale et les DF non-adressées, sans d'impact sur la santé.

## 2. Données génétiques

### a) *Génotypage pangénomique*

Les données génétiques étaient issues des échantillons de sang prélevés grâce au protocole bio-Share. Une prise de sang à jeun a été réalisée au Centre de prélèvements du centre hospitalier universitaire de Bordeaux. Un volume total de 52,5 mL de sang par participant a été prélevé avec un ordre de prélèvement déterminé pour les différents tubes afin d'éviter toute contamination entre les tubes. Après extraction, l'ADN a été envoyé au centre de génomique de l'université McGill à Montréal pour des analyses de séquençage et de génotypage pangénomiques. Un séquençage pangénomique à 30X (qui consiste à déterminer la séquence nucléotidique complète sur l'ensemble du génome) a été réalisé dans un sous-groupe de 600 participants et a pu être réalisé avec succès pour 595 participants. Un génotypage pangénomique a été réalisé pour 1 900 participants. Ce dernier a été réalisé avec succès pour 1 872 participants. Le séquençage et/ou génotypage pangénomique n'ont pas pu être réalisés sur certains échantillons, du fait d'un volume ou une quantité d'ADN insuffisants. Le génotypage pangénomique a été effectué sur une puce Affymetrix Precision Medicine Axiom

array pouvant mesurer environ 1 millions de variants génétiques distribués sur l'ensemble du génome. Une imputation sur le panel de référence Haplotype Reference Consortium a été réalisée selon les derniers standards et a permis d'imputer les génotypes des SNPs non génotypés par la puce.

#### *b) Contrôle qualité*

Un contrôle qualité des données a été réalisé par l'équipe VINTAGE. Pour les données issues du génotypage pangénomique de l'ADN, une première étape du contrôle qualité a été réalisée sur les variants (**Annexe 2**). Les variants non-autosomaux sont exclus. Un filtre sur la fréquence de l'allèle mineur (MAF : *minor allele frequency*) a été réalisé, ainsi tous les variants avec une MAF inférieure à 1 %, qui sont généralement imputés avec une moins bonne précision, ont été exclus. Tous les variants avec un taux de succès du génotypage (« call rate ») < 0,98 ont été exclus, reflétant une qualité des échantillons d'ADN. Tous les variants avec un équilibre de Hardy Weinberg (HWE) significatifs ( $p < 0,001$ ) ont été exclus. Le théorème de HWE indique que dans une population idéale, les fréquences des génotypes demeurent constantes à travers les générations, en l'absence d'autres influences évolutives et donc qu'il existe un équilibre. Donc tous les SNPs déviant de cet équilibre peuvent être dus à une erreur de génotypage ou à une stratification de la population et donc ont été exclus. Ensuite un contrôle qualité a été réalisé sur les participants. Tous les participants avec taux de succès du génotypage (« call rate ») < 0,05, ceux une hétérozygotie au-delà de plus ou moins 6 écart-type indiquant un signe de contamination des échantillons ou une consanguinité, ceux avec une incohérence de sexe, ainsi que ceux avec un génotypage en double sont exclus. Une exclusion des sujets avec un degré de parenté > 0,185 a été réalisée, correspondant environ au cousin au 3<sup>ème</sup> degré. Ces exclusions ont été nécessaires du fait d'une corrélation entre les phénotypes des individus. Cette parenté pourrait biaiser les analyses. Une analyse en composante principale a été réalisée où les outliers ont été exclus du fait d'une ascendance non-européenne. Un contrôle des variants a de nouveau été réalisé sur l'échantillon des participants ayant passé le contrôle qualité. Enfin, une vérification de l'imputation a été réalisée.

### 3. Données cognitives

Les données cognitives ont été recueillies par les tests réalisés en ligne dans le questionnaire « Comment ça va les neurones ? ». Le test de Stroop, évaluant les capacités d'inhibition des fonctions exécutives a été retenu pour évaluer la cognition chez les participants pour ce mémoire.

Un contrôle qualité des données a été réalisé par l'équipe du GIN. Ces tests s'étant déroulés sur internet, une variabilité existait entre les différents navigateurs web, les types d'ordinateurs

et les caractéristiques de connexion internet (rapidité variable) utilisés par chaque étudiant. Chaque item avait un temps de présentation prédéfini. Il était de 3 secondes pour le test de Stroop. Un seuil de +10 % du temps total de présentation de l'item a été choisi pour détecter les temps aberrants et prendre en compte la variabilité due à la machine. Les temps aberrants ont donc été exclus. Une vérification de l'utilisation de la bonne touche a été réalisée. En effet, il était indiqué de taper la réponse sur une touche du clavier (par exemple sur la touche « r » pour indiquer la couleur « rouge »). Enfin, un contrôle de la validation de la tâche par le sujet a été réalisé. Une fois la tâche terminée, il était demandé aux participants si le test s'était correctement déroulée et si non, quelles en étaient la raisons (interruptions, problèmes techniques).

#### 4. Autres données

Pour tous les participants le recueil d'information de l'étude i-Share s'est fait via des questionnaires internet auto-administrés complétés à l'inclusion, interrogeant les participants sur leurs caractéristiques personnelles, leur historique médical, leurs habitudes de vie, leur statut socio-économique, avec une attention particulière portée sur les comportements à risque, l'activité physique, l'apport nutritionnel, le sommeil et à la santé mentale. En complément de ces questionnaires, certains participants ont réalisé un check-up médical en face à face avec un examen physique.

Le questionnaire d'inclusion dans l'étude i-Share a permis de recueillir des données de type :

- sociodémographiques (âge, sexe, type d'étude),
- cliniques constituées par les facteurs de risque vasculaires tels que rapportés sur le questionnaire en ligne par le participant : marqueurs anthropométriques (poids et taille) et le statut tabagique.

Pour le sous-groupe de participants ayant réalisé le check-up médical des paramètres anthropométriques mesurés (poids et taille) et des mesures de la pression artérielle ont été recueillis par un professionnel de santé.

### III. **Variables étudiées**

#### 1. Variables d'intérêts

##### a) *PSMD (peak width of skeletonized mean diffusivity)*

Le premier évènement d'intérêt étudié, pour l'objectif 1, était le PSMD : « peak width of skeletonized mean diffusivity ». Sa construction repose sur deux étapes : les images du DTI ont été squelettisées puis est réalisée une analyse par histogramme (**Figure 6**). Ce qu'on appelle « squelette » correspond à la reconstruction des trajets des principales fibres de substance blanche d'un cerveau humain. La squelettisation correspond donc à l'extraction des



fibres de substance blanche qui forment le squelette. Cette squelettisation a été réalisée par la procédure TBSS (Tract-based Spatial Statistics procedure) (49). Très brièvement, les images du DTI issues de la FA ont été alignées sur un modèle standard en utilisant un enregistrement non linéaire (avec l'outil FNIRT), qui est un algorithme dont la procédure d'optimisation non linéaire vise à minimiser les différences entre les deux images (celle de l'échantillon et celle de référence). Ensuite, les valeurs de FA de chaque sujet ont été projetées sur un squelette, issu du modèle standard. Les images obtenues de MD ont été projetées sur le même squelette, en utilisant les paramètres de la projection dérivée du squelette de la FA. Enfin, un histogramme des valeurs de MD du squelette a été réalisé. Le PSMD correspond à la différence des 95<sup>ème</sup> et 5<sup>ème</sup> percentiles de ces valeurs de MD. Il est exprimé en  $10^{-4}\text{mm}^2/\text{s}$ . Une valeur de PSMD plus élevée indique une dispersion plus élevée des valeurs de diffusivité moyenne, indicatrices d'altérations locales au sein de la microstructure de la substance blanche.

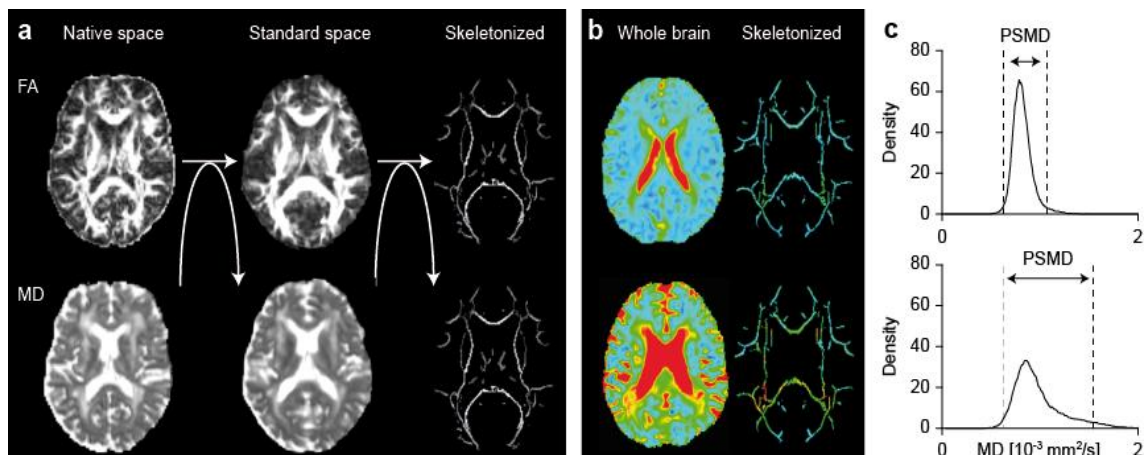


Figure 6. Procédure de la quantification du marqueur PSMD. (Baykara et al. Ann Neurol. 2016).

(A) Illustration de la procédure automatique de la squelettisation. Les images individuelles de FA ont été normalisées dans un espace standard et projetées dans squelette de référence. Ensuite, la transformation et les paramètres de projection sur le squelette ont été appliqués pour les images issues de la MD. (B) Exemple de MD cartes de sujets projetés dans le squelette de référence. (C) L'analyse par histogramme des mêmes images des exemples en B.

Un contrôle qualité a été effectué sur cette variable. Une vérification des valeurs extrêmes a été réalisée selon la méthode univariée de Tuckey basée sur l'interquartile range (IQR, Q3-Q1) pour déceler des outliers.

#### b) Tests cognitifs

Pour l'objectif 2, nous avons utilisé comme mesure d'intérêt le test de Stroop (50) qui teste les fonctions exécutives. Le principe du test de Stroop est simple : le nom d'une couleur est écrit avec une encre de couleur différente. La tâche du sujet est de nommer la couleur de l'encre de chaque mot (et non lire le mot écrit). Ce test a été réalisé en ligne, sur le site d'i-Share. 60

stimuli ont été réalisés avec un ordre de présentation déterminé. Pour chaque stimulus, le participant avait 3 secondes pour répondre. Les stimuli de ce test consistaient en :

- 20 items incongruents avec un mot (jaune ou rouge ...) écrit dans une couleur (bleu ou vert ...) différente du mot. Par exemple, la liste de mot suivant : « ROUGE BLEU VERT JAUNE » devait être lu comme ceci : « BLEU JAUNE ROUGE VERT ».
- 20 items congruents : le mot bleu coloré en bleu. Par exemple, la liste de mot suivant : « ROUGE BLEU VERT JAUNE » devait être lu comme ceci : « ROUGE BLEU VERT JAUNE ».
- 20 items neutres : un distracteur (fort, grave, plein, neuf) avec pour cible la couleur du distracteur. Par exemple, la liste de mot suivant : « FORT GRAVE PLEIN NEUF » devait être lu comme ceci : « ROUGE BLEU VERT JAUNE ».

Dans l'attente de la disponibilité de l'ensemble des scores permettant de caractériser les fonctions exécutives (contrôle qualité en cours), une première approche pour ce mémoire a été d'utiliser le temps de réaction des participants au test de Stroop comme reflet de la vitesse de traitement de l'information, dont le ralentissement est une des premières manifestations de la MPAC (51).

## 2. Expositions d'intérêts

### a) *Le Score de risque génétique (GRS)*

La principale exposition d'intérêt pour l'objectif 1 et une des expositions d'intérêt pour l'objectif 2 était le score de risque génétique des HSB (GRS). Ce score a été construit en plusieurs étapes. Tout d'abord, une sélection des top-SNPs significativement associés au volume des HSB ( $p < 5 \times 10^{-8}$ ) a été effectuée, à partir d'une étude d'association génétique pangénomique (GWAS) en cours de finalisation au sein du consortium CHARGE (*Cohorts for Heart and Aging Research in Genomic Epidemiology*), coordonnée par le Dr. Murali Sargurupremraj et le Pr. Stéphanie Debette (U1219). Cette méta-analyse de GWAS des HSB a été effectuée dans une population âgée en moyenne de 66,8 ans (écart-type = 6,8) et sur 40 869 participants, permettant de tripler le nombre de loci identifiés comme étant associés au volume d'HSB (Sargurupremraj et. Al, unpublished data.). 21 SNPs indépendant ont ainsi été sélectionnés. Le Manhattan plot d'identification de ces SNPs est présenté en **Annexe 3**. Deuxièmement, une vérification de l'indépendance de ces SNPs a été effectuée pour éviter une redondance de l'information, en éliminant les SNPs en déséquilibre de liaison avec un  $r^2 > 0.1$  (dans un locus donné le variant associé le plus significativement avec le phénotype). Ensuite, un GRS pondéré (wGRS) a été construit à partir de la formule ci-dessous :

$$wGRS = w_1 \times SNP_1 + w_2 \times SNP_2 + \dots + w_k \times SNP_k$$

Chaque SNP était codé en 0,1 ou 2 correspondant au dosage allélique, basé sur le modèle additif, et qui équivaut au nombre de copie de l'allèle à risque. Prenons l'exemple d'un SNP avec pour possibilité d'allèle A/G. Chaque individu possédant 2 copies de chaque chromosome, il y a 3 combinaisons ou génotypes possibles pour ce SNP que sont AA, AG, GG. Si G est l'allèle à risque d'HSB élevés, nous comptons donc le nombre d'allèles à risque que l'individu possède, c'est-à-dire zéro, une ou deux copies. Le GRS pondéré correspond donc à la somme des dosages alléliques, basés sur le modèle additif, des SNPs associés à un volume plus élevé d'HSB, pondérés par leur bêta ( $\beta$ ) d'association issu du GWAS des HSB, en émettant l'hypothèse que ces SNPs ont des effets indépendants (d'où le filtre susmentionné sur le degré de déséquilibre de liaison) et contribuent de manière additive au volume d'HSB. Enfin, nous avons créé un « score de risque génétique pondéré redimensionné » (wGRSres) selon la formule ci-dessous (53) :

$$wGRSres = \frac{k (w_1 \times SNP_1 + w_2 \times SNP_2 + \dots + w_k \times SNP_k)}{w_1 + w_2 + \dots + w_k}$$

Le GRS pondéré a été multiplié par le nombre de SNPs utilisés dans la construction du GRS, puis a été divisé par la somme de bêtas. Chaque augmentation d'une unité du GRS correspondait ainsi à l'augmentation d'un allèle à risque, ce qui facilite l'interprétation des données. Bien que le GRS constitue la variable explicative principale, afin de mieux comprendre et interpréter nos résultats nous avons, dans une analyse secondaire, également étudié individuellement l'association de chaque SNP constitutif du GRS avec le PSMD.

#### b) PSMD

Le PSMD a également été étudié comme exposition d'intérêt pour l'objectif 2 et sa définition correspond à celle donnée ci-dessus. La variable PSMD a été étudiée en continu puis catégorisée en quartiles.

### 3. Covariables

Les variables pouvant être associées au PSMD, à la composante génétique et à la cognition et pouvant potentiellement être des facteurs de confusion dans les associations observées ont été identifiés lors d'un entretien avec la directrice de stage de l'étude, en confrontation avec la littérature, en prenant en compte leur pertinence clinique et la disponibilité des données.

#### a) Les facteurs sociodémographiques

- Sexe
- Age en années
- Type d'étude
  - Santé

- Sciences
- Gestion, économie, commerce, marketing, droit
- Arts, Multimédia, Sciences Sociales et de l'homme
- Ecole d'ingénieurs et Sciences de l'ingénieur et Autres
- Autres

*b) Les facteurs de risques vasculaires*

- Pression artérielle systolique en cmHg
- L'indice de masse corporelle (IMC) en kg/m<sup>2</sup>, calculé à partir de la formule suivant :  
Poids / Taille<sup>2</sup>
- La consommation de tabac :
  - oui
  - oui mais en train d'arrête
  - non

Les données issues du check-up médical n'étant pour l'instant disponibles que dans un sous-groupe des participants les analyses tenant compte de ces paramètres sont considérées comme des analyses secondaires pour ce mémoire. Un contrôle qualité a été effectué sur les données de check-up avec une vérification et exclusions des valeurs aberrantes.

#### **IV. Stratégie d'analyse**

Toutes les analyses ont été réalisées avec le logiciel R© version 3.5.1 (2018-07-02), exploité par l'interface graphique RStudio et avec le logiciel SAS® version 9.3 (SAS Institute Inc., Cary, NC, USA).

##### 1. Analyse descriptive de la population d'étude

Une description des caractéristiques sociodémographiques et des facteurs de risque vasculaires des participants a été réalisée et exprimée en termes de moyenne et d'écart-type pour les variables quantitatives et d'effectif et pourcentage pour les variables qualitatives.

Des mesures d'associations, à l'aide de régressions linéaires, ont été estimées entre nos variables d'intérêt (PSMD et test de Stroop) avec les facteurs démographiques et les facteurs de risques vasculaires, en analyses univariées puis ajustées sur l'âge, le sexe et le volume total intracrânien pour le PSMD et sur l'âge, le sexe et le type d'étude pour le test de Stroop.

##### 2. Analyses principales

Les associations étudiées ont été estimées par plusieurs modèles de régressions linéaires multivariés afin de répondre aux différents objectifs. Pour le premier objectif, la variable dépendante était le PSMD et la variable d'exposition intérêt principal était la composante

génétique, d'abord traité sous la forme du wGRSres d'HSB puis pour chaque SNP, utilisé dans la construction du wGRSres d'HSB.

- Modèle 1 : Pour tester l'association du PSMD avec le wGRSres d'HSB, nous avons utilisé une régression linéaire ajustée sur l'âge, le sexe, le volume intracrânien et les quatre premières composantes principales de stratification de la population. Ces composantes principales sont utilisées pour tenir compte de la sous-structure populationnelle latente, afin de s'affranchir d'associations faussement positives (ou faussement négatives) en lien avec des différences de fréquence allélique dues à des origines géographiques différentes de sous-populations au sein de notre population d'étude (54). Le volume intracrânien est utilisé comme variable d'ajustement pour tenir compte de différences morphologiques qui pourraient impacter les mesures de PSMD.
- Modèle 1bis : Un modèle a été construit pour chaque SNP utilisé dans le wGRSres d'HSB, soit 21 modèles, avec un ajustement sur l'âge, le sexe, le volume intracrânien et les quatre premières composantes principales.

Pour répondre au deuxième objectif, la variable dépendante était le temps de réaction au test de Stroop et la variable d'exposition principale était, tout d'abord, la composante génétique (traité sous la forme de wGRSres d'HSB et avec les SNPs pris individuellement) puis le PSMD (traité d'abord en continu puis en quartiles).

- Modèle 2 : Pour tester l'association du temps de réaction au test de Stroop avec le wGRSres d'HSB, nous avons utilisé une régression linéaire ajustée sur l'âge, le sexe, les quatre premières composantes principales de stratification de la population et le niveau d'étude. Ce modèle 2 correspond à un ajustement minimal dans l'étude de la cognition et d'un facteur génétique. Les quatre premières composantes principales ont été prises en compte pour tenir compte de la sous-structure populationnelle latente. Usuellement, le niveau d'étude est utilisé pour ajuster sur la différence de niveau d'étude (primaire à études supérieures). Notre échantillon étant composé que de sujets ayant tous atteints les études supérieures, nous avons pris en compte le type d'étude.
- Modèle 2bis : Un modèle a été construit pour chaque SNP utilisés dans le wGRSres d'HSB, soit 21 modèles, avec un ajustement l'âge, le sexe, les quatre premières composantes principales de stratification de la population et le type d'étude.
- Modèle 3 : Pour tester l'association du temps de réaction au test de Stroop avec le PSMD pris en continu, nous avons utilisé une régression linéaire ajustée sur le sexe, l'âge, le volume intracrânien et le type d'étude. Ce modèle 3 correspond à un ajustement minimal dans l'étude de la cognition et d'un marqueur IRM cérébral. Le volume intracrânien est utilisé comme variable d'ajustement

pour tenir compte de différences morphologiques qui pourraient impacter les mesures de PSMD et nos participants ayant tous atteints les études supérieures, nous avons pris en compte le type d'étude.

- Modèle 3bis : Pour tester l'association du temps de réaction au test de Stroop avec le PSMD pris en quartile, nous avons utilisé une régression linéaire ajustée sur le sexe, l'âge, le volume intracrânien et le type d'étude.

Les hypothèses de normalité et d'homoscédasticité ont été vérifiées par méthodes graphiques. L'hypothèse de linéarité a été vérifiée par utilisation des polynômes fractionnaires dans les modèles de régression linéaire, et les variables quantitatives ont été modélisées selon les résultats obtenus.

Les coefficients de détermination ( $R^2$ ) ont été calculés dans les analyses de régression linéaire multivariées entre le wGRSres d'HSB et le PSMD afin de déterminer la proportion de variance dans du PSMD expliquée par le score de risques génétique d'HSB.

Pour les analyses secondaires s'intéressant à chaque locus à risque d'HSB individuellement, la correction de Bonferroni a été utilisée pour prendre en compte la multiplicité des tests. La formule suivante a été utilisée :  $\alpha = (0.05/k)$  avec k étant le nombre de tests réalisés. Le seuil de significativité était de 0,0023 (0,05/21).

### 3. Analyses complémentaires

Une première analyse complémentaire a été réalisée en prenant en compte dans nos ajustements les facteurs de risques vasculaires mesurés au check-up dans le sous-échantillon de participants ayant bénéficié de cet examen. Nous nous sommes intéressés en particulier à la pression artérielle, l'IMC, la consommation de tabac, ces facteurs étant associés au risque de maladie des petites artères (55–58) et à sa sévérité, tout particulièrement la pression artérielle. Dans cette analyse secondaire tous les modèles de notre analyse principale ont été ajustés en plus sur ces variables (sauf les modèles 1bis et 2bis pour lesquels ces analyses secondaires n'ont pas été effectuées).

Une deuxième analyse complémentaire à consister à exclure les participants pour lesquels une DF adressé et non adressée afin d'évaluer l'impact potentiel de ces découvertes fortuites sur nos analyses. Puis, nous avons également comparé les caractéristiques des participants avec et sans DF. Des tests de  $\chi^2$  ou de Fisher exact ont été réalisés pour les variables qualitatives et des tests de Student pour les variables quantitatives. Les tests ont été effectués au risque  $\alpha$  de 5%.

# RÉSULTATS

## I. Population d'étude

Tous les participants de l'étude i-Share, étudiant à Bordeaux, étaient éligibles au projet SEGWAY. Parmi les 2 000 volontaires screenés pour le projet SEGWAY, 1 999 ont été inclus. Après exclusion des participants ayant abandonné ou refusé, ceux sans tubes dédiés à l'extraction ADN, ceux ayant un consentement non conforme, ceux ayant un contrôle qualité post extraction ADN non valable, 1 900 participants possédaient des tubes pour l'extraction ADN. Après exclusion des tubes dont la quantité d'ADN était insuffisante, 1 872 participants ont eu leur ADN extraite. Après le contrôle qualité du génotypage, les données de 1 865 participants étaient disponibles pour les analyses.

Parmi les 1 999 inclus, 1 964 participants respectaient les critères d'inclusion pour le protocole IRM. Après exclusion des participants n'ayant pas réalisé l'IRM, ceux ayant retiré leur consentement et celui ayant une découverte fortuite avec des valeurs extrêmes, 1 866 participants avaient réalisé un IRM et donc les données étaient exploitables.

Après exclusion des données manquantes sur le PSMD pour les données d'imagerie cérébrales et application du filtre sur la parenté pour les données génétiques ainsi qu'une exclusion de 5 participants avec un identifiant non-correspondant, 1 692 participants avaient à la fois des données d'IRM et des génotypes pangénomiques utilisables pour l'objectif 1. Pour l'objectif 2, 1 283 participants avaient également des données pour le test de Stroop ayant satisfait les critères de contrôle qualité (**Figure 7**). Par ailleurs, parmi les participants avec données IRM et génétiques utilisables, 854 avaient également des données cliniques issues du check-up médical utilisables après contrôle qualité.

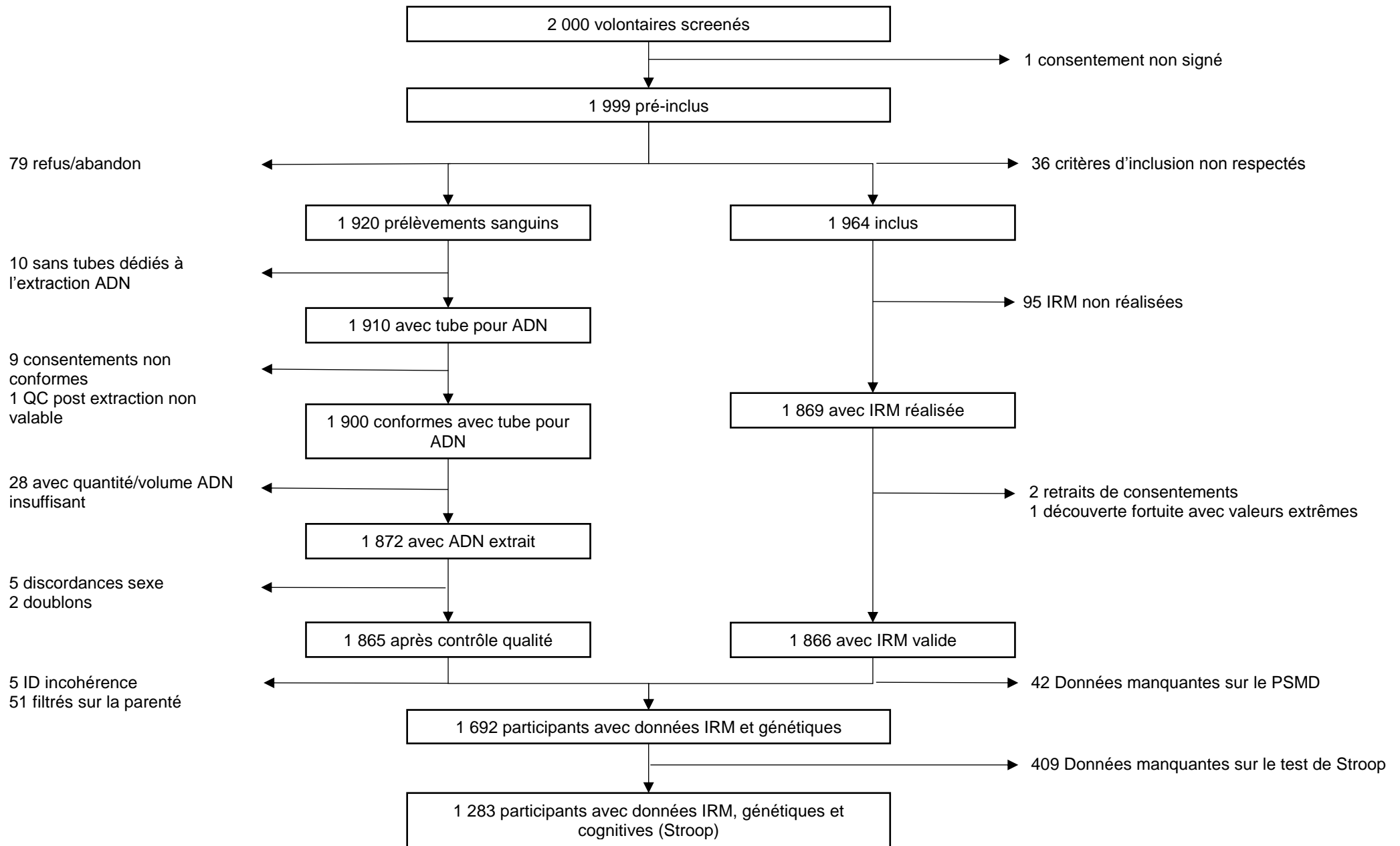


Figure 7. Diagramme de flux de la population d'étude. Cohorte i-Share, SEGWAY. Bordeaux. 2019.



## I. Description de la population

Les caractéristiques démographiques et la distribution des facteurs de risque vasculaires de la population étudiée sont décrites dans le Erreur ! Référence non valide pour un signet.. Avec une forte majorité de femmes (72,0 %), l'âge moyen des étudiants participant à l'étude SEGWAY était de 22 ans (écart-type = 2,3 ; min = 18 ; max = 34). Au moment de leur inclusion dans l'étude i-Share, la majorité des participants étudiaient dans le secteur de la santé (55,0 %), venant ensuite le secteur de l'Arts, Multimédia, Sciences Sociales et de l'homme (17,4 %). Presque d'un quart des étudiants à leur inclusion dans i-Share étaient des fumeurs (23,8 %) et 5,4 % des étudiants étaient en train d'arrêter. Un peu plus d'un tiers des participants consommaient une fois par semaine ou moins de l'alcool (37,8 %), et 22,4 % des participants consommaient entre 2 à 3 fois par semaine. Les participants répondaient en moyenne en 1,26 secondes (ET = 0,19) au test de Stroop.

Les participants ayant réalisé le check-up médical, en moyenne, mesuraient 1,68 m (ET = 0,08 ; min = 1,5 m ; max = 1,97 m), pesaient 62,3 kg (ET = 11,77) et avait un IMC de 21,9 kg/m<sup>2</sup> (ET = 3,22).

Les associations de nos variables d'intérêt que sont le PSMD et le temps de réaction au test de Stroop avec les facteurs sociodémographiques et facteurs de risques vasculaires ont été décrites dans le **Tableau 2**. Le PSMD augmentait significativement avec l'âge ( $p = 0,00001$ ) et était significativement plus élevé chez les hommes que chez les femmes ( $p = 2 \times 10^{-16}$ ). Le PSMD augmentait significativement avec la pression artérielle systolique ( $p = 7,3 \times 10^{-11}$ ) et l'IMC ( $p = 0,0001$ ). Après ajustement sur l'âge, le sexe et le volume intracrânien, ces associations restaient significatives. Aucune association n'a été retrouvée entre nos facteurs sociodémographiques et facteurs de risques vasculaires et la vitesse de traitement de l'information au test de Stroop (**Annexe 4**).

Tableau 1. Description de la population d'étude, N = 1 692. Cohorte i-Share, SEGWAY. Bordeaux, France. 2019.

Variables	Population d'étude N = 1 692		
	N	Effectif / Moyenne	Pourcentage / Ecart-type
<b>Facteurs sociodémographiques</b>			
<b>Age à l'IRM (années)</b>	1692	22	2,3
<b>Sexe</b>	1692		
Homme		473	28,0
Femme		1216	72,0
<b>Type d'études</b>	1688		
Santé		928	55,0
Sciences		179	10,6
Gestion, économie, commerce, marketing, droit		141	8,4
Arts, Multimédia, Sciences Sociales et de l'homme		290	17,2
Ecole d'ingénieurs et Sciences de l'ingénieur		30	1,8
Autres		120	7,1
<b>Consommation d'alcool</b>	1690		
Jamais		66	3,9
1 fois par an		41	2,4
Plusieurs fois par an		261	15,4
1 fois par mois		233	13,8
1 fois par semaine ou moins		638	37,8
2 à 3 fois par semaine		378	22,4
4 à 6 fois par semaine		63	3,7
Tous les jours		10	0,6
<b>Variables d'intérêts</b>			
<b>PSMD (<math>10^{-4} \text{ mm}^2/\text{s}</math>)</b>	1692	1,53	0,13
<b>Stroop (secondes)</b>	1283	1,26	0,19
<b>Facteurs de risques vasculaires</b>			
<b>Consommation de Tabac</b>	1690		
Oui		403	23,8
Oui, mais en train d'arrêter		91	5,4
Non		1196	70,8
<b>Pression Artérielle (PA)</b>			
Pression Artérielle systolique (en cmHg)	865	12,46	1,36
<b>Mesures Anthropométriques</b>			
Taille (m)	854	1,68	0,08
Poids (kg)	854	62,3	11,77
IMC ( $\text{kg}/\text{m}^2$ )	854	21,9	3,15

Tableau 2. Associations du PSMD avec des facteurs sociodémographiques et de risques vasculaires de la MPAC. Cohorte i-Share, SEGWAY. Bordeaux, France. 2019.

Variables	PSMD (10 <sup>-4</sup> mm <sup>2</sup> /s)				PSMD (10 <sup>-4</sup> mm <sup>2</sup> /s) *			
	N	β	IC 95%	p	N	β	IC 95%	p
<b>Age (/ 1 an)</b>	1 692	0,006	0,003 ; 0,009	<b>0,00001</b>	1 692	0,005 <sup>¶</sup>	0,002 ; 0,008	<b>0,00016</b>
<b>Sexe (F vs H)</b>	1 692	-0,090	-0,104 ; -0,076	<b>2x10<sup>-16</sup></b>	1 692	-0,073 <sup>§</sup>	-0,091 ; -0,055	<b>2x10<sup>-15</sup></b>
<b>Type d'études</b>	1 692				1 692			
<i>Santé</i>		1				1		
<i>Sciences</i>		0,006	-0,015 ; 0,028	0,561		-0,005	-0,026 ; 0,015	0,618
<i>Gestion, économie, commerce, marketing, communication, droit</i>		-0,019	-0,043 ; 0,005	0,127		-0,011	-0,034 ; 0,012	0,340
<i>Arts, Multimédia, Sciences Sociales et de l'homme</i>		-0,010	-0,028 ; 0,008	0,270		-0,006	-0,024 ; 0,011	0,475
<i>Ecole d'ingénieurs et Sciences de l'ingénieur</i>		0,033	-0,016 ; 0,084	0,190		0,010	-0,038 ; 0,058	0,680
<i>Autres</i>		-0,027	-0,053 ; -0,001	<b>0,044</b>		-0,018	-0,043 ; 0,006	0,156
<b>Consommation de tabac</b>	1 690				1 690			
<i>Non</i>		1				1		
<i>Oui</i>		-0,012	-0,028 ; 0,002	0,105		-0,012	-0,026 ; 0,002	0,115
<i>Oui, mais en train d'arrêter</i>		0,004	-0,025 ; 0,033	0,771		0,006	-0,021 ; 0,034	0,658
<b>Pression Artérielle systolique (/cmHg)</b>	854	0,022	0,015 ; 0,029	<b>7,3x10<sup>-11</sup></b>	854	0,018	0,004 ; 0,032	<b>0,011</b>
<b>IMC (/5 kg/m<sup>2</sup>)</b>	854	0,028	0,013 ; 0,042	<b>0,0001</b>	854	0,010	0,002 ; 0,017	<b>0,008</b>

\* Ajusté sur le sexe, l'âge et le volume intracrânien

¶ Ajusté sur le sexe et le volume intracrânien

§ Ajusté sur l'âge et le volume intracrânien

## II. Résultats des analyses d'association

### 1. Association du PSMD avec le wGRSres d'HSB

La relation entre le wGRSres d'HSB et le PSMD (variable dépendante, continue) a été étudiée dans plusieurs modèles linéaires univariés et multivariés dont les résultats sont présentés dans le **Tableau 3**.

Tableau 3. Association entre le wGRSres d'HSB et le PSMD, N = 1 692. Cohorte i-Share. SEGWAY. Bordeaux, France. 2019.

Variables	Chr:position (hg19)	Gène le plus proche	Qualité Imputation (R <sup>2</sup> )	PSMD (10 <sup>-4</sup> mm <sup>2</sup> /s)		
				β	IC 95%	p
<b>wGRSres d'HSB</b>				0,001	0,0003 ; 0,0024	<b>0,0114</b>
<b>wGRSres d'HSB (ajusté sur âge, sexe, VIC*, PC 1-4**)</b>				0,001	0,0003 ; 0,0023	<b>0,0087</b>
<b>SNPs (ajusté sur âge, sexe, VIC*, PC 1-4**)</b>						
rs786921	1:89286673	<i>PKN2</i>	0,98	0,0014	-0,0007 ; 0,0010	0,763
rs7566761	2:43118872	<i>HAAO</i>	0,97	0,0078	-0,0034 ; 0,0192	0,171
rs7596872	2:56128091	<i>EFEMP1</i>	0,97	-0,0086	-0,0238 ; 0,0065	0,263
rs62172472	2:188028317	<i>CALCRL</i>	0,85	0,0018	-0,0093 ; 0,0131	0,741
rs76122535	2:203664929	<i>ICA1L/NBEAL1</i>	0,97	-0,0036	-0,0161 ; 0,0087	0,561
rs830179	3:183380035	<i>KLHL24</i>	0,95	0,0161	0,0065 ; 0,0256	<b>0,0009</b>
rs7733216	5:82857870	<i>VCAN</i>	0,96	0,0257	0,0014 ; 0,0035	<b>0,000002</b>
rs17148926	5:121510586	<i>LOC100505841</i>	0,94	-0,0002	-0,0116 ; 0,0111	0,959
rs275350	6:151016058	<i>PLEKHG1</i>	0,99	-0,0002	-0,0090 ; 0,0087	0,957
rs17149723	8:8171487	<i>SGK223</i>	0,98	-0,0008	-0,0109 ; 0,0092	0,868
rs4630220	10:105459116	<i>SH3PXD2A</i>	0,89	0,0083	-0,0015 ; 0,0183	0,099
rs10883926	10:105608838	<i>SH3PXD2A</i>	0,99	0,0141	0,0051 ; 0,0231	<b>0,0022</b>
rs12588905	14:52595066		0,99	-0,0019	-0,0108 ; 0,0069	0,666
rs1285841	14:91881387	<i>CCDC88C</i>	0,97	0,0064	-0,0024 ; 0,0153	0,155
rs55989340	14:100635222	<i>DEGS2</i>	0,98	0,0067	-0,0031 ; 0,0166	0,179
rs1895029	16:51461433	<i>SALL1</i>	0,99	-0,0058	-0,0145 ; 0,0028	0,189
rs12921170	16:87227397	<i>C16orf95</i>	0,98	-0,0052	-0,0141 ; 0,0037	0,251
rs6503417	17:43144218	<i>NMT1/DCAKD</i>	0,99	0,0099	0,0007 ; 0,0192	<b>0,034</b>
rs34974290	17:73888354	<i>TRIM65</i>	0,99	0,0017	-0,0089 ; 0,0124	0,747
rs429358	19:45411941	<i>APOE</i>	0,99	0,0054	-0,0080 ; 0,0189	0,425
rs5762197	22:27887471	<i>MN1</i>	0,97	0,0041	-0,0051 ; 0,0132	0,382

\* VIC : volume intracrânien

\*\* 4 premières composantes principales

wGRS<sub>res</sub> d'HSB correspond au score de risque génétique des HSB pondéré et redimensionné.

Après ajustement sur l'âge, le sexe, le volume intracrânien et les quatre premières composantes principales de stratification de la population (modèle 1), le wGRSres d'HSB était significativement associé à un PSMD plus élevé ( $p=0,0087$ ). Ainsi les allèles liés à une augmentation du volume des HSB chez la personne âgée étaient associés à une augmentation des valeurs du PSMD à un âge jeune. Le wGRSres d'HSB expliquait 0,31 % de la variation du PSMD.

En étudiant l'effet des SNPs pris individuellement sur le PSMD avec un ajustement sur l'âge, le sexe, le volume intracrânien et les quatre premières composantes principales de stratification de la population (modèle 1bis), 4 SNPs étaient significativement associés au PSMD au seuil de 5%. Après application de la correction de Bonferroni pour prise en compte des tests multiples, 3 SNPs restaient significativement associés au PSMD. Le gène le plus proche du SNP rs830179 était *kelch like family member 24 (KLHL24)*, pour le SNP rs7733216 le gène le plus proche était *versican (VCAN)* et pour le SNP rs10883926 le gène le plus proche était *SH3 and PX domains 2A (SH3PXD2A)*.

## 2. Association du temps de réaction au test de Stroop avec le wGRSres d'HSB et avec le PSMD

L'association du temps de réaction au test de Stroop avec le wGRSres d'HSB a été étudiée dans plusieurs modèles linéaires multivariés (**Tableau 4**). Après ajustement sur l'âge, le sexe, le type d'étude (ou cursus) et les quatre premières composantes principales de stratification de la population (modèle 2), le wGRSres d'HSB n'était pas associé au temps de réaction des participants au test de Stroop ( $p = 0,174$ ). Néanmoins, en étudiant l'effet des SNPs pris individuellement sur le temps de réaction du test de Stroop, et en ajustant sur l'âge, le sexe, le type d'étude et les quatre premières composantes principales de stratification de la population (modèle 2bis), 3 SNPs étaient significativement associés au test de Stroop au seuil de 5%. Après application de la correction de Bonferroni, 1 SNP restait significativement associé au temps de réaction du test de Stroop ( $p = 0,0015$ ). Pour ce SNP (rs62172472) l'allèle à risque d'augmentation du volume d'HSB chez la personne âgée était significativement associé à une augmentation du temps de réaction au test de Stroop, donc un ralentissement de la vitesse de traitement de l'information. Le gène le plus proche de rs62172472 est *calcitonin receptor like receptor (CALCRL)*.

L'association du temps de réaction au test de Stroop avec le PSMD, considéré en continu puis en quartiles, a également été étudiée dans plusieurs modèles linéaires multivariés (**Tableau 4**). Après ajustement sur l'âge, le sexe, le type d'étude et le volume intracrânien (modèle 3), une augmentation des valeurs du PSMD, donc des valeurs plus pathologiques, était associée avec une augmentation du temps de réaction au test de Stroop mais de façon non significative ( $p = 0,071$ ). En étudiant le PSMD en quartile (modèle 3bis), l'effet global du PSMD était non

significativement associé au temps de réaction du test Stroop ( $p = 0,403$ ). Néanmoins, nous avons constaté que, comparativement au premier quartile, les participants dans le quatrième quartile de distribution du PSMD avaient un temps de réaction significativement plus élevé au test de Stroop ( $p = 0,031$ ).

Tableau 4. Association entre le  $wGRS_{res}$  d'HSB, le PSMD et le test de Stroop,  $N = 1283$ . Cohorte *i*-Share. SEGWAY. Bordeaux, France. 2019.

Variables	Chr:position (hg19)	Gène le plus proche	Qualité Imputation ( $R^2$ )	STROOP TR (/sec)		
				$\beta$	IC 95%	$p$
<b>wGRS<sub>res</sub> d'HSB</b>				-0,0012	-0,0029 ; 0,0005	0,167
<b>wGRS<sub>res</sub> d'HSB (ajusté sur âge, sexe, PC 1-4**, cursus)</b>				-0,0011	-0,0029 ; 0,0005	0,174
<b>SNPs (ajusté sur âge, sexe, PC 1-4**, cursus)</b>						
rs786921	1:89286673	<i>PKN2</i>	0,98	-0,0071	0,0226 ; 0,0082	0,362
rs7566761	2:43118872	<i>HAAO</i>	0,97	-0,0029	-0,0221 ; 0,0162	0,766
rs7596872	2:56128091	<i>EFEMP1</i>	0,97	-0,0279	0,0534 ; -0,0023	<b>0,032</b>
rs62172472	2:188028317	<i>CALCRL</i>	0,85	0,0307	0,0116 ; 0,0497	<b>0,0015</b>
rs76122535	2:203664929	<i>ICA1L/NBEAL1</i>	0,97	-0,0074	0,0284 ; 0,0138	0,492
rs830179	3:183380035	<i>KLHL24</i>	0,95	0,0038	-0,0123 ; 0,0199	0,645
rs7733216	5:82857870	<i>VCAN</i>	0,96	0,0067	-0,0114 ; 0,0249	0,467
rs17148926	5:121510586	<i>LOC100505841</i>	0,94	-0,0041	-0,0233 ; 0,0152	0,677
rs275350	6:151016058	<i>PLEKHG1</i>	0,99	0,0131	-0,0020 ; 0,0282	0,089
rs17149723	8:8171487	<i>SGK223</i>	0,98	-0,0054	-0,0224 ; 0,0115	0,531
rs4630220	10:105459116	<i>SH3PXD2A</i>	0,89	-0,0102	-0,0270 ; 0,0066	0,233
rs10883926	10:105608838	<i>SH3PXD2A</i>	0,99	0,0016	-0,0135 ; 0,0168	0,835
rs12588905	14:52595066		0,99	-0,0010	-0,0161 ; 0,0140	0,889
rs1285841	14:91881387	<i>CCDC88C</i>	0,97	0,0057	-0,0009 ; 0,0205	0,452
rs55989340	14:100635222	<i>DEGS2</i>	0,98	-0,0081	-0,0248 ; 0,0084	0,335
rs1895029	16:51461433	<i>SALL1</i>	0,99	-0,0051	-0,0198 ; 0,0094	0,488
rs12921170	16:87227397	<i>C16orf95</i>	0,98	-0,0121	-0,0270 ; 0,0029	0,114
rs6503417	17:43144218	<i>NMT1/DCAKD</i>	0,99	-0,0078	-0,0234 ; 0,0077	0,324
rs34974290	17:73888354	<i>TRIM65</i>	0,99	-0,0001	-0,0185 ; 0,0181	0,983
rs429358	19:45411941	<i>APOE</i>	0,99	0,0089	-0,0140 ; 0,0320	0,444
rs5762197	22:27887471	<i>MN1</i>	0,97	-0,0208	-0,0364 ; -0,0053	<b>0,008</b>
<b>PSMD (<math>10^{-4}mm^2/s</math>) (ajusté sur âge, sexe, VIC*, cursus)</b>				0,074	-0,0064 ; 0,154	0,071
<b>PSMD (<math>10^{-4}mm^2/s</math>) (ajusté sur âge, sexe, VIC*, cursus)</b>						0,403
Q1				1		
Q2				0,023	-0,006 ; 0,053	0,120
Q3				0,017	-0,012 ; 0,047	0,242
Q4				0,033	0,003 ; 0,060	<b>0,031</b>

\* VIC : volume intracrânien ; \*\* PC 1-4 : 4 premières composantes principales ; TR : temps de réaction  $wGRS_{res}$  d'HSB correspond au score de risque génétique des HSB pondéré et redimensionné.

### 3. Analyses complémentaires

La première analyse complémentaire a été d'ajuster sur les facteurs de risques vasculaires, en plus des ajustements réalisés sur les modèles de l'analyse principale. Ces résultats sont présentés dans le **Tableau 5**.

Sur cet effectif réduit, nous avons constaté que le wGRSres d'HSB restait significativement associé au PSMD ( $p = 0,044$ ), avec un ajustement sur l'âge, le sexe, les quatre premières composantes principales, le volume intracrânien, la consommation de tabac, l'IMC et la pression artérielle systolique. L'association était plus significative que dans le modèle sans ajustement sur les facteurs de risque vasculaires dans ce sous-groupe de participants ayant bénéficié du check-up médical.

Le wGRSres d'HSB n'était pas significativement associé au temps de réaction du test Stroop dans le sous-échantillon de participants avec check-up médical, avec et sans ajustement sur les facteurs de risque vasculaires. De même, le PSMD n'était pas significativement associé au temps de réaction du test Stroop dans ce même sous-échantillon, avec et sans ajustement sur les facteurs de risque vasculaires.

*Tableau 5. Association du wGRSres d'HSB et PSMD (N = 854) ainsi que du wGRSres d'HSB et PSMD sur le temps de réaction du test de Stroop (N = 741) sans et avec un ajustement sur les facteurs de risques vasculaires. Cohorte i-Share. SEGWAY. Bordeaux, France. 2019.*

Variables	N	PSMD (/sec)			
		$\beta$	IC 95%	p	
<b>wGRSres d'HSB</b> (ajusté sur âge, sexe, VIC*, PC 1-4**)	854	0,0135	-0,0001 ; 0,0028	0,069	
<b>wGRSres d'HSB</b> (ajusté sur âge, sexe, VIC*, PC 1-4**, tabac, PAS <sup>§</sup> , IMC <sup>°</sup> )	854	0,0014	0,00003 ; 0,0029	<b>0,044</b>	
		Stroop (/sec)			
		N	$\beta$	IC 95%	p
<b>wGRSres d'HSB</b> (ajusté sur âge, sexe, PC 1-4**, cursus)	741	-0,0019	-0,0042 ; 0,0003	0,086	
<b>wGRSres d'HSB</b> (ajusté sur âge, sexe, PC 1-4**, cursus, tabac, PAS <sup>§</sup> , IMC <sup>°</sup> )	741	-0,0019	-0,0042 ; 0,0003	0,100	
<b>PSMD (10<sup>-4</sup>mm<sup>2</sup>/s)</b> (ajusté sur âge, sexe, VIC*, cursus)	741	0,0499	-0,0560 ; 0,1559	0,355	
<b>PSMD (10<sup>-4</sup>mm<sup>2</sup>/s)</b> (ajusté sur âge, sexe, VIC*, cursus, tabac, PAS <sup>§</sup> , IMC <sup>°</sup> )	741	0,0556	-0,0519 ; 0,1631	0,310	

\* VIC : volume intracrânien ; \*\* PC 1-4 : 4 premières composantes principales ; § PAS : Pression Artérielle Systolique ; ° IMC : Indice de Masse Corporelle ; TR : temps de réaction  
wGRS<sub>res</sub> d'HSB correspond au score de risque génétique des HSB pondéré et redimensionné.

Une analyse de sensibilité a été réalisée en comparant les participants avec et sans découvertes fortuites (DF) sur leur IRM (**Tableau 6**) et en excluant les participants avec DF (**Tableau 7**). Les participants avec DF n'étaient pas différents des participants sans DF au niveau des facteurs démographiques et des facteurs de risque vasculaires. En revanche, les participants avec DF étaient significativement différents des participants sans DF pour le PSMD avec des valeurs du PSMD significativement plus élevées ( $p = 0,017$ ).

Tableau 6. Description des caractéristiques des sujets sans découvertes fortuites ( $n=1656$ ) et des sujets avec découvertes fortuites ( $n=36$ ). Cohorte i-Share. SEGWAY. Bordeaux, France. 2019.

Variables	Population sans DF* (N = 1656)			Population avec DF* (N = 36)			p-value**
	NA	N/Moy	% / ET	NA	N/Moy	% / ET	
<b>Facteurs Démographiques</b>							
<b>Âge à l'IRM (années)</b>	0	22	2,2	0	21	2,5	0,688
<b>Sexe</b>	0			0			0,139
Homme		459	27,7		14	38,9	
Femme		1197	72,3		22	61,1	
<b>Type d'études</b>	4			0			0,528
Santé		906	54,8		22	61,1	
Sciences		176	10,7		3	8,3	
Gestion, économie, commerce, marketing, droit		136	8,2		5	13,9	
Arts, Multimédia, Sciences Sociales et de l'homme		286	17,3		4	11,1	
Ecole d'ingénieurs et Sciences de l'ingénieur		29	1,8		1	2,8	
Autres		119	7,2		1	2,8	
<b>Facteurs de risques vasculaires</b>							
<b>Tabac</b>							
Consommation de tabac	2			0			0,367
<i>Oui</i>		398	24,0		5	13,9	
<i>Oui, mais en train d'arrêter</i>		89	5,4		2	5,6	
<i>Non</i>		1167	70,5		29	80,6	
<b>Variables d'intérêts</b>							
<b>PSMD (<math>10^{-4} \text{ mm}^2/\text{s}</math>)</b>	0	1,53	0,13	0	1,59	0,12	<b>0,017</b>
<b>Stroop (secondes)</b>	402	1,26	0,19	7	1,27	0,19	0,762

N = effectif ; Moy = moyenne ; % = pourcentage ; ET = écart-type

\*DF = découvertes fortuites.

\*\*Des tests de  $\chi^2$  ou de Fisher exact ont été réalisés pour les variables qualitatives et des tests de Student pour les variables quantitatives.



L'analyse en excluant les participants ayant eu des découvertes fortuites sur leur IRM sont présentés dans le **Tableau 7**. Nous avons constaté que le wGRSres d'HSB restait significativement associé à une valeur plus élevée de PSMD chez les participants sans DF, après un ajustement sur l'âge, le sexe, les quatre premières composantes principales de stratification de la population et du type d'étude ( $p = 0,013$ , taille de l'effet similaire).

Tableau 7. Association du wGRSres d'HSB et PSMD ( $N = 1\ 652$ ) ainsi que du PSMD et du test de Stroop ( $N = 1\ 252$ ) des participants sans DF. Cohorte i-Share. SEGWAY. Bordeaux, France. 2019.

Variables	PSMD (/sec)			
	N	$\beta$	IC 95%	p
<b>wGRSres d'HSB</b> (ajusté sur âge, sexe, VIC*, PC 1-4**)	1 652	0,0013	0,0003 ; 0,0023	<b>0,013</b>
	Stroop (/sec)			
	N	$\beta$	IC 95%	p
<b>PSMD (<math>10^{-4}\text{mm}^2/\text{s}</math>)</b> (ajusté sur âge, sexe, VIC*, cursus)	1 252	0,0659	-0,0153 ; 0,1472	0,111

\* VIC : volume intracrânien ; \*\* PC 1-4 : 4 premières composantes principales ; TR : temps de réaction  
wGRSres d'HSB correspond au score de risque génétique des HSB pondéré et redimensionné.

## DISCUSSION

Cette étude exploratoire, menée au sein de la cohorte i-Share, avait pour objectif d'analyser l'association de déterminants génétiques de la charge en HSB avec le PSMD, un marqueur précoce d'intégrité de la microstructure de la substance blanche cérébrale en IRM-DTI, et l'impact de ces modifications structurelles précoces sur les performances cognitives chez le sujet jeune.

### I. Résumé des résultats

Nous avons constaté une majorité de femmes dans la population du projet SEGWAY (ratio Femmes/Hommes de 2,6), ce qui est plus important qu'à l'échelle de la population générale, mais habituel dans les études sur la santé basées sur le volontariat. On peut émettre l'hypothèse que les femmes s'intéressent plus à leur santé que les hommes et/ou sont plus disposées à contribuer à la recherche médicale. La moitié des participants au projet SEGWAY viennent d'un cursus du secteur de la santé. Une explication potentielle pourrait être que les étudiants de la filière santé sont plus sensibilisés à l'étude de leur santé du fait de leur spécialisation et affinité pour les questions de santé.

L'intégrité de la substance blanche a été évaluée par un nouveau marqueur IRM-DTI, le PSMD, qui était significativement associé au wGRSres d'HSB en tenant compte du sexe, de l'âge, du volume intracrânien et des composantes principales de la stratification de la population. Individuellement, 3 SNPs à risque d'HSB, près des gènes *KLHL24*, *VCAN*, *SH3PXD2A*, étaient significativement associés à un PSMD plus élevé. Nous avons ensuite exploré l'impact de ce wGRSres d'HSB et du PSMD sur la cognition. Nous avons constaté que le wGRSres n'était pas significativement associé à la vitesse de traitement de l'information chez le jeune adulte. En revanche, individuellement, 1 SNP à risque d'HSB (à proximité du gène *CALCRL*, différent des SNPs associés au PSMD) était significativement associé à un ralentissement de la vitesse de traitement de l'information au test de Stroop. Le PSMD pris en continu n'était pas significativement associé à la vitesse de réaction au test Stroop mais nous avons mis en évidence une tendance à un ralentissement. Par ailleurs, les participants dans le 4<sup>ème</sup> quartile de distribution du PSMD présentaient un temps de réaction du test de Stroop significativement plus long que les participants dans le 1<sup>er</sup> quartile, en tenant compte de l'âge, du sexe, du volume intracrânien et du type d'étude.

En analyse complémentaire, nous avons constaté que le wGRSres d'HSB restait significativement associé à un PSMD plus élevé après ajustement complémentaire sur les principaux facteurs de risque vasculaires (pression artérielle systolique, IMC et consommation de tabac) et après exclusion des participants avec une découverte fortuite en IRM.

## II. Lien avec littérature

Le PSMD est un marqueur très récent, ayant montré une pertinence dans la MPAC. Baykara et al ont en effet montré que le PSMD était un marqueur plus sensible pour la MPAC que les autres marqueurs en DTI mais aussi d'autres marqueurs caractéristiques de la MPAC sur des séquences morphologiques (HSB, lacunes, microsaignements). Le PSMD montrait l'association la plus forte avec la vitesse de traitement de l'information dans les analyses multivariées. Ceci était observé dans des populations de personnes atteintes de MPAC, de formes rares monogéniques et de formes communes, et en population générale âgée. Ce dernier avait mieux capturé la progression de la MPAC que les autres marqueurs d'imagerie, dans une analyse longitudinale. Par ailleurs, le PSMD est plus sensible au changement et plus reproductible que les autres marqueurs en DTI ((17) et données non publiées de Düring et al.). Depuis, deux études, ayant étudié le PSMD, ont été publiées. Ces études ont été réalisées sur des personnes plus âgées que celle de notre étude (âgées entre 40 et 65 ans). La première étude s'est intéressée à l'association entre le PSMD et les fonctions cognitives chez des sujets hospitalisés avec des altérations de la substance blanche, avec et sans déficience cognitive en comparaison avec des sujets sains (59). Les résultats montraient que les patients avec des altérations de la substance blanche avaient des valeurs de PSMD plus élevées et des moins bonnes performances cognitives par rapport aux sujets sains. De plus, le PSMD était fortement associé à la cognition globale et les fonctions exécutives chez les patients avec des altérations de la substance blanche. La seconde étude a évalué le PSMD chez des patients atteints de sclérose en plaque par rapport à des témoins sains et des patients atteints de CADASIL (la forme génétique rare la plus fréquente de la MPAC) (60). Les résultats de cette étude ont montré : dans les deux groupes, les valeurs de PSMD étaient supérieures à celles des sujets sains et étaient étroitement corrélés au volume de lésion de la substance blanche, suggérant une sensibilité de ce marqueur dans l'évaluation des altérations des tissus cérébraux. Le PSMD semble donc un nouveau marqueur IRM prometteur pour quantifier les lésions microscopiques de la substance blanche.

Les résultats de notre étude semblent indiquer que les SNPs impliqués dans les mécanismes de formation des HSB chez la personne âgée auraient déjà un effet à un âge jeune, renforçant l'hypothèse que l'on se situe sur le même phénomène biologique. On peut émettre l'hypothèse que les changements de la microstructure de la substance blanche commenceraient à un âge jeune. Ces changements sur la microstructure de la substance blanche s'intensifieraient avec le temps, provoquant des altérations de la substance blanche qui deviendraient visibles et observables plus tardivement sur les IRM conventionnelles sous la forme des HSB. Les marqueurs du DTI, notamment le PSMD, seraient plus sensibles à un

stade précoce pour détecter des altérations de la microstructure de la substance blanche à un âge jeune. Les changements de la microstructure de la substance blanche à un âge jeune, mesurés par les marqueurs du DTI, pourraient correspondre à une phase précoce de MPAC précédant le développement des HSB. Cette hypothèse est émise dans plusieurs études. Les résultats de l'étude de Pelletier et al suggèrent que les altérations microstructurales mesurées par le DTI et la présence d'HSB sont deux phénomènes interdépendants se produisant dans le même continuum, les modifications DTI étant la phase précoce, suivi par les HSB (39). Toutefois, à ce jour, un lien entre les HSB et les modifications microstructurelles de la substance blanche en DTI n'a jamais été montré à un âge aussi jeune que dans la cohorte i-Share. De plus, l'outil génétique que nous utilisons permet de mettre en évidence des voies biologiques communes entre les HSB et ces modifications précoces en DTI. Du fait de l'âge très jeune des participants d'i-Share, dont l'âge moyen correspond au pic de maturation de la substance blanche cérébrale, les associations observées pourraient refléter soit des manifestations précoces de MPAC, soit une corrélation génétique entre des marqueurs de MPAC chez la personne âgée et des marqueurs de maturation cérébrale chez l'adulte jeune. En effet, les marqueurs en DTI sont également utilisés pour mesurer la maturation cérébrale chez l'enfant et l'adulte jeune (61), bien que le PSMD spécifiquement n'ait pas encore été étudié dans ce contexte. Il est intéressant de noter que, malgré une très forte association du PSMD avec la pression artérielle systolique, et bien que 1 des 3 SNPs les plus significativement associés avec le PSMD soient aussi un SNP à risque de pression artérielle élevée (62), l'association du GRS d'HSB avec le PSMD était indépendante de la pression artérielle systolique. En outre, nous avons pu observer une tendance d'association entre un PSMD plus élevé et un ralentissement de la vitesse de traitement de l'information, cette association étant significative pour le quartile supérieur du PSMD. Bien que l'association observée soit moins forte que celle décrite en population générale âgée (Austrian Stroke Prevention Study) et chez des patients atteints de MPAC (monogénique ou complexe) (17), il est déjà remarquable qu'une association soit détectée chez l'adulte jeune de 22 ans en moyenne. Il pourra être intéressant d'explorer plus en détail l'association du PSMD avec les performances cognitives, notamment avec les autres mesures de fonctions exécutives (test de Stroop notamment), et d'étendre ces analyses aux autres marqueurs de microstructure cérébrale en IRM-DTI (FA, MD, RD, AD).

Dans cette étude, nous avons identifiés trois locus à risque de MPAC (associés au volume d'HSB chez la personne âgée) qui sont individuellement associés à un PSMD plus élevé après correction sur tests multiples. Le locus 3q27.1 (*rs830179*) se trouve proche du gène *KLHL24* qui code pour un récepteur de substrat d'ubiquitine ligase. Ce gène aurait un taux d'expression élevé dans plusieurs tissus cérébraux. La famille *KLHL* est impliquée dans

les processus d'ubiquitination mais les rôles spécifiques de chaque membre de la famille n'ont pas encore été élucidés (63). Le locus 5q14.2 (*rs7733216*) se trouve proche du gène *VCAN* qui appartient à la famille des protéoglycanes aggrécans / versicans (64). La protéine codée est un protéoglycane, composant majeur de la matrice extracellulaire, et est impliquée dans la prolifération, la migration et l'angiogenèse cellulaires et joue un rôle central dans la morphogenèse et le maintien des tissus. De façon intéressante, ce locus a été identifié dans une GWAS sur l'intégrité microstructurale de la substance blanche du cerveau évaluée par des marqueurs du DTI au sein de la UK Biobank, dont les participants étaient âgés de 40 à 69 ans (65). Le PSMD n'était pas analysé dans cette étude. On peut faire l'hypothèse que les variants génétiques du gène *VCAN* peuvent jouer un rôle dans les mécanismes sous-tendant l'intégrité microstructurale de la substance blanche. Le locus 10q24.33 (*rs10883926*) se trouve proche du gène *SH3PXD2A* qui encode pour la protéine 2A contenant le domaine PH et SH3, c'est une protéine adaptatrice impliquée dans la formation notamment de podosomes ainsi que dans la dégradation de la matrice extracellulaire. Ce gène a précédemment été identifié dans un précédent GWAS d'HSB en population générale, sur un plus petit effectif (52), ainsi que dans un GWAS sur les AVC, ayant identifié des loci pour les AVC, les AVC de type ischémique et pour les principaux sous-types d'AVC ischémiques (66). De plus, il a été rapporté comme étant impliqué dans la neurotoxicité des peptides bêta-amyloïdes dans la maladie d'Alzheimer (67).

Enfin, nous avons observé un SNP à risque d'HSB associé de façon significative à la vitesse de traitement de l'information au test de Stroop. Ce locus 2q32.1 (*rs62172472*) se trouve proche du gène *CALCRL* qui encode un récepteur de la calcitonine. Il a été décrit comme étant impliqué dans la pathogenèse des AVC dans une population japonaise, dans une étude ayant investigué la relation entre l'adrénomédulline et les AVC (68).

### III. Forces

Cette étude est, à notre connaissance, la première à s'intéresser à l'impact de variants génétiques à risque d'HSB sur un marqueur précoce d'intégrité de la substance blanche, le PSMD, et sur les fonctions cognitives chez l'adulte jeune. C'est également la première étude décrivant l'association du PSMD avec les performances cognitives chez l'adulte jeune (et, bien que cela ne constituait pas l'objectif de notre étude, avec les facteurs de risque vasculaires). Bien que de nouvelles cohortes sur le sujet jeune commencent à se mettre en place, peu d'études à large échelle sur la santé et le bien-être des jeunes adultes existent (plus de 20 000 étudiants à ce jour inclus dans I-Share). La base de données I-Share est d'une grande richesse, ayant récolté des informations sur beaucoup de thèmes différents, permettant de caractériser la population des étudiants et leur mode de vie sur de nombreux domaines avec des données génétiques et IRM sur un sous-échantillon, faisant de la cohorte i-Share, une

cohorte unique sur cette tranche d'âge. Il est à noter que les données génétiques et d'IRM sont à la pointe du progrès. La puce utilisée dans pour le génotypage de l'ADN est plus précise pouvant génotyper environ 1 millions de variants, l'imputation quant à elle a été réalisée avec le panel de référence HRC qui regroupe des données de séquençage de multiples cohortes, permettant une précision de l'imputation du génotype et le nombre de variants imputables. En outre, les IRM cérébrales des participants ont tous été réalisés selon un protocole standardisé, sur la même machine Siemens 3T Prisma et interprétées dans le même centre de lecture. Enfin, le projet SEGWAY possède un effectif important de données génétiques et d'IRM sur une tranche d'âge peu étudiée, permettant ainsi d'avoir une puissance statistique pour déceler des associations entre un marqueur IRM, une composante génétique et la cognition.

#### **IV. Limites et perspectives**

Notre étude était exploratoire, menée afin de mieux comprendre les mécanismes sous-jacents de la MPAC par l'étude d'un nouveau marqueur. Elle possède certaines limites méritant d'être discutées et d'être prises en considération dans l'interprétation des résultats.

Tout d'abord, la première limite, inhérente au design de l'étude, concerne la sélection de la population. Le recrutement des étudiants pour les projets Bio/MRi-Share est basé sur le volontariat et n'a été proposé qu'aux étudiants d'i-Share sur le site de de Bordeaux. Les sujets volontaires sont en général en meilleure santé, qui aurait pu engendrer un biais de sélection. Ces volontaires en meilleure santé pourraient sous-estimer les associations du fait des meilleures habitudes de vies et potentiellement des valeurs de PSMD moins pathologiques. De plus, le projet SEGWAY a une surreprésentation du sexe féminin (72,0 %). Toutes les analyses ont été ajustées sur ce facteur, cependant un effet résiduel pourrait persister même après ajustement. Comme extension de ce travail nous envisageons d'effectuer des analyses secondaires stratifiées sur le sexe.

Une autre limite de notre étude réside dans le recueil partiel à ce stade de données sur les facteurs de risque vasculaires, notamment les mesures anthropométriques et la pression artérielle. Les mesures anthropométriques prises lors du check-up médical ont été réalisées par un professionnel de santé donc moins limitées aux biais. La pression artérielle avec des mesures systoliques et diastoliques ont été répétées deux fois permettant de vérifier de possibles erreurs dans les mesures. Cependant, seule une partie des participants a à ce jour réalisé ce check-up, permettant d'avoir des données fiables sur ces facteurs. De plus, les autres facteurs de risques vasculaires qu'est la consommation de tabac a été recueillie dans le questionnaire d'inclusion à l'étude i-Share et donc est basée sur la déclaration de l'individu, pouvant engendrer un biais potentiel de prévarication. En effet, bien que les données de l'étude soient rendues anonymes, ce sont des questions sensibles où les participants peuvent diminuer leur consommation, sous estimant possiblement l'effet de ce facteur. Enfin, il n'est

pas à exclure qu'un biais de confusion résiduel persiste, malgré la prise en compte des facteurs de confusions dans nos différents modèles.

Du fait d'un manque de disponibilité des données cognitives complètes à ce jour (contrôle qualité en cours de finalisation), nous n'avons pu utiliser comme variable cognitive que la vitesse de traitement des informations de participants avec le temps de réaction pour le test de Stroop. Des analyses complémentaires seront nécessaires pour confirmer ou infirmer la tendance observée d'une association entre le PSMD et la vitesse de réaction au Stroop, avec les autres tests des fonctions exécutives et de la vitesse de traitement de l'information effectués dans i-Share, soit individuellement, soit en effectuant une analyse agrégée de ces différents tests et items de type *g factor*. En effet, ce *g factor* est une technique qui permet de combiner les performances d'un individu de plusieurs tests, évaluant les fonctions exécutives ou bien la vitesse de traitement de l'information, dans un index général qui s'appelle *g factor*. Une étude récente a montré que le *g factor* de la vitesse de traitement de l'information était associé à la structure de la substance blanche, mesuré par le marqueur du DTI la FA, chez la personne âgée (69).

Pour les analyses de sensibilité avec les participants ayant des DF, nous avons pris comme première approche d'exclure tous les participants avec une DF. Des analyses supplémentaires seront nécessaires, du fait que dans cette analyse, nous avons pu exclure des participants ayant déjà des altérations de la substance blanche liées à une forme précoce de MPAC. Une analyse plus fine de l'association avec le PSMD sera réalisée, en prenant en compte les différents types de découvertes fortuites. Ces analyses seront effectuées de nouveau en n'excluant pas les participants dont la découverte fortuite traduit potentiellement une forme débutante de MPAC, c'est-à-dire des HSB sans critères caractéristiques d'une maladie inflammatoire du système nerveux central.

Il est à noter que dans l'attente de la réception des génotypes pangénomiques sur GWAS qui ont été disponibles très tardivement sur ce stage, des premières analyses ont été réalisées sur le sous-échantillon avec séquençage pangénomique (non présentées dans ce mémoire). Un premier GRS contenant 17 SNPs avaient été réalisés (SNPs issus d'une première vague de méta-analyse de GWAS d'HSB). Une actualisation de la GWAS a permis d'identifier 21 SNPs, utilisés dans ce mémoire. Cependant, un problème technique du serveur où logent et sont analysées les données, est survenu ne nous permettant pas de mettre à jour les analyses sur les données de séquençages avec les 21 au lieu de 17 SNPs. Néanmoins, ces analyses sur les données de séquençages préliminaires avaient montré qu'un signal d'association entre les variants à risque d'HSB et le PSMD était déjà présent sur un échantillon plus réduit.

## CONCLUSION

Dans une population de 1 692 d'étudiants volontaires bordelais, âgés de 22 ans en moyenne (dont près de trois quarts de femmes), avec des données d'IRM cérébrale et génétiques à la pointe du progrès, nous avons montré que des variants génétiques à risque d'HSB chez la personne âgée étaient associés à un marqueur précoce d'intégrité de la substance blanche en IRM-DTI, le PSMD, très précocement. Par ailleurs, un variant génétique à risque d'HSB chez la personne âgée, et le quartile supérieur du PSMD étaient associés à de moins bonnes performances cognitives (vitesse de traitement de l'information) chez ces mêmes adultes jeunes. Ce projet permet de renforcer l'hypothèse que les modifications microstructurales de la substance blanche chez le sujet jeune pourraient correspondre à une phase précoce de MPAC, ou éventuellement refléter un processus de maturation cérébrale qui partagerait des locus génétiques à risque avec les HSB.



## BIBLIOGRAPHIE

1. Debette S, Markus HS. The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: systematic review and meta-analysis. *BMJ*. 2010 Jul 26;341(jul26 1):c3666–c3666.
2. L'Organisation Mondiale de la Santé (OMS). [Page consultée le 05/06/2019]. La démence. [en ligne] <https://www.who.int/fr/news-room/fact-sheets/detail/dementia>
3. Ballard C, Gauthier S, Corbett A, Brayne C, Aarsland D, Jones E. Alzheimer's disease. *The Lancet*. 2011 Mar;377(9770):1019–31.
4. Glahn DC, Kent JW, Sprooten E, Diego VP, Winkler AM, Curran JE, et al. Genetic basis of neurocognitive decline and reduced white-matter integrity in normal human brain aging. *Proc Natl Acad Sci*. 2013 Nov 19;110(47):19006–11.
5. Hunter DJ. Gene–environment interactions in human diseases. *Nat Rev Genet*. 2005 Apr;6(4):287–98.
6. Kochunov P, Williamson DE, Lancaster J, Fox P, Cornell J, Blangero J, et al. Fractional anisotropy of water diffusion in cerebral white matter across the lifespan. *Neurobiol Aging*. 2012 Jan;33(1):9–20.
7. Debette S, Schilling S, Duperron M-G, Larsson SC, Markus HS. Clinical Significance of Magnetic Resonance Imaging Markers of Vascular Brain Injury: A Systematic Review and Meta-analysis. *JAMA Neurol* [Internet]. 2018 Oct 22 [cited 2019 Jan 13]; Available from: <http://archneur.jamanetwork.com/article.aspx?doi=10.1001/jamaneurol.2018.3122>
8. Heijer T den, Geerlings MI, Hoebek FE, Hofman A, Koudstaal PJ, Breteler MMB. Use of Hippocampal and Amygdalar Volumes on Magnetic Resonance Imaging to Predict Dementia in Cognitively Intact Elderly People. *Arch Gen Psychiatry*. 2006 Jan 1;63(1):57–62.
9. Poirier J. [Page consultée le 05/05/2019] Le système nerveux, Central et périphérique : formation, fonction et rôle. [en ligne]. [https://www.arsep.org/library/media/other/recherche/systeme\\_nerveux\\_central.pdf](https://www.arsep.org/library/media/other/recherche/systeme_nerveux_central.pdf)
10. Le Bihan D. Looking into the functional architecture of the brain with diffusion MRI. *Nat Rev Neurosci*. 2003 Jun;4(6):469–80.
11. Einstein A. *Investigations on the Theory of the Brownian Movement*. Courier Corporation; 1956. 148 p.
12. Le Bihan D. [Page consultée le 15/05/2019]. Du mouvement Brownien aux images de la pensée : l'IRM de diffusion [en ligne]. <http://www.academie-medecine.fr/du-mouvement-brownien-aux-images-de-la-pensee-lirm-de-diffusion/>
13. Baumann PS, Hagmann P, Marquet P. De l'eau et des neurones... IRM de diffusion, cartographie cérébrale et brèves perspectives psychiatriques. *Inf Psychiatr*. 2010;Volume 86(6):493–504.
14. Guillevin R, Menuel C, Vallée J-N, Chiras J. Imagerie par tenseur de diffusion : principes et applications. *Neurologie.com*. 2010 Apr;2(4):100–4.

15. Le Bihan D, Mangin JF, Poupon C, Clark CA, Pappata S, Molko N, et al. Diffusion tensor imaging: concepts and applications. *J Magn Reson Imaging JMRI*. 2001 Apr;13(4):534–46.
16. Exarchos TP, Papadopoulos A, Fotiadis DI [Page consultée le 06/05/2019] Handbook of Research on Advanced Techniques in Diagnostic Imaging and Biomedical Applications [en ligne]. <http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-60566-314-2>
17. Baykara E, Gesierich B, Adam R, Tuladhar AM, Biesbroek JM, Koek HL, et al. A Novel Imaging Marker for Small Vessel Disease Based on Skeletonization of White Matter Tracts and Diffusion Histograms: Novel SVD Imaging Marker. *Ann Neurol*. 2016 Oct;80(4):581–92.
18. Bush WS, Moore JH. [Page consultée le 27/05/2019] Chapter 11: Genome-Wide Association Studies. [en ligne]. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3531285/>
19. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet*. 2018;50(9):1219–24.
20. Flechsig Of Leipsic P. Developmebtal (myelogenetic) localisation of the cerebral cortex in the human subject. *The Lancet*. 1901 Oct;158(4077):1027–30.
21. Small GW. Use of neuroimaging to detect early brain changes in people at genetic risk for Alzheimer's disease. *Adv Drug Deliv Rev*. 2002 Dec 7;54(12):1561–6.
22. Amieva H, Jacqmin-Gadda H, Orgogozo J-M, Le Carret N, Helmer C, Letenneur L, et al. The 9 year cognitive decline before dementia of the Alzheimer type: a prospective population-based study. *Brain*. 2005 May 1;128(5):1093–101.
23. Fornage M, Debette S, Bis JC, Schmidt H, Ikram MA, Dufouil C, et al. Genome-wide association studies of cerebral white matter lesion burden: the CHARGE Consortium. *Ann Neurol*. 2011 Jun;69(6):928–39.
24. Longstreth WT. Brain vascular disease overt and covert. *Stroke*. 2005 Oct;36(10):2062–3.
25. Pantoni L. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. *Lancet Neurol*. 2010 Jul;9(7):689–701.
26. Tan R, Traylor M, Rutten-Jacobs L, Markus H. New insights into mechanisms of small vessel disease stroke from genetics. *Clin Sci*. 2017 Apr 1;131(7):515–31.
27. Wang R, Laveskog A, Laukka EJ, Kalpouzos G, Bäckman L, Fratiglioni L, et al. MRI load of cerebral microvascular lesions and neurodegeneration, cognitive decline, and dementia: *Neurology*. 2018 Sep;1.
28. Cottier J-P, Edjlali M, Gaillard M-A, Domengie F, Aljishi A, Casals X, et al. Pathology of small cerebral arteries demonstrated by MRI: a marker of aging? *Gériatrie Psychol Neuropsychiatr Viellissement*. 2011 Dec;(4):465–475.
29. Cuadrado-Godia E, Dwivedi P, Sharma S, Ois Santiago A, Roquer Gonzalez J, Balcells M, et al. Cerebral Small Vessel Disease: A Review Focusing on Pathophysiology, Biomarkers, and Machine Learning Strategies. *J Stroke*. 2018 Sep 30;20(3):302–20.

30. INSERM. [Page consultée le 20/05/2019] Accident vasculaire cérébral (AVC). [en ligne]. <https://www.inserm.fr/information-en-sante/dossiers-information/accident-vasculaire-cerebral-avc>
31. Chauhan G, Adams HHH, Satizabal CL, Bis JC, Teumer A, Sargurupremraj M, et al. Genetic and lifestyle risk factors for MRI-defined brain infarcts in a population-based setting. *Neurology*. 2019 Jan 16;
32. Mishra A, Chauhan G, Violleau M-H, Vojinovic D, Jian X, Bis JC, et al. Association of variants in HTRA1 and NOTCH3 with MRI-defined extremes of cerebral small vessel disease in older subjects. *Brain J Neurol*. 2019 Apr 1;142(4):1009–23.
33. Tzourio C, Laurent S, Debette S. Is Hypertension Associated With an Accelerated Aging of the Brain? *Hypertension*. 2014 May;63(5):894–903.
34. Lawrence AJ, Patel B, Morris RG, MacKinnon AD, Rich PM, Barrick TR, et al. Mechanisms of Cognitive Impairment in Cerebral Small Vessel Disease: Multimodal MRI Results from the St George's Cognition and Neuroimaging in Stroke (SCANS) Study. Baron J-C, editor. *PLoS ONE*. 2013 Apr 22;8(4):e61014.
35. Cremers LGM, de Groot M, Hofman A, Krestin GP, van der Lugt A, Niessen WJ, et al. Altered tract-specific white matter microstructure is related to poorer cognitive performance: The Rotterdam Study. *Neurobiol Aging*. 2016 Mar;39:108–17.
36. Deslandre E, Lefebvre G, Girard C, Lemarchand M, Minouni A. [Page consultée le 01/06/2019]. Evaluation neuropsychologique - Les fonctions exécutives. [en ligne] <https://www.em-consulte.com/en/article/83794>
37. Alber J, Alladi S, Bae H-J, Barton DA, Beckett LA, Bell JM, et al. White matter hyperintensities in vascular contributions to cognitive impairment and dementia (VCID): Knowledge gaps and opportunities. *Alzheimers Dement Transl Res Clin Interv*. 2019 Apr 9;5:107–17.
38. Zeestraten EA, Lawrence AJ, Lambert C, Benjamin P, Brookes RL, Mackinnon AD, et al. Change in multimodal MRI markers predicts dementia risk in cerebral small vessel disease. *Neurology*. 2017 Oct 31;89(18):1869–76.
39. Pelletier A, Periot O, Dilharreguy B, Hiba B, Bordessoules M, Chanraud S, et al. Age-Related Modifications of Diffusion Tensor Imaging Parameters and White Matter Hyperintensities as Inter-Dependent Processes. *Front Aging Neurosci* 2016.
40. Lebel C, Treit S, Beaulieu C. A review of diffusion MRI of typical white matter development from early childhood to young adulthood. *NMR Biomed*. 2017 Sep 8;e3778.
41. van Norden AGW, de Laat KF, Fick I, van Uden IWM, van Oudheusden LJB, Gons RAR, et al. Diffusion tensor imaging of the hippocampus and verbal memory performance: The RUN DMC Study. *Hum Brain Mapp*. 2012 Mar;33(3):542–51.
42. den Heijer T, der Lijn F van, Vernooij MW, de Groot M, Koudstaal PJ, der Lugt A van, et al. Structural and diffusion MRI measures of the hippocampus and memory performance. *NeuroImage*. 2012 Dec;63(4):1782–9.
43. Maillard P, Seshadri S, Beiser A, Himali JJ, Au R, Fletcher E, et al. Effects of systolic blood pressure on white-matter integrity in young adults in the Framingham Heart Study: a cross-sectional study. *Lancet Neurol*. 2012 Dec;11(12):1039–47.

44. de Groot M, Verhaaren BFJ, de Boer R, Klein S, Hofman A, van der Lugt A, et al. Changes in Normal-Appearing White Matter Precede Development of White Matter Lesions. *Stroke*. 2013 Apr;44(4):1037–42.
45. de Groot M, Ikram MA, Akoudad S, Krestin GP, Hofman A, van der Lugt A, et al. Tract-specific white matter degeneration in aging: The Rotterdam Study. *Alzheimers Dement*. 2015 Mar;11(3):321–30.
46. Svärd D, Nilsson M, Lampinen B, Lätt J, Sundgren PC, Stomrud E, et al. The effect of white matter hyperintensities on statistical analysis of diffusion tensor imaging in cognitively healthy elderly and prodromal Alzheimer's disease. *PLoS ONE*. 2017.
47. Maillard P, Fletcher E, Lockhart SN, Roach AE, Reed B, Mungas D, et al. White Matter Hyperintensities and Their Penumbra Lie Along a Continuum of Injury in the Aging Brain. *Stroke*. 2014 Jun;45(6):1721–6.
48. Objectifs | i-Share: La plus grande étude jamais réalisée sur la santé des étudiants [Internet]. [cited 2018 Apr 25]. Available from: <http://www.i-share.fr/etude/objectifs>
49. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, et al. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*. 2006 Jul;31(4):1487–505.
50. Stroop JR. Studies of interference in serial verbal reactions. *J Exp Psychol*. 1935;18(6):643–62.
51. Jouvent E, Reyes S, De Guio F, Chabriat H. Reaction Time is a Marker of Early Cognitive and Behavioral Alterations in Pure Cerebral Small Vessel Disease. *J Alzheimers Dis JAD*. 2015;47(2):413–9.
52. Verhaaren BFJ, Debette S, Bis JC, Smith JA, Ikram MK, Adams HH, et al. Multiethnic Genome-Wide Association Study of Cerebral White Matter Hyperintensities on MRI. *Circ Cardiovasc Genet*. 2015 Apr;8(2):398–409.
53. Lin X, Song K, Lim N, Yuan X, Johnson T, Abderrahmani A, et al. Risk prediction of prevalent diabetes in a Swiss population using a weighted genetic score—the CoLaus Study. *Diabetologia*. 2009 Apr;52(4):600–8.
54. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006 Aug;38(8):904–9.
55. Jeerakathil T, Wolf PA, Beiser A, Massaro J, Seshadri S, D'Agostino RB, et al. Stroke risk profile predicts white matter hyperintensity volume: the Framingham Study. *Stroke*. 2004 Aug;35(8):1857–61.
56. Elias MF, Sullivan LM, D'Agostino RB, Elias PK, Beiser A, Au R, et al. Framingham stroke risk profile and lowered cognitive performance. *Stroke*. 2004 Feb;35(2):404–9.
57. Au R, Massaro JM, Wolf PA, Young ME, Beiser A, Seshadri S, et al. Association of white matter hyperintensity volume with decreased cognitive functioning: the Framingham Heart Study. *Arch Neurol*. 2006 Feb;63(2):246–50.
58. Uiterwijk R, Staals J, Huijts M, Leeuw PW de, Kroon AA, Oostenbrugge RJ van. Framingham Stroke Risk Profile is related to cerebral small vessel disease progression

- and lower cognitive performance in patients with hypertension. *J Clin Hypertens*. 2018;20(2):240–5.
59. Wei N, Deng Y, Yao L, Jia W, Wang J, Shi Q, et al. A Neuroimaging Marker Based on Diffusion Tensor Imaging and Cognitive Impairment Due to Cerebral White Matter Lesions. *Front Neurol*. 2019.
  60. Vinciguerra C, Giorgio A, Zhang J, Di Donato I, Stromillo ML, Tappa Brocci R, et al. Peak width of skeletonized mean diffusivity (PSMD) as marker of widespread white matter tissue damage in multiple sclerosis. *Mult Scler Relat Disord*. 2019 Jan 1;27:294–7.
  61. Bouziane C, Caan MWA, Tamminga HGH, Schrantee A, Bottelier MA, de Ruiter MB, et al. ADHD and maturation of brain white matter: A DTI study in medication naive children and adults. *NeuroImage Clin*. 2018;17:53–9.
  62. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat Genet*. 2018;50(10):1412–25.
  63. Dhanoa BS, Cogliati T, Satish AG, Bruford EA, Friedman JS. Update on the Kelch-like (KLHL) gene family. *Hum Genomics*. 2013 May 15;7(1):13.
  64. Wight TN. Versican: a versatile extracellular matrix proteoglycan in cell biology. *Curr Opin Cell Biol*. 2002 Oct;14(5):617–23.
  65. Rutten-Jacobs LCA, Tozer DJ, Duering M, Malik R, Dichgans M, Markus HS, et al. Genetic Study of White Matter Integrity in UK Biobank (N=8448) and the Overlap With Stroke, Depression, and Dementia. *Stroke*. 2018 Jun;49(6):1340–7.
  66. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet*. 2018 Apr;50(4):524–37.
  67. Harold D, Jehu L, Turic D, Hollingworth P, Moore P, Summerhayes P, et al. Interaction between the ADAM12 and SH3MD1 genes may confer susceptibility to late-onset Alzheimer's disease. *Am J Med Genet Part B Neuropsychiatr Genet Off Publ Int Soc Psychiatr Genet*. 2007 Jun 5;144B(4):448–52.
  68. Koyama T, Kuriyama N, Ozaki E, Matsui D, Watanabe I, Takeshita W, et al. Genetic Variants of RAMP2 and CLR are Associated with Stroke. *J Atheroscler Thromb*. 2017 Dec 1;24(12):1267–81.
  69. Kuznetsova KA, Maniega SM, Ritchie SJ, Cox SR, Storkey AJ, Starr JM, et al. Brain white matter structure and information processing speed in healthy older age. *Brain Struct Funct*. 2016 Jul;221(6):3223–35.

## ANNEXES

### ANNEXE 1 : Présentation de la structure d'accueil

Ce stage, réalisé dans le cadre du Master 2 Sciences, Technologies, Santé : mention Santé Publique – spécialité Épidémiologie, a été effectué au sein de l'équipe VINTAGE « Epidémiologie intégrative et génétique du vieillissement cérébral et vasculaire » du Centre de Recherche U1219.

L'Institut national de la santé et de la recherche médicale (INSERM) a vu le jour en 1964. Cet établissement public à caractère scientifique et technologique, placé sous la double tutelle de ministère de la Santé et du ministère de la Recherche, est dédié à la recherche biologique, médicale et à la santé humaine. Le centre de Recherche INSERM U1219 « Bordeaux population Health » (ex-U 897) est dirigé par le professeur Christophe TZOURIO, et est composé des 11 équipes labélisées et des 2 équipes émergentes. Leurs travaux de recherches couvrent un large champ de pathologies (neurologie, cancérologie, maladies infectieuses dont le VIH ...), d'expositions (facteurs nutritionnelles, génétiques ...), de méthodes (épidémiologie, biostatistiques ...) et de populations (jeunes adultes, personnes âgées ...). Ce stage a été effectué au sein de l'équipe VINTAGE, sous la responsabilité du Pr. Stéphanie Debette. Cette équipe a pour but d'améliorer la compréhension des déterminants environnementaux et génétiques des AVC et démence, avec un intérêt particulier pour l'identification et l'évaluation de nouveaux biomarqueurs (génétiques, en imagerie, etc.) des AVC et de la démence, avec l'objectif d'accélérer l'identification de nouvelles cibles thérapeutiques, de faciliter la détection d'individus à haut risque de développer ces maladies, et d'optimiser les stratégies de prévention. Cette s'équipe s'intéressent notamment à l'étude en imagerie des marqueurs d'altérations structurelles cérébrales, en particulier des marqueurs de lésions cérébrovasculaires, en imagerie par résonance magnétique car ceux-ci, pouvant être mesurés en population avant l'apparition de symptômes cliniques, représentent un puissant prédicteur de survenue d'AVC et de démence. Les travaux de recherche s'appuient sur des données d'études multicentriques, telles que des études en population, chez la personne âgées (cohorte 3C-Dijon) et chez les jeunes étudiants (cohorte i-Share), ou bien une cohorte multicentrique de personnes présentant des plaintes cognitives ou des déficits cognitifs (Memento), ou encore une étude multicentrique sur la dissections des artères cervicales, une des causes les plus fréquentes d'AVC (CADISP) du sujet jeune.

#### Sources :

*L'Inserm en un coup d'oeil [Internet]. Inserm. [cité 3 juin 2018]. Disponible sur : <https://www.inserm.fr/connaître-inserm/inserm-en-coup-oeil>*

*Le Centre | Bordeaux population Health [Internet]. Inserm. [cité 3 juin 2018]. Disponible sur : <https://www.bordeaux-population-health.center/le-centre/>*

## ANNEXE 2

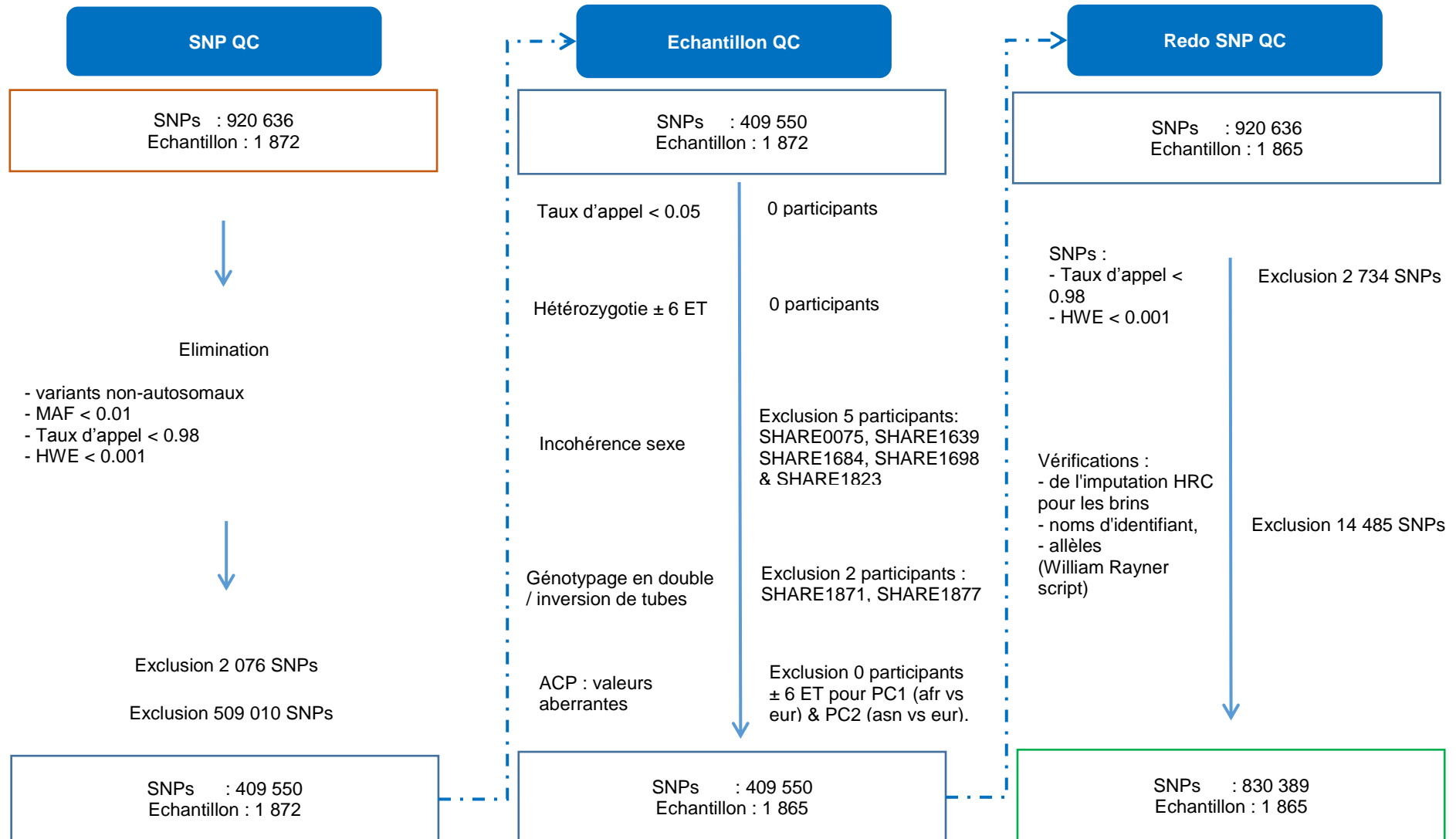


Figure. Procédure du contrôle qualité des données issues du génotypage pangénomique ADN. Cohorte i-Share. SEGWAY. Bordeaux, France. 2019.

ET\* : écart-type  
 QC : quality check (contrôle qualité)

### ANNEXE 3

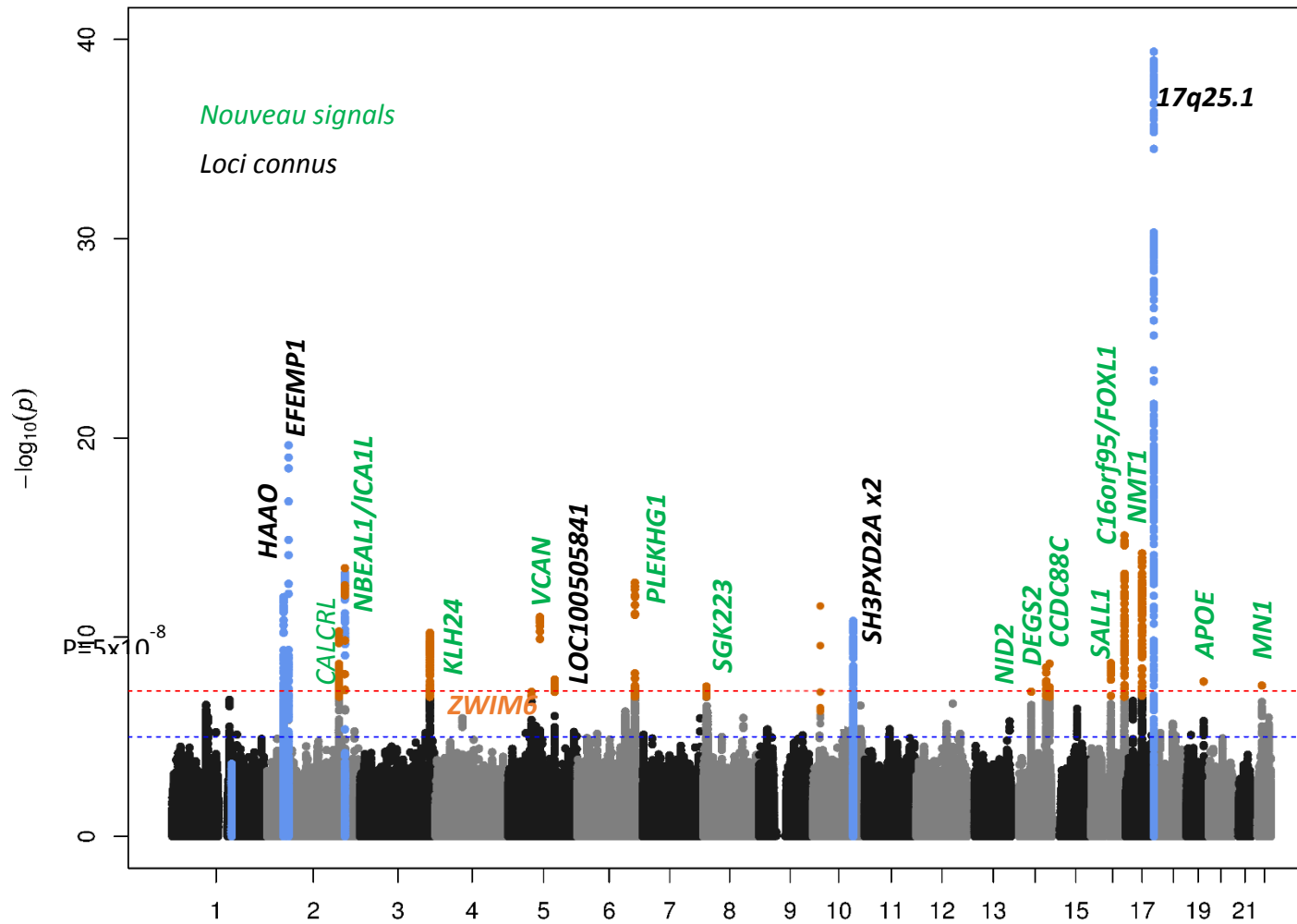


Figure. Manhattan plot du GWAS des HSB en cours de finalisation au sein du consortium CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology), coordonnée par le Dr. Murali Sargurupremraj et le Pr. Stéphanie Debette. 21 loci au seuil de significativité pangénomique ( $p < 10^{-8}$ ), donc 15 nouvellement découverts.



## ANNEXE 4

Tableau. Associations du temps de réaction au test de Stroop avec des facteurs sociodémographiques et de risques vasculaires de la MPAC. Cohorte i-Share, SEGWAY. Bordeaux, France. 2019.

Variables	Stroop (secondes)				Stroop (secondes) *			
	N	$\beta$	IC 95%	<i>p</i>	N	$\beta$	IC 95%	<i>p</i>
<b>PSMD</b> (/10 <sup>-4</sup> mm <sup>2</sup> /s)	1 283	0,070	-0,0067 ; 0,146	0,074	1 281	0,071	-0,008 ; 0,151	0,081
<b>Age</b> (/ 1 an)	1 283	0,003	-0,001 ; 0,008	0,11	1 281	-0,003 <sup>¶</sup>	-0,0006 ; 0,008	0,091
<b>Sexe</b> (F vs H)	1 283	0,002	-0,022 ; 0,026	0,876		0,003 <sup>§</sup>	-0,021 ; 0,027	0,809
<b>Type d'études</b>	1 283				1 281			
<i>Santé</i>		1				1		
<i>Sciences</i>		-0,017	-0,051 ; 0,017	0,329		-0,018	-0,052 ; 0,016	0,297
<i>Gestion, économie, commerce, marketing, communication, droit</i>		0,026	-0,012 ; 0,066	0,185		0,026	-0,013 ; 0,065	0,189
<i>Arts, Multimédia, Sciences Sociales et de l'homme</i>		-0,001	-0,029 ; 0,027	0,928		-0,004	-0,031 ; 0,025	0,819
<i>Ecole d'ingénieurs et Sciences de l'ingénieur</i>		0,067	-0,013 ; 0,149	0,103		0,067	-0,013 ; 0,150	0,097
<i>Autres</i>		0,009	-0,032 ; 0,050	0,674		0,008	-0,033 ; 0,050	0,685
<b>Consommation de tabac</b>	1 283				1 281			
<i>Non</i>		1				1		
<i>Oui</i>		-0,002	-0,027 ; 0,023	0,892		-0,001	-0,027 ; 0,024	0,903
<i>Oui, mais en train d'arrêter</i>		-0,035	-0,086 ; 0,014s	0,166		-0,033	-0,085 ; 0,017	0,197
<b>Pression Artérielle systolique</b> (/cmHg)	741	0,0004	-0,009 ; 0,010	0,923	741	-0,001	-0,023 ; 0,020	0,900
<b>IMC</b> (/5 kg/m <sup>2</sup> )	741	0,0003	-0,021 ; 0,021	0,973	741	-0,001	-0,012 ; 0,010	0,863

\* Ajusté sur le sexe, l'âge et le type de cursus

<sup>¶</sup> Ajusté sur le sexe et le type de cursus

<sup>§</sup> Ajusté sur l'âge et le type de cursus

## Résumé

**Introduction :** La maladie des petites artères cérébrales (MPAC) est associée à un risque accru d'AVC, de déclin cognitif et de démence. De nombreuses données suggèrent que les processus physiopathologiques conduisant à la démence débutent de nombreuses années avant la survenue de symptômes cliniques. Afin de mieux appréhender les mécanismes très précoces des modifications structurelles cérébrales amenant à une démence ou un déclin cognitif, nous nous intéressons à l'association de déterminants génétiques de la charge en HSB avec un marqueur précoce d'intégrité de la microstructure de la substance blanche cérébrale en IRM-DTI, le PSMD (peak-width of skeletonized mean diffusivity), et à l'impact de ces modifications structurelles précoces sur les performances cognitives dans la cohorte d'étudiants i-Share.

**Méthodes :** Environ 1 700 étudiants ont été inclus dans l'analyse. Nous avons réalisé plusieurs modèles linéaires multivariés pour étudier la relation entre un score de risque génétiques des variants associés à un seuil significatif pangénomique ( $p < 5 \times 10^{-8}$ ) aux HSB et un marqueur IRM, le PSMD, puis entre ce score de risques génétique et le PSMD sur le temps de réaction du test Stroop.

**Résultats :** le score de risque génétique des HSB était significativement associé au PSMD ( $\beta : 0,001$  ; IC95 % : 0,0003 ; 0,0024 ;  $p = 0,0087$ ) après ajustement. Bien que le score de risque des HSB et le PSMD n'étaient pas significativement associés à la vitesse de traitement de l'information chez le jeune adulte, une tendance d'association a été observé entre un PSMD plus élevé et un ralentissement de la vitesse de traitement de l'information, cette association étant significative pour le quartile supérieur du PSMD.

**Conclusion :** Les associations observées entre le PSMD, la cognition et le score de risque génétique des HSB apportent des arguments supplémentaires en faveur de l'hypothèse que les modifications microstructurales de la substance blanche chez le sujet jeune pourraient correspondre à une phase précoce de MPAC, ou éventuellement refléter un processus de maturation cérébrale qui partagerait des locus génétiques à risque avec les HSB.

**Mots clés :** Maladies des petites artères cérébrales, PSMD, cognition, jeunes adultes, SEGWAY, microstructure de la substance blanche.

## Abstract

**Context:** Small vessels disease (SVD) is an important risk factor for stroke, cognitive impairment and dementia. Epidemiological studies suggest that pathological processes leading to dementia start many years before the onset of clinical symptoms. In order to improve our understanding of the underlying the mechanisms and time-courses of structural brain changes throughout a lifetime, we are interested in the association of genetic determinants of white matter hyperintensities with an early marker of integrity of the white matter microstructure in MRI-DTI, peak-width of skeletonized mean diffusivity (PSMD), and the impact of these early structural changes on cognitive performance in the i-Share study, cohort of students.

**Methods:** We included ~ 1,700 individuals, from i-Share cohort. Linear multivariable regressions have been used to study, firstly, the association between genetic risk score (GRS) of variants associated with white matter hyperintensities (WMH) at a genome-wide significance threshold ( $p < 5 \times 10^{-8}$ ) and a new MRI marker, PSMD. And secondly, between GRS and PSMD on processing speed information of Stroop test.

**Results:** after adjustment, GRS of WMH was significantly associated with PSMD ( $\beta: 0.001$ , 95% CI: 0.0003, 0.0024,  $p = 0.0087$ ). Although GRS of WMH and PSMD was not significantly associated with the processing speed in young adults, an association trend was observed between higher PSMD and slower processing speed. This association was significant for the upper quartile of the PSMD.

**Conclusion:** The associations between PSMD, cognition function and GRS of WMH in this study, add insights in favor of the hypothesis on microstructural changes of the white matter in young subjects could be an early phase of SVD, or reflecting a process of brain maturation that would share genetic loci at risk with WMH; still additional studies are needed to confirm these exploratory findings.

**Key words:** Small vessels disease, PSMD, cognition, young adults, white matter microstructure, SEGWAY.