



Segmentation automatique des anomalies de la substance blanche du sujet âgé

Thomas Samaille

► To cite this version:

Thomas Samaille. Segmentation automatique des anomalies de la substance blanche du sujet âgé. Neurosciences. Paris 6, 2013. Français. NNT: . tel-01565592

HAL Id: tel-01565592

<https://inria.hal.science/tel-01565592>

Submitted on 19 Jul 2017

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.



THÈSE DE DOCTORAT DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

Spécialité

Neurosciences

École Doctorale Cerveau, Cognition, Comportement (ED3C)

Présentée par

Thomas SAMAILLE

Segmentation automatique des anomalies de la substance blanche du sujet âgé.

soutenue le 4 juin 2013

Membres du jury

Pr. Bernard MAZOYER	Rapporteur
Dr. Michel DOJAT	Rapporteur
Dr. Christian BARILLOT	Examineur
Dr. Eric JOUVENT	Examineur
Pr. Claude MARSAULT	Examineur
Dr. Olivier COLLIOT	Membre invité
Dr. Marie CHUPIN	Co-encadrante
Pr. Didier DORMONT	Directeur de thèse

Remerciements

Je tiens tout d'abord à exprimer mes remerciements à l'ensemble de mes encadrants Marie Chupin, Didier Dormont et Olivier Colliot. Un immense merci à Marie pour sa gentillesse, sa rigueur et son souci du détail. Malgré tes nombreux "petits" et tes innombrables sollicitations, tu as toujours su te montrer disponible tout en me faisant suffisamment confiance pour expérimenter par moi-même. Merci à Didier pour son apport clinique aux travaux effectués. Merci, enfin, à Olivier pour sa capacité de synthèse et à remettre en perspective les développements effectués. Je mesure toute la chance d'avoir eu des encadrants aussi attentifs au bien-être de leurs étudiants.

Mes remerciements vont également à l'ensemble des membres du jury. Je souhaite tout particulièrement remercier Michel Dojat et Bernard Mazoyer pour avoir accepté d'être rapporteurs de mon travail de thèse et pour leurs commentaires instructifs. Merci également à Eric Jouvent et Christian Barillot pour leur présence dans ce jury et au professeur Claude Marsault pour en avoir assuré la présidence.

Mes remerciements vont ensuite bien évidemment à l'ensemble des membres du laboratoire que j'ai eu la chance de croiser tout au long de ces années. Les noms d'équipe ont évolué mais la bonne humeur et la sympathie sont, elles, demeurées constantes.

Merci aux différents chercheurs, ingénieurs, administratifs et cliniciens, participant au succès de l'équipe : Laurent Hugueville, Denis Schwartz, Jean-Didier Le Maréchal, Stanley Durrleman, Joan Glaunès, Mario Chavez, Jacques Martinerie, Corinne Omer, Christine Delmaire, Charlotte Rosso, Dominique Hasboun pour ses exposés fascinants et tout particulièrement à Lionel Thivard, pour son optimisme à toute épreuve.

Merci à mes prédécesseurs et anciens doctorants : Guillaume A. pour son enseignement du moonwalk, Guillaume D. pour ses présentations holistiques, Yohan pour sa présence rassurante ; à Thibaud (futur Marc Lesggy), José-Luis, Patricia et Catalina, entre autres pour la visite de l'île de Batz ; à Emilie pour les chorégraphies de Claude François ; à Rémi pour les jeux de mots plus ou moins réussis. Merci également aux plus jeunes, que j'espère avoir accueillis de manière aussi agréable que je l'ai moi-même été : merci à Ludovic pour mon apprentissage d'encadrant, à Claire pour ses jus citronnés, à Nicolas, Takoua, Pietro, Alexandre, Souheilya, Josefin, Ayna, Fanny, Pauline, Lorraine. Les remerciements poussant

à l'établissement de catégories, je remercie également tous ceux étiquetés "CATI" : Ali, Chabha, Xavier, Joanne, Hugo, Thomas et Catherine, qui vous tiendra informés de la prochaine t-conf. Merci également à Priska, Typhaine, Morgane, Sophie, Fabrice & co pour les enschufla, dame, sombrero et autres setienta complicado, qui ont égayé les mercredi soirs.

Je tiens également à remercier mes amis proches, Miguel, Micka, Jojo, Lolo et François pour les verres au Lucernaire, les calculs de barycentre et les nombreux débats, posés mais toujours passionnants (et merci à ceux qui ont pu se libérer de leurs lourdes obligations professionnelles pour assister à ma soutenance). Merci également à Joseph et Claire pour les discussions, les déménagements de minuit et les parties de tennis.

Enfin, je souhaiterais remercier l'ensemble de ma famille : mes parents pour leur soutien et la confiance qu'ils me témoignent, mes frères et soeur (Laurent pour les robustes références gaulliennes) et mes grands-parents. Et, bien évidemment, Larissa, pour sa patience et sa présence tout au long de ces années.

Table des matières

Remerciements	iii
Table des matières	v
Introduction	1
1 Contexte clinique : vieillissement et démences	5
1.1 La démence, un enjeu majeur de santé public au XXI ^e siècle	5
1.2 Diagnostic des démences	6
1.2.1 Définition de la démence	6
1.2.2 La démence de type Alzheimer	7
1.2.3 La démence vasculaire	9
1.3 Mécanismes pathologiques	10
1.3.1 Lésions de type Alzheimer	10
1.3.2 Lésions vasculaires	13
1.3.3 Des mécanismes non exclusifs	16
1.4 Apport de l'imagerie par résonance magnétique pour le diagnostic des démences	17
1.4.1 IRM cliniques usuelles	18
1.4.2 IRM et maladie d'Alzheimer	19
1.4.3 IRM et démence vasculaire ischémique sous-corticale	20
1.5 Objectif de la thèse	23
2 Quantification des hyperintensités de la substance blanche	25
2.1 Principales études de cohortes	25
2.2 Echelles visuelles et cotations	27
2.3 Segmentation manuelle	28
2.4 Segmentation automatique	28
2.4.1 Méthodes non supervisées	30

2.4.2	Méthodes supervisées	34
2.4.3	Comparaison des méthodes	37
2.5	Discussion	38
3	Méthode de segmentation des hyperintensités de la substance blanche (WHASA)	41
3.1	Schéma général de la chaîne de traitements	42
3.2	Prétraitements	42
3.2.1	Segmentation unifiée de la T1	44
3.2.2	Recalage de l'image T1 sur l'image FLAIR	45
3.2.3	Correction de biais de l'image FLAIR	46
3.2.4	Masque de la région intracrânienne	47
3.2.5	Bilan des prétraitements	47
3.3	Segmentation de l'image <i>mFLAIR</i>	47
3.3.1	Diffusion non linéaire	48
3.3.2	Paramètre de contraste	52
3.3.3	Ligne de partage des eaux	54
3.3.4	Critère d'arrêt	56
3.3.5	Fusion de régions adjacentes	59
3.3.6	Segmentation détaillée de l'image <i>mFLAIR</i>	59
3.3.7	Bilan de la segmentation de l'image <i>mFLAIR</i>	59
3.4	Sélection des régions sur $LPE_{grossière}$	60
3.4.1	Critère d'intensité	61
3.4.2	Critère de localisation	62
3.4.3	Elimination de régions corticales	64
3.4.4	Bilan de la sélection des régions	65
3.5	Etape de croissance	65
3.5.1	Insuffisances de la segmentation HSB_{init}	65
3.5.2	Dilatation géodésique	67
3.5.3	Ensemble marqueur pour la dilatation géodésique	68
3.5.4	Obtention de HSB_{final}	69
3.6	Conclusion	70
4	Evaluation de la méthode de segmentation	73
4.1	Données	74
4.1.1	Jeu de données <i>Hippocampe</i>	74
4.1.2	Jeu de données <i>Cadasil</i>	75

4.2	Référence manuelle	75
4.3	Outils d'évaluation	77
4.3.1	Coefficient de corrélation intraclasse (ICC)	77
4.3.2	Régression lineaire	77
4.3.3	Graphe de Bland-Altman	78
4.3.4	Validation spatiale	80
4.4	Résultats	82
4.4.1	ICC et régression lineaire	82
4.4.2	Graphes de Bland-Altman	86
4.4.3	Validation spatiale	88
4.5	Discussion	89
4.5.1	Sujets ayant un mauvais accord spatial	89
4.5.2	Sujets ayant un mauvais accord volumique	89
4.5.3	Interêt de l'étape de croissance	91
4.5.4	Sensibilité à l'estimation du paramètre de contraste	95
4.5.5	Sensibilité au critère d'intensité	95
4.6	Conclusion	97
5	Comparaison aux autres méthodes de la littérature	99
5.1	Choix des méthodes	100
5.1.1	Méthodes non supervisées	100
5.1.2	Méthodes supervisées	101
5.2	Implémentation des méthodes non supervisées	101
5.2.1	Freesurfer	101
5.2.2	Seuillage	102
5.2.3	SIPS	103
5.3	Implémentation des méthodes supervisées	103
5.3.1	Définition de l'ensemble d'apprentissage	104
5.3.2	Définition du vecteur de caractéristiques	105
5.3.3	Classification par k plus proches voisins (kNN)	108
5.3.4	Classification par SVM	109
5.4	Résultats	113
5.4.1	Comparaison avec les méthodes non supervisées	113
5.4.2	Comparaison avec les méthodes supervisées	118
5.5	Discussion	131
5.5.1	Comparaison avec les méthodes non supervisées	131
5.5.2	Comparaison avec les méthodes supervisées	133

5.5.3	Temps de calcul	134
5.6	Conclusion	135
6	Application sur d'autres jeux de données	137
6.1	Collaboration avec le CHRU de Lille	137
6.1.1	Description des données	138
6.1.2	Segmentation des HSB par WHASA	139
6.1.3	Corrélation avec les échelles visuelles	139
6.1.4	Relation avec l'âge	142
6.2	Mise en place au sein du Centre pour l'Acquisition et le Traitement des Images (CATI)	144
6.2.1	Cohorte MEMENTO	144
6.2.2	Centre pour l'Acquisition et le Traitement des Images (CATI)	146
6.2.3	Protocole d'acquisition	146
6.2.4	Implémentation dans la suite logicielle Brainvisa	148
6.2.5	Validation de la méthode WHASA sur la cohorte MEMENTO	151
6.3	Conclusion	154
	Conclusion	157
A	Critères diagnostiques des démences	163
A.1	Critères diagnostiques de la démence de type Alzheimer, DSM-IV-TR	163
A.2	Critères diagnostiques de la maladie d'Alzheimer, NINCDS- ADRDA	165
A.3	Critères de démence vasculaire, DSM-IV-TR	167
A.4	Critères diagnostiques de la démence vasculaire probable ou possible, NINDS-AIREN	169
B	Notions de base sur l'imagerie par résonance magnétique (IRM)	171
B.1	Origine du signal	171
B.2	Pondération des images	174
	Table des figures	178
	Liste des tableaux	183

Bibliographie	185
----------------------	------------

Introduction

Avec le vieillissement de la population, la multiplication des cas de démence est appelée à devenir un enjeu majeur de santé publique au XXI^{ème} siècle. La plus fréquente et la plus connue des pathologies occasionnant la démence chez le sujet âgé est la Maladie d'Alzheimer (MA). Elle est suivie des démences vasculaires.

Les étiologies de ces deux pathologies sont distinctes : la maladie d'Alzheimer est une maladie neurodégénérative caractérisée par l'accumulation anormale de protéines dans le tissu cérébral ; les démences vasculaires sont provoquées par des atteintes des vaisseaux sanguins, qui altèrent à différents degrés l'irrigation du cerveau. Néanmoins, l'âge est un facteur important de ces deux pathologies. Leur coexistence est donc fréquemment observée chez les personnes de plus de 70 ans. La question se pose alors de savoir quels peuvent être les interactions entre ces pathologies et leur impact sur la survenue de la démence.

* *
*

Les progrès de l'imagerie par résonance magnétique (IRM) permettent de visualiser *in vivo* de nombreuses structures auparavant inaccessibles. Le développement de méthodes d'analyse permettant d'exploiter ces nouvelles informations fait actuellement l'objet de nombreuses recherches. L'objectif est de parvenir à identifier à partir de ces images des caractéristiques traduisant un mécanisme pathologique sous-jacent. De manière similaire au dosage d'une protéine dans le sang, la mesure précise et reproductible de ces caractéristiques permettrait de contribuer au diagnostic ou de prédire l'évolution probable d'une maladie. On parle alors de biomarqueurs d'imagerie.

Dans le cas de la MA, par exemple, la perte neuronale causée par la maladie serait à l'origine d'une atrophie de certaines structures cérébrales.

La quantification du volume de l'hippocampe, structure sous-corticale touchée dès les stades précoces, ou l'estimation de l'épaisseur corticale permettent alors, dans une certaine mesure, d'évaluer l'avancée de la MA. De nombreuses méthodes ont ainsi été développées pour effectuer de telles analyses. Dans le cas des démences vasculaires, en revanche, peu de méthodes de référence existent actuellement.

* *
*

L'IRM permet de détecter trois principaux types d'anomalies vasculaires, relativement fréquentes chez les personnes âgées. Les lacunes sont des zones de nécrose du tissu cérébral, dont le diamètre est généralement compris entre 3 et 15 mm ; les microsaignements désignent des hémorragies périvasculaires, de diamètre inférieur à 10mm ; les hyperintensités de la substance blanche (HSB) regroupent des atteintes diffuses de la substance blanche.

La prévalence des HSB est particulièrement élevée, puisqu'elle est de près de 95% dans une population générale de plus de 65 ans. Les études histologiques ont montré que les HSB étaient très probablement liées à l'artériolosclérose des petites artères perforantes du cerveau. Les atteintes tissulaires associées regroupent entre autres la démyélinisation, la perte axonale et la gliose.

La quantification des HSB est traditionnellement effectuée par l'intermédiaire d'échelles visuelles utilisées par des neuroradiologues. Cette approche est limitée par son caractère subjectif et le manque de précision, particulièrement dans le cadre d'un suivi longitudinal. Une méthode plus complète reposerait sur la segmentation manuelle des HSB, qui pose également des problèmes de subjectivité, et qui peut difficilement être appliquée pour l'étude de grandes cohortes, regroupant plusieurs milliers de sujets. Pour répondre à cette problématique, plusieurs méthodes de segmentation automatique des HSB ont alors été proposées. Elles ont cependant la plupart du temps été développées pour une étude spécifique et leurs performances sur de nouvelles données ont rarement été évaluées. Il existe en effet de multiples sources de variations dans les conditions d'acquisition des images IRM, qui peuvent perturber les méthodes de segmentation automatiques. Bien que de nombreuses procédures de standar-

disation des acquisitions soient actuellement mises en place, la variabilité résiduelle demeure néanmoins importante.

* *
*

Cette thèse vise à proposer une méthode automatique robuste de segmentation des HSB, qui puisse être utilisée de manière fiable sur un grand nombre d'images cliniques dans un environnement multicentrique.

Nous introduisons une nouvelle méthode, dénommée WHASA (*White matter Hyperintensities Automatic Segmentation Algorithm*), permettant la segmentation des HSB à partir d'une image T1 et d'une image FLAIR. Afin de pouvoir traiter de manière robuste les images provenant de différents centres, WHASA s'appuie sur le contraste des HSB avec les tissus environnants plutôt que sur les intensités. L'image T1 est utilisée pour caractériser, pour chaque sujet, le contraste moyen entre la substance grise et la substance blanche. Cette information est ensuite utilisée pour paramétrer un filtrage par diffusion non linéaire sur l'image FLAIR, renforçant le contraste des HSB. L'utilisation couplée de ce filtrage avec une segmentation par ligne de partage des eaux, permet d'obtenir deux représentations de l'image FLAIR à différents niveaux de détails. L'utilisation de trois critères simples (intensité, localisation dans la substance blanche, non connexité à l'interface substance grise/liquide cérebrospinal) sur le niveau de détails le plus faible fournit une initialisation des HSB. Cette initialisation est finalement affinée par reconstruction géodésique en utilisant le niveau de détail plus fin.

La validation de la méthode proposée est effectuée sur un ensemble de 67 sujets présentant une grande variabilité dans les conditions d'acquisition et de charge lésionnelle. Nous avons également réimplémenté plusieurs méthodes de segmentation présentées dans la littérature et comparé les résultats obtenus sur ces mêmes données.

Enfin, nous montrons que WHASA peut être appliqué de manière directe sur de nouvelles données. Tout d'abord, les résultats de la segmentation automatique sont comparés à l'utilisation d'une échelle visuelle pour 59 sujets dans le cadre d'une collaboration avec le Centre Hospitalier Régional Universitaire de Lille. WHASA a également été intégré comme module dans la plateforme logicielle Brainvisa et est utilisé sur des sujets

de la phase pilote (87 sujets) et de la phase de déploiement (47 sujets) d'une importante étude multicentrique, la cohorte MEMENTO.

* *

*

Le manuscrit est organisé de la manière suivante. Le **chapitre 1** présente le contexte clinique général et décrit notamment les différentes altérations structurelles cérébrales caractéristiques de la MA et des démences vasculaires. Le **chapitre 2** constitue une revue de littérature des approches existantes consacrées à la quantification des HSB. Le **chapitre 3** détaille la nouvelle méthode WHASA que nous avons proposée. L'évaluation quantitative et détaillée de WHASA sur 67 sujets est décrite dans le **chapitre 4**. Le **chapitre 5** est consacré à la comparaison des résultats obtenus par WHASA et ceux obtenus par d'autres méthodes proposées dans la littérature sur ces mêmes données. Enfin, le **chapitre 6** présente l'application de WHASA sur de nouvelles données.

Chapitre 1

Contexte clinique : vieillissement et démences

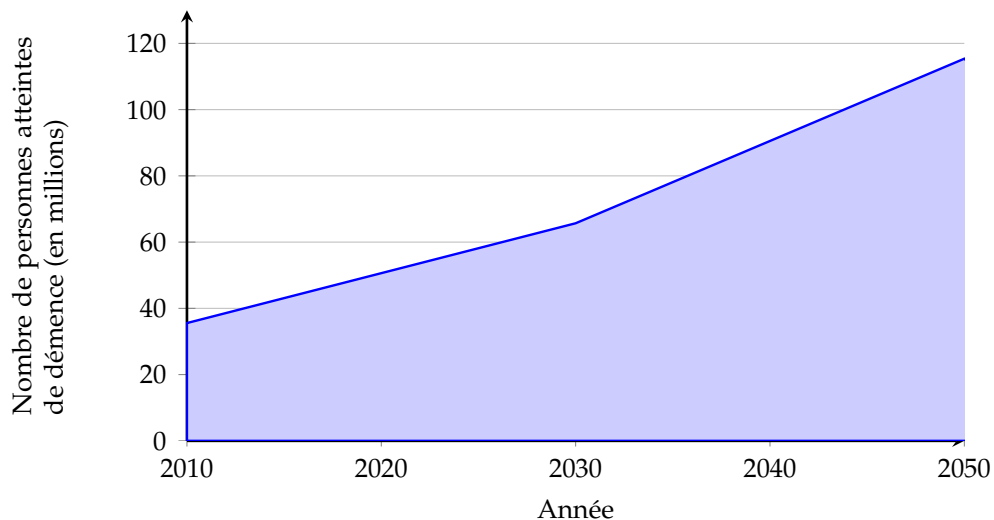


FIGURE 1.1 – Estimation du nombre de personnes dans le monde atteintes de démence entre 2010 et 2050 (source : [World Health Organization & Alzheimer’s Disease International, 2012]).

1.1 La démence, un enjeu majeur de santé public au XXI^e siècle

En 2012, l’Organisation Mondiale de la Santé (OMS) a publié, conjointement avec l’organisation Alzheimer’s Disease International, un rapport intitulé *Dementia : a public health priority* dans lequel elle insiste sur la nécessité d’adapter les systèmes de santé à l’explosion des cas de démences qui

se produira dans les années et décennies à venir. En 2011, le nombre d'individus atteints de démence s'élève ainsi à 35,6 millions dans le monde. Ce chiffre est appelé à croître de manière importante et devrait doubler environ tous les 20 ans, atteignant 65,7 millions en 2030 et 115,4 millions en 2050 (figure 1.1).

Cette augmentation est essentiellement liée au vieillissement de la population. En effet, la démence touche principalement les personnes de plus de 65 ans, seuls 2 à 10% des cas de démences commençant avant cet âge. Après 65 ans, la prévalence¹ des démences croît exponentiellement avec l'âge, en doublant à peu près tous les 5 ans.

Face à la multiplication attendue des cas de démence dans les prochaines décennies et au coût important que ces pathologies engendrent (estimé à 604 milliards de dollars en 2010), l'OMS appelle à une prise de conscience globale de ce problème et au développement de politiques de santé publique permettant la meilleure prise en charge possible. La question du diagnostic des démences est d'une importance particulière pour cet objectif.

1.2 Diagnostic des démences

1.2.1 Définition de la démence

Alors que le terme de démence fait référence dans le langage courant à un comportement aberrant ou irraisonné, synonyme de folie, sa signification dans un contexte médical est cependant toute autre. L'OMS définit ainsi la démence (ou syndrome démentiel) comme "l'altération acquise progressive de la mémoire et d'au moins une autre fonction supérieure, suffisamment marquée pour handicaper les activités de la vie quotidienne, apparue depuis au moins six mois".

L'établissement du diagnostic de syndrome démentiel exige donc :

- un trouble de la mémoire ;
- au moins un autre trouble intellectuel. Il peut s'agir d'un trouble du langage (aphasie), d'un trouble du mouvement (apraxie), d'un trouble du jugement ou du raisonnement ;
- une perte d'autonomie ;

1. La prévalence désigne la proportion d'individus atteints d'une affection dans une population à un moment donnée.

- l'absence de confusion mentale ou de dépression pouvant à elles seules expliquer le tableau clinique.

La démence est ainsi un terme purement descriptif désignant un ensemble de troubles cognitifs menant à une perte d'autonomie, fût-elle légère. Elle recouvre un grand nombre de pathologies différentes, qui sont différenciées par les mécanismes pathologiques sous-jacents ou les zones cérébrales atteintes. Parmi les principaux sous-types de démence, on peut citer la maladie d'Alzheimer (MA), les démences vasculaires, la démence à corps de Lewy (DCL), la démence fronto-temporale (DFT) ou encore la maladie de Parkinson avec démence.

La figure 1.2 illustre la répartition des différents diagnostics de démence en fonction de l'âge. Alors que les démences fronto-temporales peuvent représenter une proportion importante des cas de démence chez les personnes jeunes, les démences affectant les personnes de plus de 60 ans sont pour une très large majorité celles de type Alzheimer (entre 60 et 70%), suivies des démences vasculaires (15 à 20 %).

Il faut néanmoins préciser que le diagnostic précis des sous-types de démence est loin d'être évident. Nous détaillons dans les paragraphes suivants les critères diagnostiques pour les deux types principaux de démence, à savoir la démence de type Alzheimer et la démence vasculaire.

1.2.2 La démence de type Alzheimer

Historiquement, la démence de type Alzheimer a été décrite pour la première fois par Aloïs Alzheimer à propos du cas de Auguste Deter, une patiente de 51 ans admise à l'Hôpital de Francfort le 26 novembre 1901 et présentant des symptômes de démence. Après sa mort le 8 avril 1906, Aloïs Alzheimer obtint l'autorisation d'effectuer une autopsie et décrit deux anomalies du tissu cérébral : les plaques amyloïdes et les dégénérescences neurofibrillaires (DNF) (voir section 1.3). L'association de ces deux anomalies et leur extension dans certaines zones cérébrales est ainsi devenu le critère de diagnostic anatomo-pathologique de la maladie d'Alzheimer.

Cependant, un tel diagnostic ne peut être porté que post-mortem. Afin de pouvoir traiter et prendre en charge ces patients, différents critères diagnostiques ont depuis été proposés pour caractériser la maladie d'Alzheimer sans avoir recours à l'autopsie. Les plus connus et les plus

1. CONTEXTE CLINIQUE : VIEILLISEMENT ET DÉMENCES

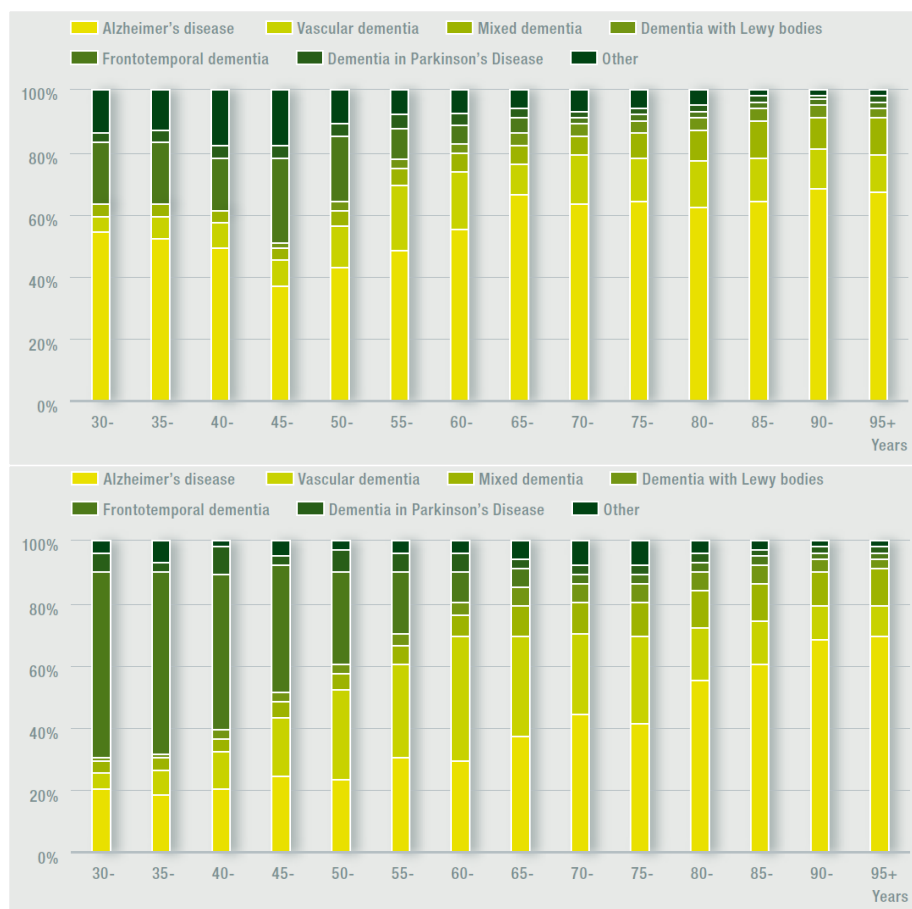


FIGURE 1.2 – Répartition des diagnostics de sous-types de démence en fonction de l'âge (Source : [World Health Organization & Alzheimer's Disease International, 2012]).

Haut : Femmes, **Bas :** Hommes.

couramment utilisés aujourd'hui sont ceux :

- du *Diagnostic and Statistical Manual of Mental Disorders (DSM)*. La version actuelle est une révision de la 4ème édition, dont l'abréviation est *DSM-IV-TR*.
- du *National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA)*.

Il ne s'agit bien évidemment pas ici de détailler de manière exhaustive les différents consensus diagnostiques, ni de discuter de leur sensibilité et spécificité. Le tableau 1.1 présente une synthèse très générale des critères diagnostiques de maladie d'Alzheimer. Les critères détaillés du DSM-IV-TR et du NINCDS-ADRDA sont donnés dans l'annexe A.

Au cours de la dernière décennie, les nombreux progrès dans la com-

Démence de type alzheimer		
Possible	Probable	Certaine
• Syndrome démentiel • Absence d'autre pathologie pouvant expliquer la démence	• Syndrome démentiel • Altérations progressives de la mémoire • Âge de début entre 40 et 90 ans, souvent après 65 ans • Absence d'autre pathologie pouvant expliquer la démence	• Maladie d'Alzheimer probable • Preuve histologique de plaques amyloïdes et DNF

TABLEAU 1.1 – Synthèse des critères diagnostiques de maladie d'Alzheimer.

préhension de la MA ont conduit à proposer de nouveaux critères, notamment par la prise en compte de marqueurs biologiques ou de données d'imagerie [Dubois et al., 2007, 2010; McKhann et al., 2011; Jack et al., 2011]. Ces nouveaux critères attendent néanmoins toujours confirmation par corrélation avec des données postmortem.

1.2.3 La démence vasculaire

A la différence de la maladie d'Alzheimer, la démence vasculaire n'est pas associée à un type particulier de lésions histologiques mais plutôt à un ensemble d'atteintes provoquées par des anomalies affectant les vaisseaux sanguins.

Le diagnostic de démence vasculaire repose donc, pour sa part, sur l'existence d'un syndrome démentiel, d'une maladie cérébro-vasculaire et d'une relation causale entre ces deux affections. Là encore, différents critères diagnostiques ont été développés en fonction de la finalité clinique ou de recherche, dont une synthèse est présentée dans le tableau 1.2. Les critères plus complets développés par le *National Institute of Neurological Disorders and Stroke* et l'*Association Internationale pour la Recherche et l'Enseignement en Neurosciences* (NINDS-AIREN) [Román et al., 1993] ou ceux du DSM-IV-TR sont donnés dans l'annexe A.

CADASIL Une partie des résultats présentés dans cette thèse ont été obtenus sur des patients atteints de CADASIL (*Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoaraiosis*), un sous-type

Démence vasculaire		
Possible	Probable	Certaine
• Syndrome démentiel • Signes focaux	• Syndrome démentiel • Signes focaux et confirmation en imagerie de lésion vasculaire • Relation temporelle	• Syndrome démentiel • Preuve histologique de lésion vasculaire • Absence de plaques amyloïdes et DNF

TABLEAU 1.2 – Synthèse des critères diagnostiques de démence vasculaire.

de démence vasculaire particulièrement étudié . Il s’agit en effet de la forme génétique la plus courante de démence vasculaire, causée par une mutation du gène *Notch3* sur le chromosome 19. La détection de la mutation permet un diagnostic précis du vivant des malades et constitue donc un modèle de choix pour l’étude des mécanismes et de l’évolution des démences vasculaires. [Tournier-Lasserre et al., 1993 ; Boussier & Tournier-Lasserre, 1994 ; Chabriat et al., 2009]

Les critères diagnostics présentés ci-dessus visent à séparer deux processus pathologiques différents pouvant mener tous deux à la démence chez les personnes âgées, qui devraient donc appeler des réponses thérapeutiques différentes. Nous allons présenter ici brièvement les mécanismes pathologiques liées au vieillissement qui sont à l’oeuvre dans les démences de maladie d’Alzheimer et les démences vasculaires.

1.3 Mécanismes pathologiques

Parmi les principes altérations du tissu cérébral qui apparaissent fréquemment chez les personnes âgées se trouvent les plaques amyloïdes, les dégénérescences neurofibrillaires, qui sont reliées à la maladie d’Alzheimer, et les lésions vasculaires.

1.3.1 Lésions de type Alzheimer

1.3.1.1 Plaques amyloïdes

Les plaques amyloïdes sont des accumulations de protéines situées dans le milieu extracellulaire (figure 1.3). Identifiées en microscopie dès

la fin du XIXe siècle, Divry [1927] montre qu'elles sont marquées électivement par le colorant Rouge Congo, de la même manière que certaines substances amylacées², ce qui leur donnera le nom de plaques amyloïdes.

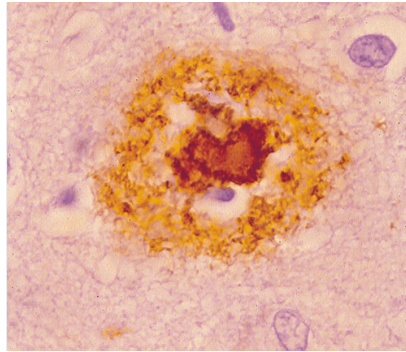


FIGURE 1.3 – Plaque amyloïde (Source : Dessi et al. [1997]).
Le marquage en marron du peptide $A\beta$ est obtenue à l'aide d'un anticorps (grandissement $\times 330$).

Cinquante ans plus tard, Glenner & Wong [1984] montrent que le principal constituant des ces agrégats est un polypeptide composé de 40 ou 42 acides aminés, appelé peptide $A\beta$. Il sera montré par la suite que ce peptide est le produit de la dégradation d'une protéine transmembranaire, la *protéine précurseur de l'amyloïde* (APP pour *Amyloid protein precursor*).

1.3.1.2 Protéines Tau et dégénérescences neurofibrillaires

Alors que les plaques amyloïdes s'accumulent dans le milieu extracellulaire, la dégénérescence neurofibrillaire (DNF) se produit dans le corps des neurones, où l'on observe l'accumulation de petites fibres, les fibrilles (figure 1.4). Kidd [1963], grâce à la microscopie électronique, montre que ces fibrilles sont constituées de filaments enroulés en forme d'hélice (PHF, pour *paired helicoïdal filaments*). Il faudra attendre le milieu des années 80 pour que le constituant principal de ces filaments soit identifié, à l'aide de techniques d'immunologie, comme étant des protéines tau [Brion et al., 1985].

Les protéines tau sont connues pour interagir en temps normal avec les microtubules du cytosquelette des neurones, qui contribuent à la forme et aux propriétés mécaniques du neurone (figure 1.5). L'hyperphosphorylation de la protéine tau bloquerait cette fonction habituelle et entrainerait la mort cellulaire.

2. Se dit des composés ayant la composition de l'amidon

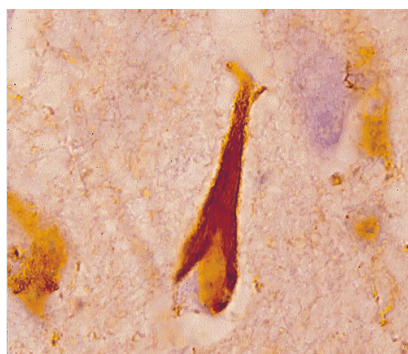


FIGURE 1.4 – Dégénérescence neurofibrillaire dans un neurone du cortex. La protéine tau est marquée par anticorps, apparaissant en marron. (grandissement $\times 330$, source : [Dessi et al., 1997])

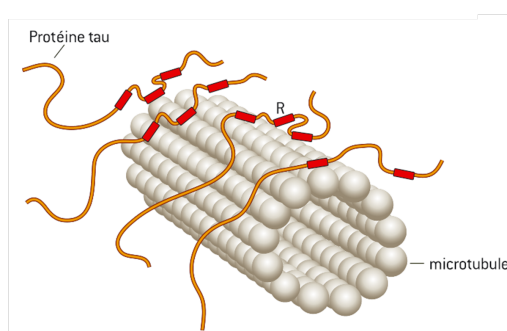


FIGURE 1.5 – Illustration du rôle physiologique de la protéine tau. Les microtubules forment en quelque sorte le squelette du neurone et les protéines tau participent à la stabilisation de ce squelette (source : Delacourte [2003]).

1.3.1.3 Evolution des lésions de type Alzheimer

La simple présence des plaques amyloïdes et des DNF ne suffit néanmoins pas à expliquer l'apparition de la maladie d'Alzheimer. Plus précisément, c'est leur accumulation et leur extension cérébrale qui semblent le mieux expliquer l'évolution des symptômes cliniques.

Braak & Braak [1995] ont été les premiers à identifier le schéma d'évolution caractéristique de la maladie d'Alzheimer et à proposer différents stades de gravité en fonction des régions atteintes. Une classification similaire, mais plus détaillée, a également été proposée par André Delacourte concernant l'évolution de la pathologie tau, qui corrèle plus précisément avec les troubles cliniques [Delacourte et al., 1999] (figure 1.6).

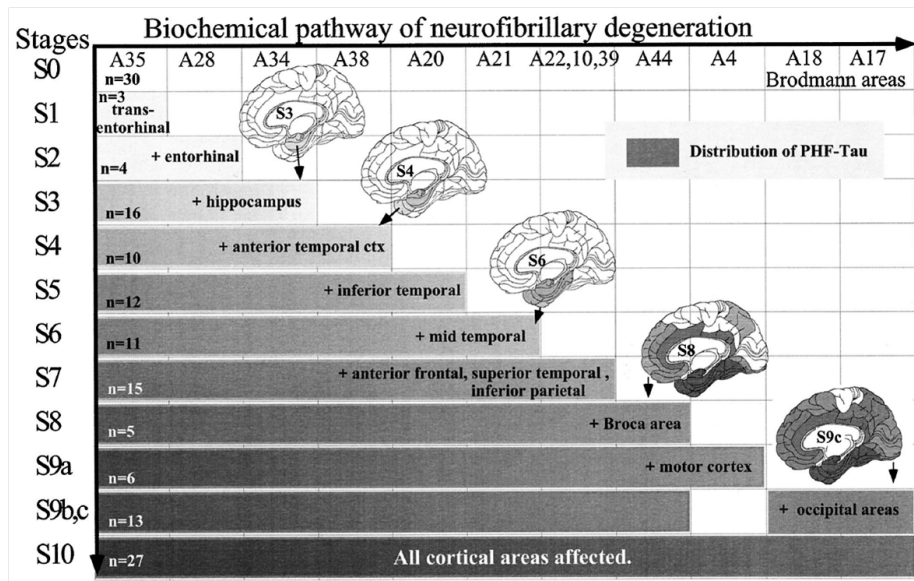


FIGURE 1.6 – Evolution de la pathologie tau en 10 étapes selon les régions cérébrales touchées (Source : [Delacourte et al., 1999]).

1.3.2 Lésions vasculaires

Les lésions vasculaires sont causées par des anomalies des vaisseaux sanguins cérébraux. L'irrigation sanguine du cerveau est assurée par un système complexe qui fait intervenir des vaisseaux de différentes tailles dont certains assurent une redondance dans les territoires pris en charge (figure 1.7). En fonction des vaisseaux touchés, les atteintes du tissu cérébral seront différentes ainsi que les conséquences cliniques.

1.3.2.1 Différentes atteintes des vaisseaux

Il existe globalement trois grandes catégories d'atteintes des vaisseaux cérébraux [Grinberg & Thal, 2010].

L'**athérosclérose** affecte principalement les grandes et moyennes artères. L'épaississement de la paroi interne des vaisseaux, l'*intima*, et l'accumulation de lipides entraînent la formation d'une plaque athérosclérotique. Cette plaque peut se rompre, entraînant une réaction de coagulation au niveau de la rupture. L'agrégation des plaquettes, composants sanguins importants dans le phénomène de coagulation, entraîne la formation d'un caillot sanguin, ou *thrombus*. Dans les cas asymptomatiques, ce thrombus s'intègre dans la plaque d'athérome, contribuant à sa stabilisation. Si le thrombus devient trop important, il peut entraîner une occlusion du

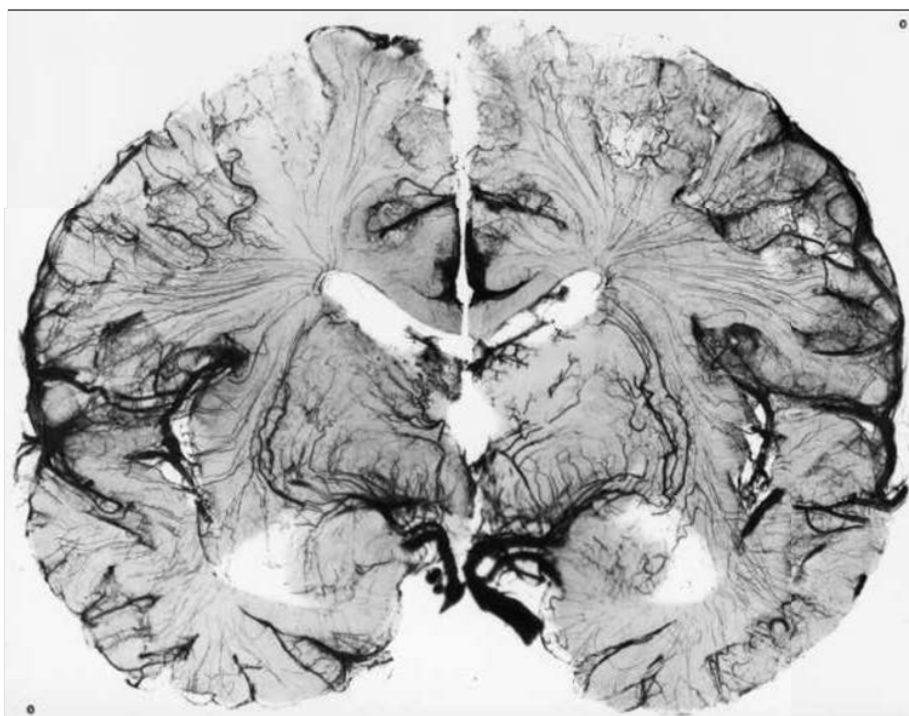


FIGURE 1.7 – Préparation anatomique mettant en évidence les grandes, moyennes et petites artères cérébrales (Source : [Marc & Dubas, 2005])

vaisseau, réduisant le flux sanguin (phénomène de *thrombose*). Enfin, le thrombus peut également se détacher et être relâché dans la circulation sanguine, où il risque alors d'obstruer une autre artère.

La **maladie des petits vaisseaux** (*small vessel disease*) regroupe un ensemble d'altérations affectant les vaisseaux de plus petit diamètre (entre 40 et 800 μm). Ces altérations, parmi lesquelles l'*artériolosclérose* (figure 1.8), consistent essentiellement en un épaississement de la paroi des vaisseaux, entraînant un rétrécissement de l'ouverture du vaisseau. Les vaisseaux localisés dans la substance blanche sont fréquemment touchés.

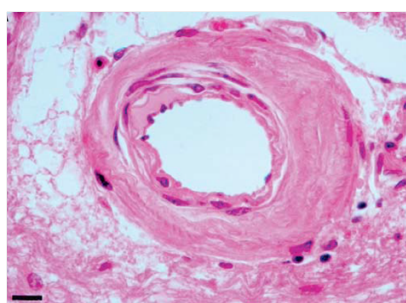


FIGURE 1.8 – Artériolosclérose, caractérisée par un épaississement concentrique de la paroi du vaisseau. (source : [Lammie, 2002])

L'**angiopathie amyloïde cérébrale** (figure 1.9) désigne le dépôt de peptides $A\beta$ dans la paroi de vaisseaux cérébraux de petits et moyens calibres. De manière générale, les vaisseaux de la substance blanche et de la moëlle sont épargnés.

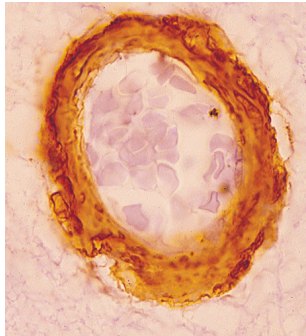


FIGURE 1.9 – Angiopathie amyloïde affectant la paroi d'un vaisseau. Le marquage en marron du peptide $A\beta$ est obtenue à l'aide d'un anticorps. (grandissement $\times 330$, source : [Dessi et al., 1997])

1.3.2.2 Différentes lésions du tissu cérébral

Les **infarctus** sont des zones de nécrose causées par l'insuffisance de l'afflux sanguin (ou *ischémie*). Ils sont souvent caractérisés selon leur taille. Les infarctus plus grands que 15 mm^3 sont la plupart du temps causés par des embolies causés par l'athérosclérose. Entre 5 et 15 mm^3 , les infarctus sont alors appelés *lacunes* et sont associés à la maladie des petits vaisseaux mais pas à l'angiopathie amyloïde. Ils sont le plus souvent situés dans la substance blanche et les structures sous-corticales. Enfin, des *microinfarctus*, inférieurs à 5 mm^3 peuvent également se produire.

Une deuxième catégorie de lésions concerne les hémorragies. On les distingue là encore en fonction de leur taille. Les **hémorragies cérébrales** mesurent plus de 10 mm de diamètre et peuvent être consécutives, par exemple, à une rupture d'anévrisme. Les **microsaignements** mesurent généralement moins de 10 mm de diamètre. Ils sont particulièrement reliés à l'angiopathie amyloïde, bien qu'ils soient également associés à l'hypertension artérielle [Cordonnier et al., 2007].

Enfin, on peut observer des **atteintes de la substance blanche** regroupant différentes altérations telles que la démyélinisation, la perte axonale et la gliose, altérations qui sont liées à la maladie des petits vaisseaux.

1.3.2.3 Différents types de démences vasculaires.

Ces différentes lésions vasculaires affectent différemment les vaisseaux en fonction de leur localisation. Il est ainsi possible de classer les démences vasculaires en trois sous-types (figure 1.10) [Thal et al., 2012].

La **démence par infarctus multiples** est causée par la multiplication d'infarctus de grande taille. La **démence avec infarctus stratégique** est due à un seul infarctus, localisé dans une zone stratégique dont l'atteinte seule peut expliquer la démence, telle que l'hippocampe ou le thalamus. Enfin, la **démence vasculaire ischémique sous-corticale** est causée par l'atteinte générale de la substance blanche, fréquemment associée à la présence, également sous-corticale, de lacunes.

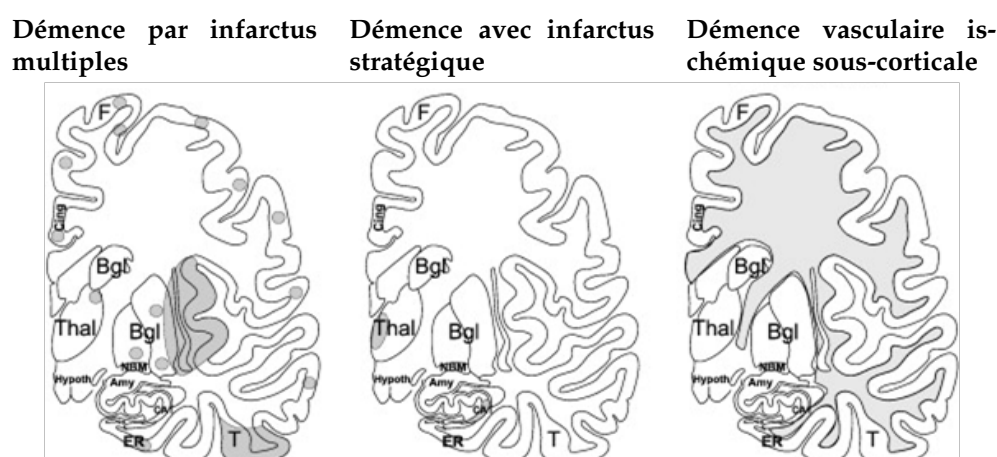


FIGURE 1.10 – Représentation schématique des trois principaux sous types de démence vasculaire (Source : [Thal et al., 2012]).

Abbreviations : Amy = amygdale, Bgl = ganglions de la base, CA1 = 1^{er} champ de la corne d'Ammon de l'hippocampe, Cing = gyrus cingulaire, ER = cortex entorhinal, F = cortex frontal, Hypoth = hypothalamus, NBM = noyau basal de Meynert, T = cortex temporal, Thal = thalamus.

1.3.3 Des mécanismes non exclusifs

Il semblerait donc y avoir deux mécanismes pathologiques distincts : l'accumulation de protéines dans le milieu extracellulaire et dans les corps des neurones résultant d'un dysfonctionnement cellulaire d'une part, la dégradation des parois des vaisseaux sanguins menant à des ischémies, des démyélinisations ou des inflammations d'autre part. L'association des plaques amyloïdes et des DNF marquerait ainsi la MA alors que

1.4. Apport de l'imagerie par résonance magnétique pour le diagnostic des démences

l'accumulation des lésions vasculaires serait la marque d'une démence vasculaire.

Cette dichotomie séduisante a néanmoins été mise à mal par plusieurs études anatomopathologiques menées sur des populations générales âgées [Brayne et al., 2001 ; Snowden et al., 1997 ; Petrovitch et al., 2005 ; Esiri et al., 1999]. Dans toutes ces études, l'autopsie de plusieurs centaines de personnes décédées après 70 ans a démontré la prévalence importante à la fois des lésions de type Alzheimer et des anomalies vasculaires. En particulier, l'étude du *Medical Research Council* [Brayne et al., 2001] a montré que des sujets non-déments pouvaient présenter fréquemment des densités de DNF et de plaques amyloïdes similaires à celles de sujets déments. Cette étude a également mis en évidence la très grande prévalence des lésions vasculaires dans cette population âgée (environ 75%). Snowden et al. [1997] ; Petrovitch et al. [2005] ; Esiri et al. [1999] ont, quant à eux, tous mis en évidence l'importance des lésions cérébrovasculaires dans la sévérité des symptômes cliniques, particulièrement associés à des charges légères ou modérées des lésions de la maladie d'Alzheimer.

Cependant, puisqu'elles sont réalisées postmortem, ces études ne permettent d'analyser ni l'évolution ni les interactions possibles existant entre les différents types de lésions. Elles ne permettent pas non plus l'évaluation de nouveaux traitements et sont de plus souvent limitées à un nombre assez restreint de sujets. L'imagerie, au contraire, bien que moins détaillée et précise, peut être réalisée de manière non invasive sur un grand nombre d'individus. Au cours des dernières décennies, l'étude du cerveau a particulièrement profité des progrès accomplis dans le domaine de l'imagerie par résonance magnétique (IRM).

1.4 Apport de l'imagerie par résonance magnétique pour le diagnostic des démences

Ainsi que nous l'avons évoqué lors de la présentation des critères diagnostiques de démence (tableau 1.2), l'imagerie peut jouer un rôle important dans la détermination des causes de démences, notamment vasculaires. Son rôle est également de plus en plus important dans le diagnostic de la MA.

Parmi les différentes modalités d'imagerie, l'imagerie par résonance magnétique (IRM) est actuellement la plus sensible pour l'étude du cerveau. Les principes physiques de l'IRM sont exposés dans l'annexe B. Nous nous limiterons ici à présenter les principales séquences IRM utilisées, puis indiquons leur rôle dans le diagnostic de la MA et de la démence vasculaire ischémique sous-corticale. Nous nous restreindrons à ces deux types de démence, qui coexistent fréquemment et dont le diagnostic est plus difficile, l'apparition des symptômes se faisant de manière lente et insidieuse.

1.4.1 IRM cliniques usuelles

En s'appuyant sur les principes présentés dans l'annexe B, quatre types de séquences IRM ont été développées et sont particulièrement utilisées pour l'imagerie du cerveau (figure 1.11).

Les **images pondérées en T1** ou images T1 sont aussi parfois appelées IRM anatomique. L'aspect observé est en effet similaire à ce qui est visualisé dans les coupes anatomiques obtenues par dissection. La substance grise apparaît grise, la substance blanche est plus claire et le liquide cébrospinal est sombre. Le mécanisme du contraste T1 permet également d'obtenir plus facilement une image 3D quasiment isotrope. Les images pondérées en T1 permettent de délimiter clairement les structures du cerveau, telles que le cortex.

Les **images pondérées en T2** ont un contraste inversé par rapport à celui des images T1. Le LCS apparaît blanc, la substance grise est plus claire que la substance blanche. Puisque les liquides sont mis en évidence, les images T2 sont particulièrement utiles pour visualiser les processus d'inflammation ou les œdèmes.

Il est possible d'obtenir des images n'ayant ni de pondération T1 ni de pondération T2. Il suffit pour cela d'utiliser un TR long, qui entraînera la repousse complète du vecteur d'aimantation, et un TE court, afin de mesurer directement la valeur du vecteur résultant avant que ne se produise la relaxation T2. Dans ce cas, la valeur du signal indique alors la densité de protons, c'est pourquoi on parle d'**images en densité de proton (DP)**.

Les **images FLAIR** (*FLuid Attenuated Inversion Recovery*) utilisent le principe d'inversion-récupération avec un TI adapté pour annuler le signal provenant du LCS. La pondération est ensuite équivalente à celle d'une T2.

1.4. Apport de l'imagerie par résonance magnétique pour le diagnostic des démences

Ces images permettent donc de visualiser les mêmes phénomènes que les images T2 mais l'annulation du signal du liquide permet de s'affranchir de l'effet de volume partiel lié à l'hypersignal du liquide cérébrospinal, particulièrement autour des ventricules ou au niveau du cortex cérébral.

Enfin, les **images T2*** sont les images où la correction des inhomogénéités de champ magnétique n'a pas été effectuée. Elles sont donc très sensibles aux inhomogénéités locales. Alors que ce phénomène est généralement indésirable, il peut néanmoins être utile pour détecter certaines anomalies, telles que les hémorragies.

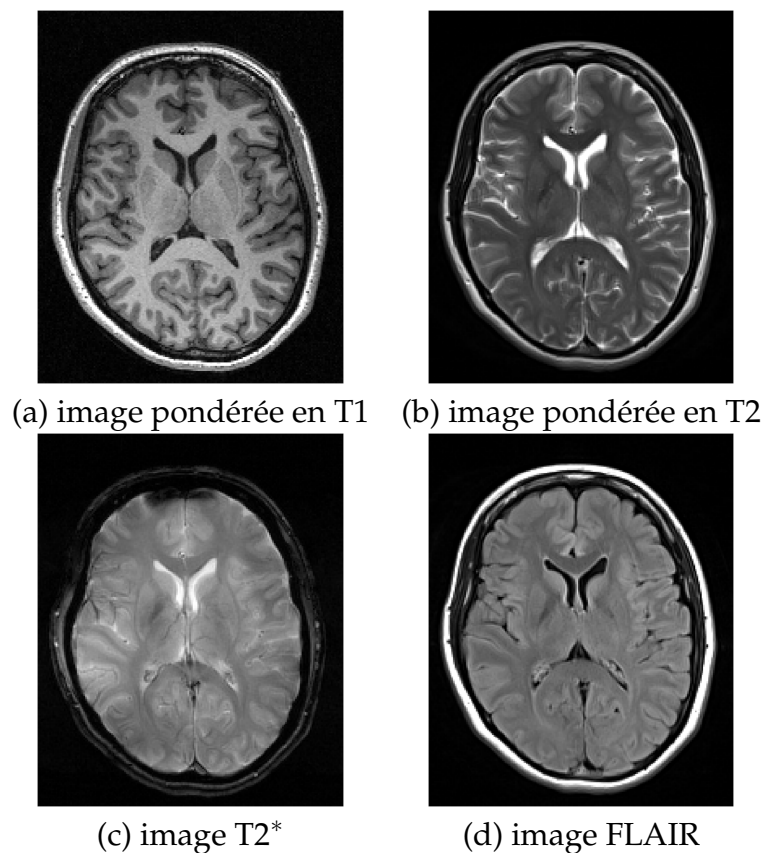


FIGURE 1.11 – Exemples de différents contrastes IRM sur le même plan de coupe pour un même sujet.

1.4.2 IRM et maladie d'Alzheimer

Dans le cas de la MA, l'imagerie est longtemps restée cantonnée à une démarche d'exclusion, visant à détecter d'autres causes possibles de syndromes démentiels (tumeur, AVC, hématome sous-dural...). Son rôle a ensuite évolué vers la recherche d'indices positifs de la maladie

d'Alzheimer. Dans ce but, les images pondérées en T1 sont particulièrement informatives, car elles permettent d'analyser l'atrophie de différentes structures cérébrales.

En effet, contrairement aux analyses histologiques, l'IRM dans un contexte clinique ne peut pas, à l'heure actuelle, détecter directement le peptide $A\beta$ ni les protéines tau. Cependant, les accumulations de plaques séniles et de DNF dans la maladie d'Alzheimer entraînent par la suite une mort neuronale et donc l'apparition d'une atrophie des structures cérébrales touchées. Les images pondérées en T1 permettent d'obtenir un bon contraste entre la substance blanche (SB) et la substance grise (SG) et d'évaluer ainsi cette atrophie.

De nombreuses méthodes d'analyse ont ainsi été développées pour analyser les différences d'atrophie entre les sujets sains et les patients Alzheimer, basées soit sur l'analyse du cortex ou de structures plus spécifiques tels que l'hippocampe [Jack et al., 1998; Thompson et al., 2003; Colliot et al., 2008; Davatzikos et al., 2008; Klöppel et al., 2008; Vemuri et al., 2008; Chupin et al., 2009; Gerardin et al., 2009; Dickerson et al., 2009; McDonald et al., 2009]. De manière générale, ces analyses permettent une bonne sensibilité et spécificité pour les patients Alzheimer mais sont moins performants au stade précoce de la maladie [Cuingnet et al., 2011].

1.4.3 IRM et démence vasculaire ischémique sous-corticale

L'IRM permet également d'étudier les trois principaux types de lésions vasculaires qui se produisent dans le cas de la maladie des petits vaisseaux : les lacunes, les microsaignements et les hyperintensités de la substance blanche (HSB). Ces trois anomalies sont illustrées sur la figure 1.12.

1.4.3.1 Lacunes

Les lacunes correspondent à des mini-infarctus de diamètre supérieurs à 3 mm et inférieurs à 15 mm [Vernooij et al., 2007]. Il s'agit donc de tissus nécrosés qui ont le même signal que le liquide cérébro-spinal sur l'ensemble des séquences IRM. Ils peuvent être visualisées en hyposignal sur les séquences pondérées en T1 et les séquences FLAIR ou en hypersignal sur les séquences pondérées en T2.

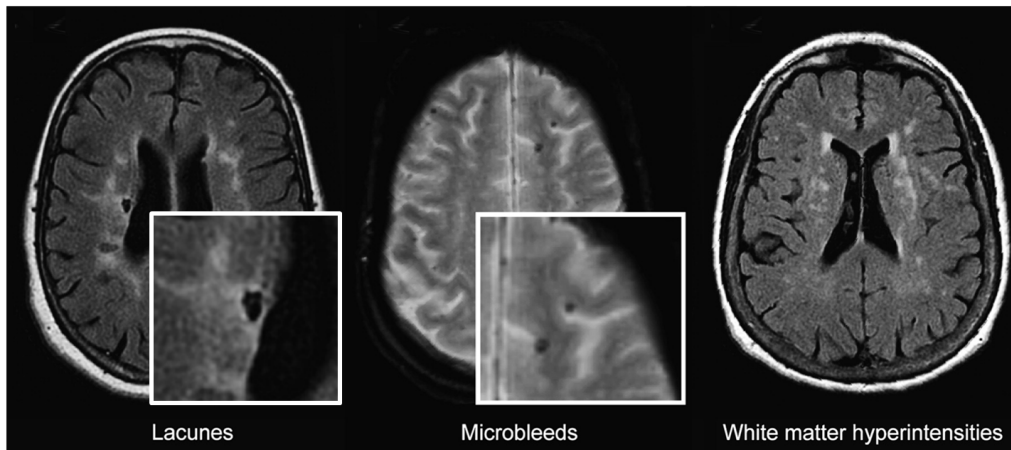


FIGURE 1.12 – Illustrations des trois types d'anomalies vasculaires en IRM.

Gauche : image FLAIR . **Milieu** : image T2* . **Droite** : image FLAIR. (source : [Gouw et al., 2011]).

La définition radiologique des lacunes peut varier selon les études, par exemple concernant leur taille maximale. Cet absence de consensus peut affecter la prévalence rapportée dans la littérature. Il apparaît néanmoins que la prévalence des infarctus lacunaires est extrêmement dépendante de l'âge. Dans une population de 63 ans de moyenne d'âge, seuls 7% des sujets présentaient au moins une lacune [Vernooij et al., 2007]. Cette proportion s'élève à près de 50% lorsque la moyenne d'âge est de 75 ans [Poggesi et al., 2012].

1.4.3.2 Microsaignements

Lorsque des microsaignements se produisent, des macrophages riches en hémossidérine, un composé contenant du fer, s'accumulent à la périphérie de l'hémorragie. Provoquant des inhomogénéités de champ magnétique, ces dépôts d'hémossidérine peuvent ainsi être visualisés sur des séquences pondérées en T2*. Leur détection est cependant difficile à cause de leur aspect très similaire aux vaisseaux ou aux calcifications [Greenberg et al., 2009]. Seules deux échelles visuelles ont été proposées pour permettre leur quantification [Gregoire et al., 2009 ; Cordonnier et al., 2009].

La prévalence des microsaignements varie selon les études entre 3,1% et 23,5%. Cette variation peut s'expliquer par les populations considérées (facteurs de risque vasculaire élevés ou non), par la différence d'âge entre les études, et par les séquences IRM plus ou moins sensibles utilisées.

Pour des populations âgées de 60 ans et plus, la prévalence totale semble être aux environs de 15-20% [Poels et al., 2010].

De nouvelles méthodes de traitement sont actuellement en cours de développement pour mieux détecter ces microsaignements en utilisant l'information contenue dans la phase du signal RMN [Barnes et al., 2011 ; Liu et al., 2012].

1.4.3.3 Hyperintensités de la substance blanche

Historiquement, les hyperintensités de la substance blanche (HSB) ont d'abord été décrits sous le terme de leucoaraïose [Hachinski et al., 1986]. Du grec "leuko" (blanc) et "araïos" (rare), ce terme correspondait à des zones situées dans la substance blanche, de plus faible densité au scanner ou ayant un signal altéré en IRM. Par la suite, les séquences IRM T2 se sont montrées particulièrement sensibles à la détection de ces anomalies. Le développement des séquences FLAIR (*FLuid Attenuated Inversion Recovery*) a encore amélioré cette sensibilité en augmentant le contraste de ces zones hyperintenses par la suppression du signal du liquide cérébrospinal (LCS).

Comme nous le développerons dans le chapitre 2, l'extension de ces hyperintensités est très variable et leur aspect peut également varier de petites taches focales à de larges étendues diffuses. De très nombreuses échelles visuelles ont été proposées pour quantifier ces HSB. Leur prévalence est très importante puisque que près de 95% des individus d'une population générale âgés de plus de 65 ans présentent des HSB [Longstreth et al., 1996].

Face à une prévalence aussi importante, plusieurs études se sont intéressées aux corrélats histologiques des HSB [Awad et al., 1986 ; Marshall et al., 1988 ; Fazekas et al., 1991 ; van Swieten et al., 1991 ; Chimowitz et al., 1992 ; Fazekas et al., 1993 ; Smith et al., 2000 ; Bronge et al., 2002 ; Gouw et al., 2008]. Une revue de ces études histopathologiques peut être trouvée dans [Gouw et al., 2011]. Ces études de corrélation entre l'histopathologie et les IRM post-mortem ont mis en évidence l'existence d'altérations variées, qui se manifestent par différents types de HSB. Les HSB modérées périventriculaires présentent ainsi des discontinuités de l'épendyme, une gliose sous-épendymaire modérée, une densité inférieure du réseau de fibres et une perte de myéline autour des dilatations des espaces périvasculaires. Les HSB périventriculaires irrégulières présentent des phénomènes

de gliose et de démyélinisation auxquelles s'ajoutent parfois des infarctus associés à la présence de vaisseaux présentant une fibrohyalinose. Le caractère confluent des HSB de la substance blanche profonde a, lui, été associé à une sévérité croissante des atteintes des tissus cérébraux. De plus, Moody et al. [2004] ont également mis en évidence une densité de vaisseaux afférents inférieure chez les patients présentant des HSB importants. Des exemples des principales altérations observées dans les HSB périventriculaires et celles de la substance blanche profonde sont montrés sur la figure 1.13.

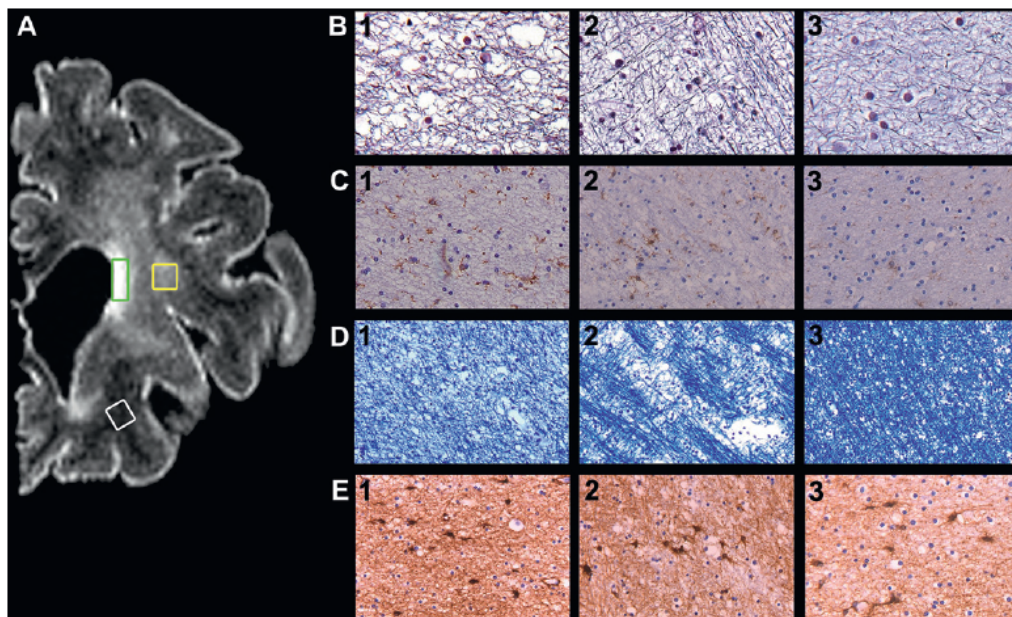


FIGURE 1.13 – Analyse postmortem des HSB.

A : IRM FLAIR postmortem d'une femme décédée à 88 ans, atteinte de la maladie d'Alzheimer. Chacune des régions indiquées est détaillée dans une des colonnes. Rectangle vert : colonne 1, rectangle jaune : colonne 2, rectangle blanc : colonne 3. **Ligne B** : mise en évidence de la densité axonale plus faible en B1 et B2 qu'en B3. **Ligne C** : Mise en évidence de l'activité microgliale en marron, plus importante en C1 et C2 qu'en C3. **Ligne D** : Mise en évidence de la myéline, moins dense en D1 et D2 qu'en D2. **Ligne E** : Les inflammations des cellules astrocytaires sont marquées en marron et sont à peu près similaires dans les trois régions. (Source : [Gouw et al., 2011])

1.5 Objectif de la thèse

L'augmentation attendue des démences constitue un enjeu majeur de santé publique du XXI^e siècle, liée essentiellement au vieillissement de la population. Cependant, la démence n'est pas un processus normal et inévitable de l'âge. Au contraire, elle est provoquée par un ensemble de

pathologies, dont les plus courantes sont la maladie d'Alzheimer et les démences vasculaires.

Ces deux types de démences semblent provoquées par des mécanismes pathologiques différents. La MA est provoquée par l'accumulation de deux types de lésions : les plaques amyloïdes et les dégénérescences neurofibrillaires. Les démences vasculaires sont, elles, causées par plusieurs types d'atteintes des vaisseaux cérébraux. En particulier, dans le cas du sous-type de démence vasculaire ischémique sous-corticale, ces atteintes entraînent la présence de lacunes, d'hémorragies et d'atteintes de la substance blanche.

L'âge est un facteur de risque important commun à l'ensemble de ces lésions. Plusieurs études histologiques ont ainsi démontré la prévalence importante de toutes ces lésions dans des populations âgées (au-delà de 70 ans). Plus particulièrement, la coexistence des lésions semblent avoir un impact important sur le déclenchement et la sévérité des symptômes cliniques de démence. Néanmoins, ces études restent limitées pour l'analyse des évolutions de lésions et de leur relations causales puisqu'elles ne peuvent être effectuées que post mortem.

L'imagerie, notamment à travers le développement de l'IRM, permet des études longitudinales sur un grand nombre de patients. En particulier, l'IRM permet d'étudier l'atrophie des structures cérébrales ainsi que les différentes lésions de la démence vasculaire ischémique sous-corticale : les lacunes, les microsaignements et les hyperintensités de la substance blanche (HSB), et ce, à différentes dates dans l'évolution de la pathologie.

Alors que de nombreuses méthodes ont été développées pour l'analyse de l'atrophie cérébrale, les approches permettant de prendre en compte les anomalies vasculaires sont plus rares. Cette thèse vise à mettre au point un outil permettant de façon automatique la détection et la quantification de l'un des trois types de lésions vasculaires, les HSB. Ces dernières sont en effet celles dont la prévalence est la plus forte. De plus, elles sont radiologiquement mieux définies que les lacunes, dont les critères de définition ne font pas toujours consensus, ou que les microsaignements, pour lesquels de nouvelles séquences plus sensibles et spécifiques sont en cours de développement. Enfin, elles sont un critère important pour le diagnostic différentiel des démences.

Chapitre 2

Quantification des hyperintensités de la substance blanche

Depuis leur description sous le terme de leucoaraïose par Hachinski en 1986, la découverte d'anomalies de la substance blanche chez les sujets âgés est devenue de plus en plus courante en pratique clinique avec les progrès de l'imagerie. Face à cette prévalence importante, de nombreuses études ont été menées pour évaluer leur impact.

Dans ce chapitre, nous présentons un état de l'art des méthodes utilisées pour la quantification des HSB. De plus, la problématique de segmentation des HSB est proche de celle, plus ancienne, de segmentation des lésions de la sclérose en plaques. Nous présentons également des approches dédiées à cette problématique et indiquons, lorsqu'elles existent, les différences entre l'application aux lésions de sclérose en plaques et celle aux hyperintensités de la substance blanche.

2.1 Principales études de cohortes

Nous listons ici les principales cohortes qui se sont intéressées à la quantification des HSB et auxquelles nous ferons référence par la suite.

Cardiovascular Health Study (CHS) [Fried et al., 1991] : Cette étude vise à étudier les facteurs de risque principaux des maladies vasculaires. 5888 sujets âgés de plus de 65 ans furent recrutés dans quatre centres américains entre 1989 et 1993. Parmi ces 5888 sujets, 3660 ne présentaient

2. QUANTIFICATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE

pas de contre-indications et acceptèrent de passer une IRM cérébrale entre 1991 et 1994.

Atherosclerosis Risk in Communities Study (ARIC) [investigators, 1989] : 15 972 sujets, âgés de 45 à 64 ans, participèrent à cette étude, débutée en 1987 et répartie entre quatre centres américains. En 1993, 1920 participants, parmi ceux d'au moins 55 ans, furent sélectionnés aléatoirement dans deux de ces centres pour une IRM cérébrale.

Austrian Stroke Prevention Study (ASPS) [Schmidt et al., 1994] : Cette étude vise à évaluer les facteurs de risque vasculaire sur les structures et fonctions cérébrales dans une population générale âgée. Elle se déroula sur un seul centre à Graz en Autriche. Incluant 2007 personnes au total, 1076 participants, âgés de 45 à 85 ans, furent aléatoirement sélectionnés et passèrent également une IRM cérébrale.

Framingham Heart Study (FHS) [Kannel et al., 1979] : Initiée en 1948, cette étude inclut environ les deux tiers de la population adulte résidant alors à Framingham aux Etats-Unis, soit 5209 personnes. Elle fut poursuivie en étudiant également les descendants de la cohorte originelle et est désormais composée de trois générations (seconde étude en 1971 et troisième étude en 2000). Parmi les deux premières générations, environ 2300 passèrent une IRM cérébrale entre 1999 et 2005.

Leukaraiosis And DISability (LADIS) [Pantoni et al., 2005] : Cette étude européenne multicentrique a démarré en 2001 et vise à évaluer l'impact de la leucoaraïose sur la perte d'autonomie chez les personnes âgées entre 65 et 84 ans. Onze centres européens sont impliqués dans cette étude, qui inclut 639 participants.

Aging Genes-Environnement Susceptibility - Reykjavik (AGES - Reykjavik) [Harris et al., 2007] : Débutée en 1967, cette étude incluait 30 795 personnes vivant à Reykjavik et nées entre 1907 et 1935. Entre 2002 et 2006, parmi les 11 539 participants encore en vie, 5764 furent aléatoirement sélectionnés et invités à passer une IRM cérébrale.

Epidemiology of Vascular Ageing (EVA) [Dufouil et al., 2001] : L'étude EVA incluait 1389 sujets recrutés sur les listes électorales de la ville de Nantes en France. Les sujets, âgés entre 63 et 75 ans lors de l'inclusion, ont eu un suivi après deux et quatre ans. Parmi les 1389 sujets recrutés, une IRM cérébrale fut acquise pour 844 sujets.

Etude des 3 Cités (3C) [3C Study Group, 2003] : L'étude des 3 Cités est une étude française débutée en 1999. Elle implique les villes de Bor-

deaux, Montpellier et Dijon, où 9294 personnes de plus de 65 ans ont été recrutées. Environ 3000 sujets ont également été invités à passer une IRM.

Rotterdam Scan Study [Hofman et al., 2009] : Cette étude est basée dans la ville de Rotterdam aux Pays-Bas et s'est déroulée en plusieurs vagues. La première vague (Rotterdam Scan Study I) a commencé en 1990 et inclut 7983 personnes, suivies tous les 3 à 4 ans. Un sous-groupe de 563 sujets passèrent une IRM cérébrale en 1995-1996. Durant la deuxième vague (Rotterdam Scan Study II), 3 011 participants supplémentaires furent ajoutés en 2000-2001. Parmi ceux-ci, 895 furent invités à passer une IRM.

La plupart des études mentionnées ci-dessous font également partie du consortium CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) [Psaty et al., 2009]. Cette méta-cohorte internationale permet, grâce à l'importante population étudiée, d'effectuer des études d'associations pangénomiques. Les premiers résultats ont ainsi permis d'identifier un locus situé sur le chromosome 17, associé à l'augmentation de la charge lésionnelle de HSB [Fornage et al., 2011].

2.2 Echelles visuelles et cotations

Parmi les premières échelles proposées pour la mesure des HSB se trouve l'échelle dite de Fazekas & Schmidt [Fazekas et al., 1987], qui est encore couramment utilisée aujourd'hui (figure 2.2). Cette échelle différencie deux types de HSB, les HSB périventriculaires et celles de la substance blanche profonde, et attribue ensuite une note de plus en plus grande suivant l'extension et le caractère confluent des lésions.

Plusieurs versions de cette échelle de Fazekas existent et une version simplifiée classe les lésions en trois stades : Léger, Modéré et Sévère (figure 2.1). C'est cette version simplifiée qui a été utilisée pour l'évaluation visuelle dans le cadre de l'étude LADIS.

Pour les études CHS et ARIC, une autre échelle visuelle a été utilisée, qui distinguait 10 stades différents (figure 2.3). La cotation était effectuée par comparaison avec huit IRM, qui servirent de référence, partagées entre tous les centres [Manolio et al., 1994].

De très nombreuses autres échelles visuelles ont été proposées [Gerard & Weisberg, 1986; Shimada et al., 1990; van Swieten et al., 1990; Wahlund et al., 1990; Mirsen et al., 1991; Scheltens et al., 1993; Erkinjuntti et al., 1994;

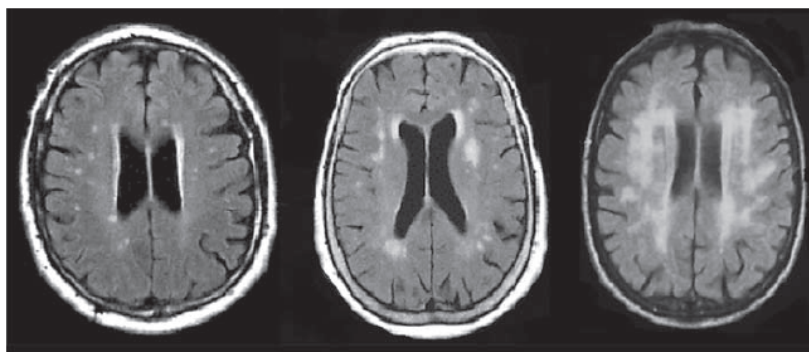


FIGURE 2.1 – Echelle modifiée de Fazekas. (Source : [Poggesi et al., 2011])
Gauche : stade léger. **Milieu** : stade modéré. **Droite** : stade sévère.

Wahlund et al., 2001 ; Bocti et al., 2005]. Mäntylä et al. [1997] ont comparé treize de ces échelles et ont conclu à l'incohérence de certaines de ces échelles entre elles, et à la nécessité d'une harmonisation des procédures de quantification des HSB.

De plus, van Straaten et al. [2006] ont comparé les effets de différentes échelles visuelles sur les corrélations avec des symptômes cliniques. En particulier, les échelles visuelles utilisées étaient moins sensibles pour la détection des troubles de mémoire, mieux mis en évidence par des méthodes volumétriques.

2.3 Segmentation manuelle

Afin d'obtenir une mesure quantitative du volume des HSB, certaines études ont établi des protocoles de segmentation manuelle.

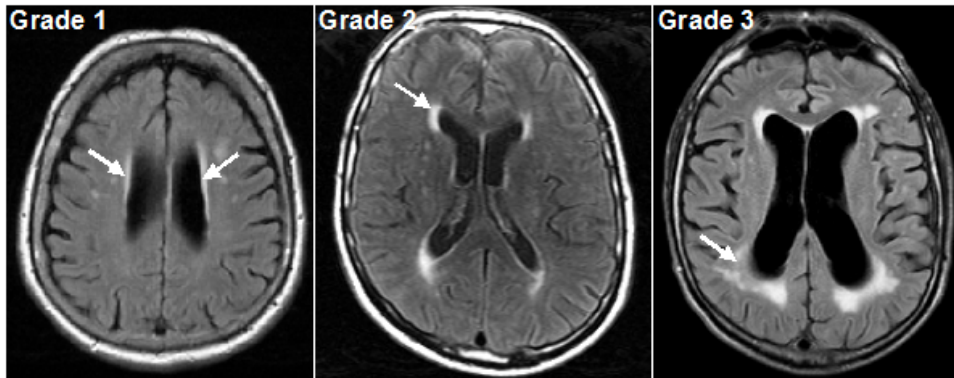
C'est par exemple le cas de l'étude ASPS, où les HSB ont été segmentées à la main sur les plusieurs centaines d'IRM. Ce travail est évidemment particulièrement fastidieux et cette méthode ne peut être utilisée de manière répétée sur différentes études de milliers de patients. En plus de la charge de travail que cela représente, cette méthode manuelle pose également la question de la reproductibilité et de la subjectivité de l'opérateur.

2.4 Segmentation automatique

Les méthodes de segmentation automatique sont ainsi apparues progressivement pour obtenir des mesures quantitatives fiables qui puissent

Grade des hypersignaux périventriculaires (illustration sur FLAIR)

- 0 pas d'anomalies
- 1 caps ou hypersignal fin bordant les ventricules
- 2 hypersignal paraventriculaire plus étendu, à bord régulier
- 3 hypersignaux à bords irréguliers s'étendant dans la substance blanche



Grade des hypersignaux dans la substance blanche profonde (illustration sur T2)

- 0 pas d'anomalies
- 1 foyers ponctuels
- 2 début de confluence des foyers
- 3 larges plages confluentes

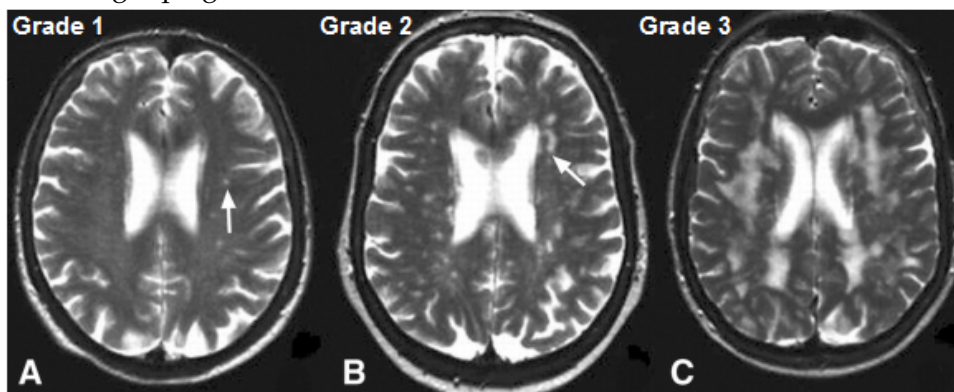


FIGURE 2.2 – Echelle de Fazekas & Schmidt. Protocole utilisé au Centre de Neuroimagerie de Recherche - Hôpital Pitié-Salpêtrière.

2. QUANTIFICATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE

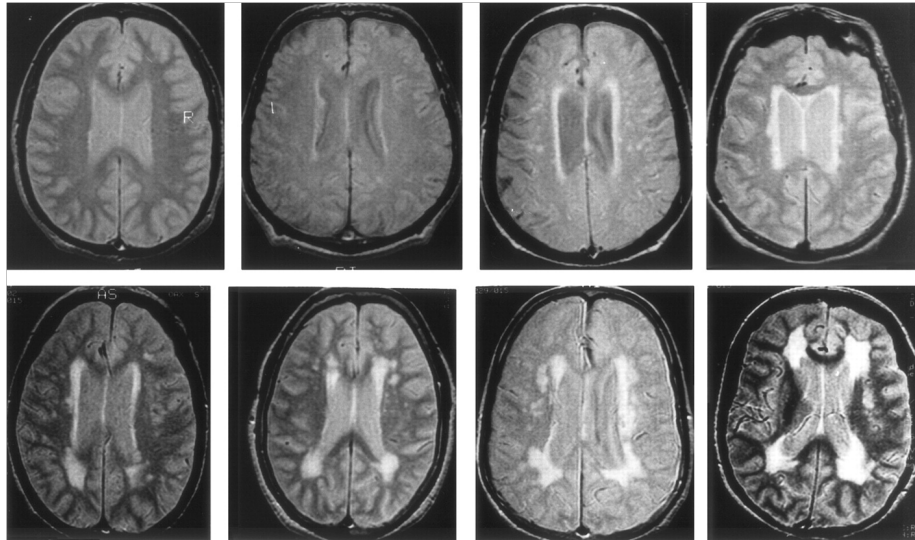


FIGURE 2.3 – Exemples de coupes des IRM de référence utilisée par les études CHS et ARIC pour la cotation des HSB. (Source : [Longstreth et al., 1996])

Ligne du haut : score de 1 à 4. **Ligne du bas** : score de 5 à 8. Les cas sans aucune atteinte de substance blanche étaient notés 0 et ceux où les changements étaient plus importants que le dernier cas étaient notés 9.

être obtenues facilement sur des cohortes de plusieurs centaines ou milliers de patients.

De manière générale, nous pouvons diviser ces méthodes en deux grandes catégories : les méthodes non supervisées et les méthodes supervisées. Les premières s'appuient sur une modélisation ou des règles de classification, qui ont été spécialement adaptées au problème considéré. Les secondes, en revanche, utilisent une partie des données où une segmentation manuelle est disponible, appelée *ensemble d'apprentissage*, et infèrent, par différents algorithmes de classification, les règles permettant d'obtenir ces segmentations. Ces règles sont ensuite appliquées au reste des données.

2.4.1 Méthodes non supervisées

2.4.1.1 Méthode par seuillage

Parmi les méthodes non supervisées, nous trouvons une première catégorie s'appuyant sur le seuillage des intensités. Dès 2001, Jack et al. [2001] ont proposé de déterminer un seuil sur l'histogramme des intensités de l'image FLAIR afin de détecter les HSB (figure 2.4). Le seuil optimal, T_U ,

est déterminé par une analyse de régression sur un ensemble de données simulées. Bien que les auteurs rapportent une erreur moyenne volumique de 6,6%, l'évaluation n'a porté que sur dix sujets simulés et apparaît peu robuste à la variabilité interindividuelle.

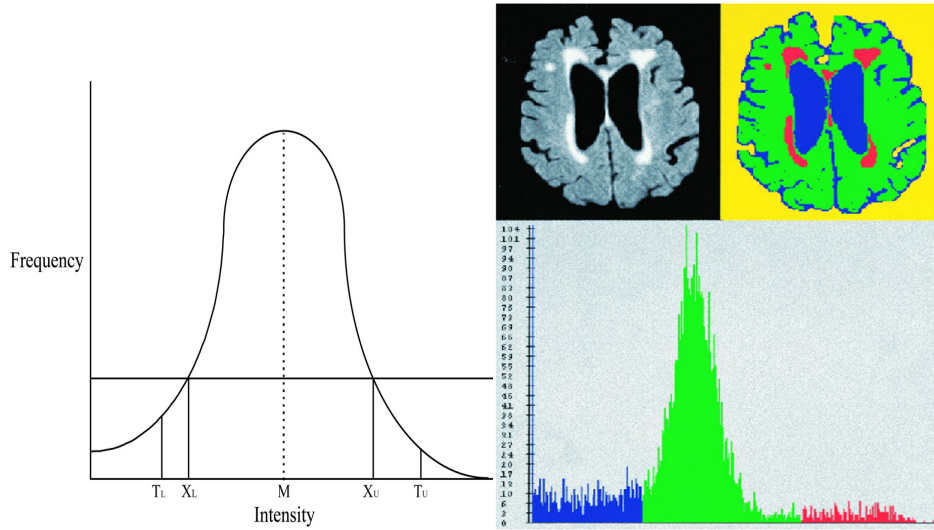


FIGURE 2.4 – Seuillage de l'histogramme FLAIR [Jack et al., 2001]

Par la suite, Wen & Sachdev [2004] ont proposé de déterminer le seuil pour la détection des HSB à l'aide d'un atlas probabiliste du cerveau. Cet atlas est recalé sur l'image à segmenter, permettant de calculer la moyenne μ_{WM}^{FLAIR} et l'écart-type σ_{WM}^{FLAIR} des intensités FLAIR sur la substance blanche. Le seuil caractérisant les HSB est alors déterminé comme $\mu_{WM}^{FLAIR} + 3\sigma_{WM}^{FLAIR}$. L'utilisation de l'atlas de substance blanche permet de pondérer l'importance de la localisation et un seuil déterminé sur l'image T1 permet d'éliminer certaines erreurs de segmentation.

Une approche similaire a été utilisée pour l'étude FHS [Jeerakathil et al., 2004]. Une segmentation manuelle de la cavité intracrânienne était effectuée pour ne conserver que les tissus cérébraux. Un premier seuil était déterminé permettant de segmenter le liquide cébrospinal. Puis les deux échos des séquences T2 étaient additionnés et la somme était modélisée par une distribution log-normale [DeCarli et al., 1992, 1999]. La segmentation des HSB était alors obtenue en seuillant les intensités à $m + 3,5sd$, où m est la moyenne de la distribution modélisée et sd son écart-type.

Enfin, la deuxième vague de l'étude Rotterdam Scan Study s'appuie sur une approche similaire [de Boer et al., 2009]. Un algorithme de classi-

fication par k-plus proches voisins (voir section 5.3.3) est d'abord utilisé pour segmenter les trois principaux types de tissus (substance blanche, substance grise et liquide cérebrospinal) à partir des images T1 et PD. Les images FLAIR sont ensuite utilisées pour la segmentation des HSB. L'histogramme des intensités FLAIR des voxels classifiés comme substance grise est lissé par convolution avec une gaussienne afin de déterminer le mode principal correspondant à la substance grise considérée comme normale. Les intensités FLAIR sont ensuite seuillées à $m + 2,3sd$, m étant défini comme la valeur du pic d'intensité de la distribution lissée et sd comme la largeur du pic à mi-hauteur. La valeur de 2,3 est fixée de manière empirique sur un ensemble test. Une contrainte est ajoutée sur la localisation en imposant qu'un pourcentage minimal des voisins du voxel considéré soit inclus dans la substance blanche.

2.4.1.2 Logique floue

Une deuxième catégorie de méthodes s'appuie sur les principes de logique floue et l'algorithme *fuzzy C-means* (FCM). Dans le cadre de la logique floue, un voxel peut être considéré comme appartenant à plusieurs classes par l'intermédiaire de son degré d'appartenance à une classe, degré qui peut prendre une valeur comprise entre 0 et 1.

Gibson et al. [2010] combinent une approche de type seuillage avec l'algorithme FCM. Cette approche mixte est également utilisée dans [Wu et al., 2006], où les voxels hyperintenses détectés par seuillage sont considérés comme graines d'un algorithme flou de croissance de régions.

Admiraal-Behloul et al. [2005] s'appuient également sur les principes de logique floue (figure 2.5). Les intensités des images sont ainsi transformées en variables linguistiques en fonction des résultats de classification du FCM : les intensités des images pondérées en densité de proton sont ainsi classées en *bright* ou *dark* alors que les intensités des images T2 et FLAIR sont classées en *bright*, *medium-bright* et *dark*. Un système d'inférence floue est ensuite utilisé pour segmenter le LCS et les HSB à partir des deux règles d'inférence floue suivantes :

- Si *position_voxel* est IC et *intensité_T2* est *bright* et *intensité_FLAIR* est *dark*, alors *voxel* est CSF.
- Si *position_voxel* est SB et *intensité_T2* est *bright* et *intensité_FLAIR* est *bright*, alors *voxel* est HSB.

où la valeur de la variable *position_voxel* (IC pour intracrânien et SB pour substance blanche) est définie à partir du recalage d'un atlas sur l'image à segmenter.

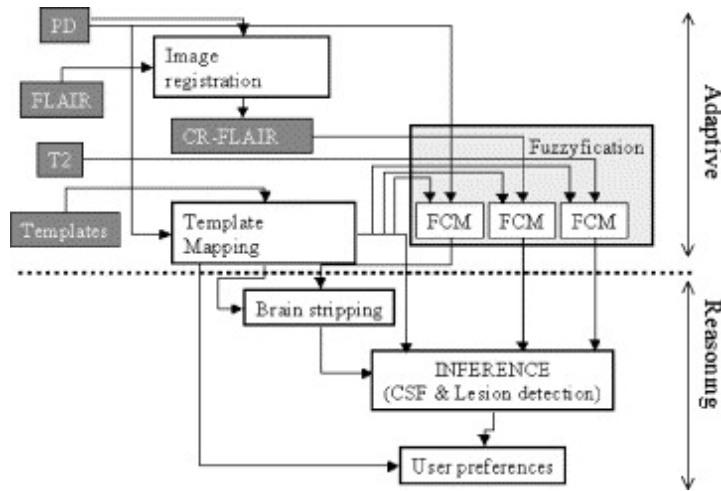


FIGURE 2.5 – Principe de la segmentation des HSB par Admiraal-Behloul et al. [2005]

2.4.1.3 Modélisation probabiliste

Une troisième catégorie de méthodes non supervisées regroupe celles s'appuyant sur une modélisation probabiliste, le plus souvent à travers un modèle de mélange de gaussiennes.

Ainsi, Gicquel [1998] proposent une segmentation bayésienne multispectrale basée sur les images T1, T2 et PD. Cette approche est reprise et améliorée par Maillard et al. [2008] et a été appliquée sur l'étude 3C. Après des étapes classiques de prétraitement (recalage et correction des inhomogénéités d'intensité), une segmentation en 7 classes (liquide cérébrospinal, substance grise, noyau caudé, noyau lenticulaire, thalamus, substance blanche et HSB) est obtenue. Une série d'opérations morphologiques est ensuite appliquée de manière à corriger les erreurs de segmentation liées à l'imprécision du recalage et aux effets de volume partiel. Un masque regroupant la substance blanche et les HSB est ainsi obtenu. La méthode de segmentation multispectrale est de nouveau appliquée sur ce masque en distinguant trois classes : substance blanche, HSB faiblement contrastées et HSB fortement contrastées. Finalement, certains groupes de voxels, classés de manière erronée comme HSB, sont éliminés de la segmentation si plus de la moitié de leur volume est compris dans un masque du LCS

ou si leur intensité moyenne sur l'image T2 est inférieure à celle de la substance blanche.

Khayati et al. [2008] ont proposé une approche similaire dans le cas des lésions de sclérose en plaques. Un modèle de mélange de gaussiennes est combiné avec une régularisation par champ de Markov dans un classifieur bayésien. A la différence de la méthode précédente, le modèle de mélange est cette fois adaptatif, le nombre de gaussiennes, leur moyenne et variance étant estimés de manière itérative à partir des données.

Enfin, Van Leemput et al. [2001] s'appuient également sur un modèle de mélange de gaussiennes appliqué sur les images T1, T2 et PD. Cependant, les HSB ne sont plus modélisées dans le modèle de mélange de gaussiennes mais sont au contraire caractérisées comme des points aberrants du modèle. Quatre classes sont ainsi utilisées pour la modélisation ("substance blanche", "substance grise", "liquide cérébrospinal" et "autre"). Les paramètres du modèle sont obtenus par un algorithme d'*Espérance-Maximisation* rendu robuste aux points aberrants par l'introduction de poids quantifiant l'accord d'un point avec le modèle. Ces poids, appelés *poids de typicité*, dépendent de la classe considérée et permettent de déterminer un niveau de significativité à partir duquel les points seront considérés comme aberrants.

2.4.2 Méthodes supervisées

Les méthodes supervisées s'appuient sur les techniques issues de l'apprentissage automatique. En effet, la segmentation des HSB peut aussi être vue comme une problématique de classification binaire où, pour chaque voxel d'une image, on cherche à savoir s'il appartient à la classe "non HSB" (notée par exemple 0) ou à la classe "HSB" (notée 1) à partir d'une ou plusieurs caractéristiques associées à ce voxel (par exemple, les intensités de ce voxel sur différentes séquences IRM). Pour effectuer cette tâche, les méthodes supervisées vont s'appuyer sur un *ensemble d'apprentissage*, constitué d'exemples de voxels pour lesquels la classe est déjà connue. Dans le cas de la segmentation des HSB, une segmentation manuelle (ou l'accord entre plusieurs segmentations manuelles) effectuée par un expert est fréquemment choisie comme référence puisque la vérité histologique n'est pas connue. Les données que l'on cherche à classer sont appelées *ensemble test*.

Plusieurs cadres mathématiques ont été proposés pour construire des règles de classification à partir d'un ensemble d'apprentissage de manière à pouvoir ensuite appliquer cette règle sur l'ensemble test. Les réseaux de neurones artificiels, la classification par k plus proches voisins et les machines à vecteurs de support ont notamment été utilisés dans le cas des HSB.

2.4.2.1 Classification par réseaux de neurones artificiels (ANN)

Les réseaux de neurones artificiels (ANN pour *Artificial Neural Networks*) sont une méthode de classification inspirée du fonctionnement des neurones biologiques. Schématiquement, un réseau de neurones est composé de plusieurs couches, chacune composée de différents noeuds. Chaque couche possède comme entrées les valeurs de sortie produites par la couche précédente. La couche d'entrée possède autant de noeuds que de valeurs d'entrée, la couche de sortie fournit une valeur correspondant au problème de classification (dans notre cas, deux valeurs, 0 pour "non HSB" et 1 pour "HSB" par exemple). Pour chaque noeud, les valeurs d'entrée sont pondérées par un coefficient, appelé *poids synaptique*, puis sommées et une fonction dite d'activation calcule une valeur de sortie en fonction de cette somme. En comparant la valeur prédite avec la valeur attendue, il est possible d'adapter les poids synaptiques des noeuds de chaque couche pour obtenir une meilleure fonction de prédiction.

Zijdenbos et al. [2002] ont proposé d'utiliser une classification par réseaux de neurones artificiels pour effectuer la segmentation des tissus cérébraux. La classification fait suite à un ensemble de prétraitements permettant d'homogénéiser les images à travers un recalage non linéaire dans un espace stéréotaxique, une normalisation des intensités entre les coupes, un débruitage et une correction des inhomogénéités d'intensité. Un réseau de neurones à trois couches est ensuite entraîné sur un ensemble d'apprentissage composé de dix images. Les variables d'entrée sont les valeurs d'intensités pour trois contrastes IRM (T1, T2 et PD) ainsi que la probabilité d'appartenance à l'un des trois tissus principaux (substance blanche, substance grise ou liquide cébrospinal). Ces trois dernières valeurs sont obtenues à partir d'un atlas statistique construit dans le même espace. Cette méthode de classification fut tout d'abord appliquée au cas de la sclérose en plaques sur un grand nombre d'images ($\simeq 1000$)

2. QUANTIFICATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE

provenant de plusieurs centres d'acquisition. Les résultats obtenus par cette méthode étaient statistiquement identiques à ceux obtenus par des segmentations manuelles.

Suite aux bons résultats observés pour la sclérose en plaques, cet algorithme a ensuite été modifié pour être appliqué à l'étude LADIS [Dyrby et al., 2008]. Plusieurs structures de réseaux de neurones ont été évaluées, prenant en compte soit les images T1 et T2, soit les images FLAIR, soit leur combinaison. Des informations concernant la localisation ou le voisinage du voxel à classifier ont également été ajoutées. Tous les modèles présentaient des performances à peu près similaires et relativement décevantes, bien que l'utilisation des toutes les séquences IRM améliorait légèrement les résultats.

Les ANN ont également été appliqués dans l'étude AGES - Reykjavik en prenant en compte quatre séquences IRM (T1, T2, PD et FLAIR) [Sigurdsson et al., 2011]. La classification visait ici la segmentation de tous les types de tissus, y compris substance blanche, substance grise et liquide cébrospinal. Alors que les résultats concernant les autres types de tissus étaient plutôt satisfaisants, ceux concernant les HSB étaient plus faibles (indice de similarité moyen : 0,62).

2.4.2.2 Classification par k plus proches voisins (kNN)

Les classifications par k plus proches voisins (kNN pour *k Nearest Neighbors*) déterminent la classe d'un voxel de l'ensemble test par comparaison avec la classe de ses k plus proches voisins dans l'ensemble d'apprentissage.

Une classification par k plus proches voisins pour les HSB a été proposée par Anbeek et al. [2004] et fut utilisée pour la première vague de l'étude Rotterdam Scan Study. Elle s'appuie sur différents contrastes IRM (T1, T2, PD, FLAIR et IR). En particulier, Anbeek et al. [2005] ont analysé les différentes combinaisons possibles de ces contrastes et montré que les images FLAIR étaient particulièrement discriminantes (tableau 2.1).

2.4.2.3 Classification par Machines à Vecteurs de Support (SVM)

Les Machines à Vecteur de Support (SVM pour *Support Vector Machines*) font partie des techniques les plus récentes d'apprentissage supervisé. Le principe général des SVM est de construire une séparatrice qui ne

	Tous contrastes	IR	FLAIR	IR/FLAIR	T1/PD/T2
Substance blanche	0,893	0,865*	0,775*	0,892	0,850*
Substance grise	0,830	0,745*	0,652*	0,832	0,741*
LCS	0,819	0,718*	0,795	0,842	0,737*
Ventricules	0,888	0,845*	0,917	0,909	0,870
HSB	0,805	0,407*	0,803	0,808	0,631*

TABLEAU 2.1 – Indice de similarité moyen entre la segmentation obtenue par kNN et une segmentation de référence en fonction des contrastes IRM utilisés (Source : [Anbeek et al., 2005]).

L'indice de similarité permet de caractériser le recouvrement spatial entre deux segmentations (voir section sec :dice). Une valeur de 0 indique qu'il n'y a aucun recouvrement, une valeur de 1 indique un recouvrement parfait. * : indice de similarité statistiquement différent de celui obtenu en utilisant tous les contrastes IRM (*t-tests appariés*, $p < 0,05$)

dépende que d'un petit nombre de vecteurs, les vecteurs dits supports, qui sont supposés être les plus difficiles à classer. Cette propriété devrait notamment permettre une meilleure robustesse des SVM. De plus, les SVM permettent de gérer facilement les espaces de grande dimension.

Une application aux HSB a été proposée par Lao et al. [2008]. La classification s'effectue à partir de quatre modalités IRM (T1, T2, PD et FLAIR). L'évaluation n'a porté que sur 45 sujets provenant de trois centres différents. 10 sujets ont été utilisés pour constituer l'ensemble d'apprentissage, ne laissant donc que 35 sujets pour l'évaluation des performances.

2.4.3 Comparaison des méthodes

Il est très difficile de comparer les résultats des différentes méthodes présentées ci-dessous. En effet, la plupart du temps, les méthodes ont été appliquées sur des jeux de données différents. La qualité des acquisitions, la charge lésionnelle totale, les méthodes de prétraitement utilisées, la définition de l'ensemble d'apprentissage pour les méthodes supervisées peuvent tous avoir une influence importante sur les résultats obtenus.

Klöppel et al. [2011] ont évalué différentes méthodes de segmentation des HSB sur un même jeu de données de 20 sujets. 10 sujets ont été aléatoirement choisis pour constituer l'ensemble test et les 10 sujets restants servaient à l'évaluation des performances. Les séquences IRM utilisées étaient les images T1 et FLAIR. Parmi les méthodes évaluées se trouvaient plusieurs méthodes de seuillage, les méthodes de classification

par k plus proches voisins et celles par machines à vecteurs de support. Pour l'ensemble des méthodes, l'accord moyen des segmentations automatiques avec une segmentation manuelle était inférieur à l'accord entre deux segmentations manuelles. De plus, les résultats indiquaient une très mauvaise performance des méthodes de seuillage et une supériorité du SVM par rapport au kNN. Cette étude était toutefois limitée par le caractère monocentrique des données utilisées, qui ne permettent donc pas d'évaluer la robustesse de ces approches pour des études multicentriques, ainsi que par le faible nombre de sujets dans l'ensemble test (10 sujets). De plus, pour les méthodes supervisées, l'impact de l'ensemble d'apprentissage n'a pas non plus été évalué. Il s'agit pourtant d'un aspect essentiel pour la généralisabilité de ces méthodes.

2.5 Discussion

Le tableau 2.2 synthétise les méthodes utilisées dans les grandes cohortes présentées au début de ce chapitre. Il est à noter que la cotation visuelle demeure couramment utilisée dans toutes ses études. Une seule étude a utilisé une procédure de segmentation manuelle sur toutes ses images, ce qui confirme que cette procédure n'est pas utilisable de manière systématique à cause du temps important qu'elle requiert et de son caractère fastidieux.

Cinq études ont développé des méthodes de segmentation automatique des HSB pour l'analyse de la cohorte (FHS, LADIS, AGES - Reykjavik, 3C, Rotterdam Scan Study). Cependant, trois d'entre elles étaient des études monocentriques. Parmi ces dernières, FHS et Rotterdam Scan Study II ont mis en place des procédures à base de seuillage alors que Klöppel et al. [2011] ont montré que ces approches étaient peu robustes. Rotterdam Scan Study I et AGES - Reykjavik utilisent, elles, des méthodes supervisées (respectivement kNN et ANN). Cependant, les méthodes supervisées sont particulièrement dépendantes de la représentativité de l'ensemble d'apprentissage et leur généralisabilité à de nouveaux centres nécessite une harmonisation importante des conditions d'acquisition et une relative homogénéité des aspects des lésions.

Cela ne semble pas être le cas pour les HSB, puisque l'étude LADIS, qui s'est appuyée sur la même approche que l'étude AGES - Reykjavik, rapporte des résultats décevants pour les ANN. L'étude 3C est la seule

Etude	Nb centres	Séquences IRM	Méthode de quantification
CHS	2	T1, T2 et PD	Cotation 0-10 ([Manolio et al., 1994])
ARIC	2	T1, T2 et PD	Cotation 0-10 ([Manolio et al., 1994])
ASPS	1	T1, T2 et PD	Cotation visuelle ou segmentation manuelle
FHS	1	T1, T2 et PD	Modélisation probabiliste et seuillage [Jeerakathil et al., 2004]
LADIS	11	T1, T2 et FLAIR	Cotation visuelle ou ANN [Zijdenbos et al., 2002; Dyrby et al., 2008]
AGES-Reykjavik	1	T1, T2, PD et FLAIR	Cotation visuelle ou ANN [Zijdenbos et al., 2002; Sigurdsson et al., 2011]
3C / EVA	3	T1, T2 et PD	Cotation visuelle ou segmentation bayésienne multispectrale [Maillard et al., 2008]
Rotterdam Scan Study I	1	T1, T2, PD, IR et FLAIR	Cotation visuelle ou kNN [Anbeek et al., 2005]
Rotterdam Scan Study II	1	T1, PD et FLAIR	Cotation visuelle ou kNN pour les tissus normaux suivi d'un seuillage pour les HSB [de Boer et al., 2009]

TABLEAU 2.2 – Synthèse des méthodes de quantification utilisées pour les principales cohortes.

étude multicentrique utilisant une méthode non supervisée. Elle n'utilise cependant pas le FLAIR, alors que cette séquence IRM a depuis été démontrée comme étant particulièrement informative pour la segmentation des HSB (tableau 2.1). De plus, elle n'a pas été évaluée par rapport à une segmentation manuelle mais uniquement en étudiant la corrélation avec une échelle visuelle.

Il est également à noter que différents travaux ont également porté sur une meilleure caractérisation des HSB par l'utilisation d'autres séquences IRM. On peut ainsi citer le transfert de magnétisation ([Fazekas et al., 2005]), l'utilisation d'indices de diffusivité tels que l'anisotropie fraction-

2. QUANTIFICATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE

nelle, la diffusivité radiale ou moyenne [Vernooij et al., 2008 ; Burzynska et al., 2010] ou encore l'utilisation de séquences DIR (Double Inversion Recovery) [Turetschek et al., 1998 ; Madhuranthakam et al., 2011]. Cependant, les études épidémiologiques ayant permis de mettre en évidence les conséquences cliniques des HSB s'appuient pour l'essentiel sur les séquences pondérées en T2 et les séquences FLAIR, qui demeurent donc celles couramment utilisées.

Dans cette thèse, nous proposons une nouvelle méthode de segmentation des HSB. Cette méthode sera non supervisée afin de ne pas dépendre de l'ensemble d'apprentissage et devra être robuste aux différences d'acquisition entre différents centres pour être applicable dans un environnement multicentrique. Elle s'appuiera essentiellement sur les images FLAIR, modalité la plus informative. Nous évaluerons cette méthode par comparaison avec une segmentation manuelle de référence sur des données multicentriques présentant une grande variabilité de charge lésionnelle. Enfin, nous comparerons nos résultats avec ceux de plusieurs autres méthodes de segmentation, aussi bien non supervisées que supervisées. Cette comparaison sera faite sur un même jeu de données et en utilisant les mêmes prétraitements afin de mieux évaluer les performances propres aux méthodes.

Chapitre 3

Méthode de segmentation des hyperintensités de la substance blanche (WHASA)

L'état de l'art présenté au chapitre précédent nous conduit au développement d'une nouvelle méthode de segmentation qui s'appuie sur les images T1 et FLAIR. Un critère important pour le développement de la méthode doit être sa robustesse aux conditions d'acquisition (type de machine, intensité du champ) ainsi qu'à l'importante variabilité de la pathologie (allant de l'absence totale de HSB jusqu'à la présence de HSB confluentes occupant la quasi totalité de la substance blanche). Le but de cette méthode est également d'être applicable à des données de qualité clinique et pouvant être acquises dans des protocoles incluant un bilan classique pour les sujets âgés.

Ce chapitre est organisé de la manière suivante : nous présentons tout d'abord le schéma général de la chaîne de traitements proposée (section 3.1). Nous détaillons ensuite chacune de ces composantes. La section 3.2 résume les différentes étapes de prétraitements effectués permettant de faciliter l'analyse automatique des images. La section 3.3 détaille la manière de combiner diffusion non linéaire et ligne de partage des eaux pour obtenir deux segmentations présentant différents niveaux de détails de l'image FLAIR. La section 3.4 précise les critères utilisés pour sélectionner les HSB au niveau de détail le plus grossier. Enfin, la section 3.5 détaille l'utilisation de la dilatation géodésique pour améliorer la segmentation obtenue en utilisant le niveau de détail plus fin.

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

Une version antérieure de cette méthode a été décrite dans [Samaille et al., 2012]. L’approche détaillée dans ce manuscrit diffère sur certains points, notamment afin d’améliorer la robustesse et de diminuer le nombre de paramètres à fixer.

3.1 Schéma général de la chaîne de traitements

La méthode de segmentation des HSB développée au cours de cette thèse, dénommée WHASA (*White matter Hyperintensities Automatic Segmentation Algorithm*), est résumée dans la figure 3.1. Elle repose sur 4 étapes principales :

1. une étape de **prétraitements**, qui permet de corriger le biais des images IRM et d’obtenir une première segmentation des principaux types de tissus (substance blanche (SB), substance grise (SG) et liquide cébrospinal (LCS)). Cette étape s’appuie sur le logiciel SPM¹.
2. une étape de **segmentation de l’image FLAIR**. Cette étape est obtenue par la combinaison de deux algorithmes classiques en traitement des images : la diffusion non linéaire et l’algorithme de ligne de partage des eaux (LPE).
3. une étape de **sélection des régions** obtenues à l’étape précédente, afin de conserver celles correspondantes aux HSB.
4. une étape de **croissance**, qui permet de rendre plus précise la segmentation des bords des HSB obtenues à l’étape précédente.

3.2 Prétraitements

Le traitement et l’analyse des images médicales ont fait l’objet de nombreuses recherches au cours des dernières décennies. En ce qui concerne les images cérébrales, l’analyse d’un certain nombre de problèmes a permis d’aboutir à des solutions relativement standard qui sont disponibles à travers diverses suites logicielles, telles que Freesurfer (<http://surfer.nmr.mgh.harvard.edu/>), SPM (<http://www.fil.ion.ucl.ac.uk/spm/>), FSL (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>) ou

1. <http://www.fil.ion.ucl.ac.uk/spm/>

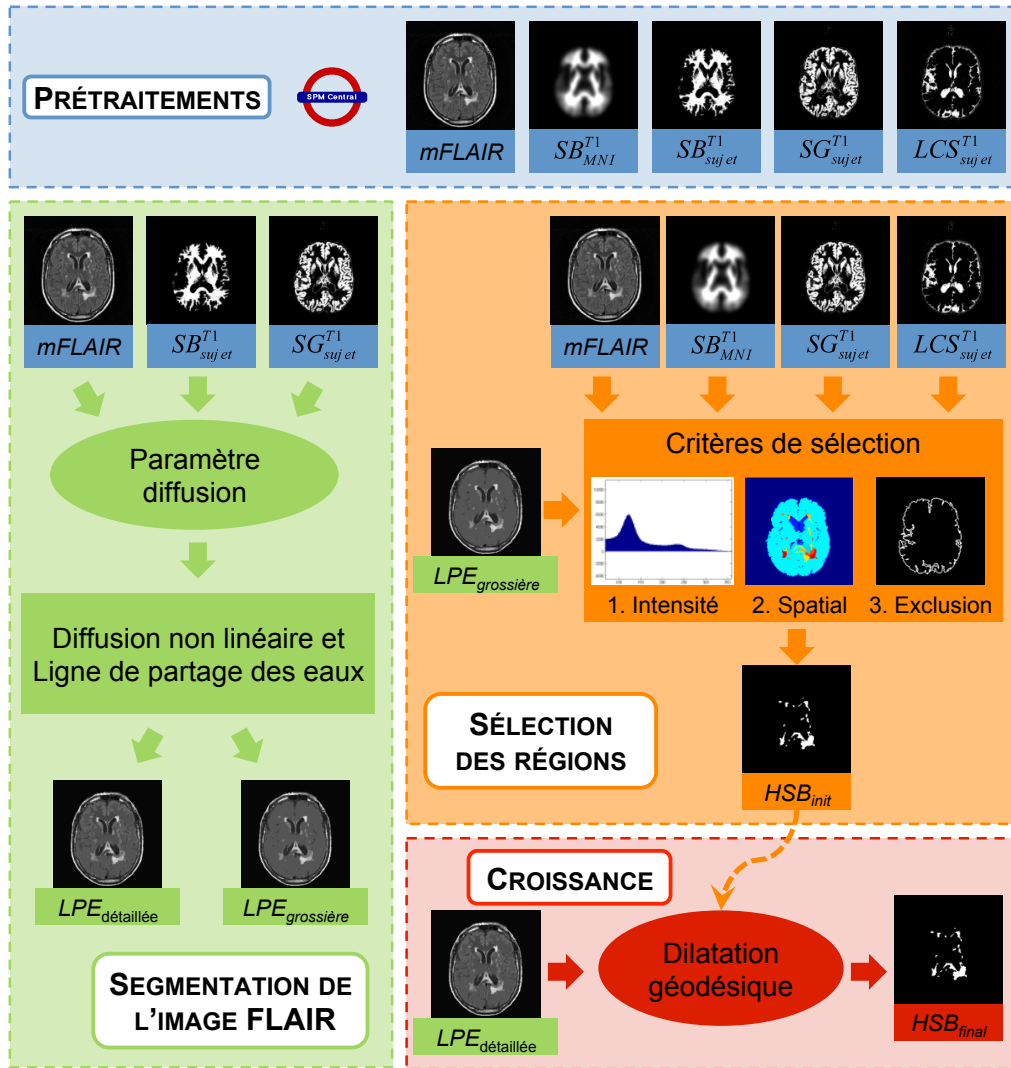


FIGURE 3.1 – Schéma général de la chaîne de traitements WHASA.

BrainVisa (http://brainvisa.info/index_f.html). En particulier, les problématiques de segmentation en tissus (substance grise, substance blanche et liquide cérebrospinal), de correction de biais pour les images IRM et de recalage inter et intramodalités peuvent être traitées par différentes solutions, du moins dans les cas non pathologiques. Nous nous appuyons donc sur ces résultats pour notre problématique de segmentation des HSB.

Parmi les différentes suites logicielles disponibles pour le traitement des images cérébrales, nous avons choisi d'utiliser SPM (*Statistical Parametric Mapping*), l'un des logiciels actuellement les plus utilisés en neuroimagerie. SPM permet en effet d'obtenir de manière rapide et robuste des cartes de probabilité des principaux tissus cérébraux, ainsi que différents

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

recalages. La version utilisée est la version SPM8.

3.2.1 Segmentation unifiée de la T1

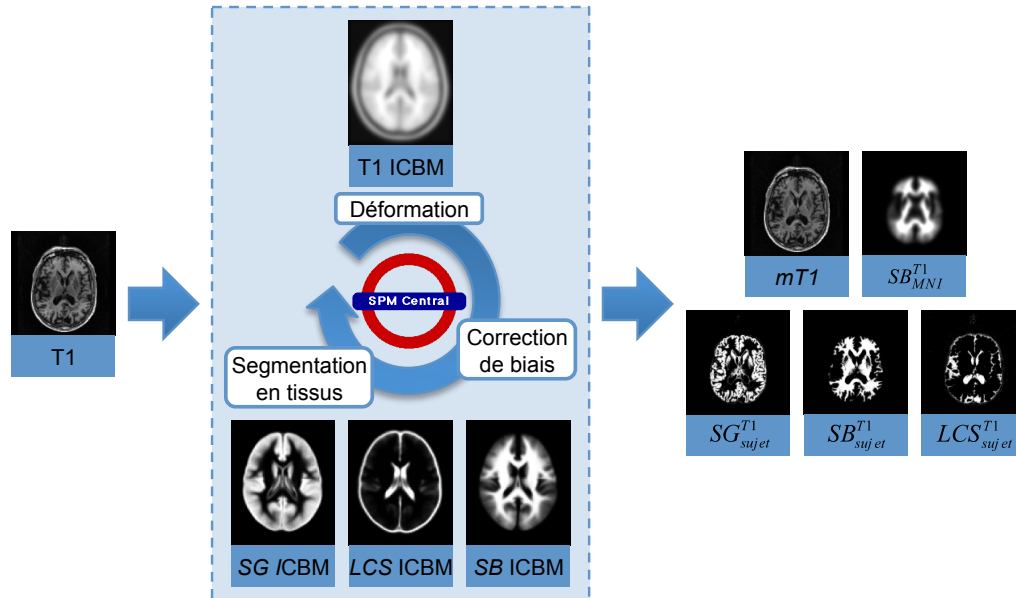


FIGURE 3.2 – Schéma de principe de la segmentation unifiée

La segmentation unifiée [Ashburner & Friston, 2005] est une procédure implémentée dans SPM qui permet le recalage non linéaire d'une image T1 dans un espace de référence (l'espace standard du MNI - *Montreal Neurological Institute*), la segmentation en tissus et la correction de biais de manière unifiée à partir du même modèle génératif. De manière succincte, la segmentation en tissus est obtenue à partir d'un modèle de mélanges de gaussiennes. Un biais multiplicatif et un recalage non linéaire sont ajoutés à cette modélisation. La vraisemblance du modèle total est alors minimisée, aboutissant à la détermination des paramètres optimaux du biais, de la déformation et du modèle de mélange de gaussiennes. Pour plus de détails sur le modèle utilisé pour la segmentation unifiée et les méthodes d'optimisation associées, nous référons le lecteur intéressé à [Ashburner & Friston, 2005].

En particulier, la segmentation unifiée utilise les atlas de l'*International Consortium for Brain Mapping*, comme initialisation pour le calcul de la déformation et de la segmentation en tissus. Plus précisément, il s'agit de l'atlas ICBM452² pour la version SPM8. Cet atlas a été créé à partir

2. http://www.loni.ucla.edu/ICBM/Downloads/Downloads_Atlases.shtml

des acquisitions IRM T1 de 452 sujets sains jeunes. Ces 452 images ont été recalées ensemble dans un espace standard, l'espace MNI (*Montreal Neurological Institute*), afin d'obtenir un modèle moyen d'IRM T1, puis segmentées manuellement en substance grise, substance blanche et liquide cérebrospinal.

Nous avons utilisé le module *New Segment* du logiciel SPM, qui implémente une version améliorée de la segmentation unifiée, en particulier en modélisant de manière plus précise les tissus graisseux et le crâne. Le principe général de l'algorithme demeure néanmoins similaire. Au final, la procédure de segmentation unifiée nous permet d'obtenir :

- une image T1 corrigée de biais, notée $mT1$,
- des cartes de probabilités de substance blanche, de substance grise et de liquide cérebrospinal, notées respectivement SB_{sujet}^{T1} , SG_{sujet}^{T1} et LCS_{sujet}^{T1} ,
- une transformation non linéaire inversible permettant de recalculer l'image du sujet dans l'espace standard MNI, $t_{MNI \rightarrow sujet}$.

Bien que la procédure de segmentation unifiée fournisse généralement de très bons résultats sur des sujets sains ou atrophés, la présence d'HSB, qui n'est pas modélisée, peut altérer la qualité des résultats de SPM [Levy-Cooperman et al., 2008]. Les HSB sont ainsi le plus souvent classées dans SB_{sujet}^{T1} ou SG_{sujet}^{T1} en fonction de leur localisation. Afin de pouvoir corriger par la suite les erreurs de localisation de la substance blanche, nous utilisons la transformation $t_{MNI \rightarrow sujet}$ pour recalculer également l'atlas de substance blanche ICBM dans l'espace de l'image T1 du sujet, que nous noterons SB_{MNI}^{T1} .

3.2.2 Recalage de l'image T1 sur l'image FLAIR

Afin d'amener dans l'espace FLAIR les informations anatomiques obtenues par la segmentation en tissus de l'image T1, il est nécessaire de recalculer l'image T1 sur l'image FLAIR. Ceci est accompli à l'aide de l'outil de corecalage du logiciel SPM. Cette procédure estime la transformation rigide, caractérisée par 3 rotations et 3 translations, maximisant une fonction objectif à déterminer. Dans le cas du recalage intermodal, un choix classique de fonction objectif est celui de l'information mutuelle normalisée.

Une fois cette transformation calculée, nous pouvons appliquer cette

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

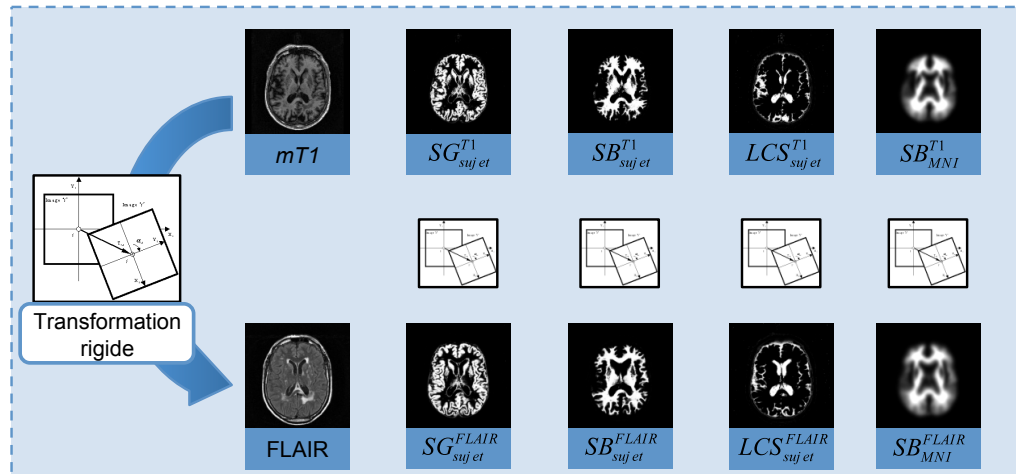


FIGURE 3.3 – Recalage rigide entre l'espace T1 et l'espace FLAIR.

La transformation rigide maximisant l'information mutuelle normalisée entre l'image T1 et l'image FLAIR est calculée à l'aide du module *Coreg* de SPM. La transformation ainsi obtenue est ensuite appliquée aux autres images situées dans l'espace T1 pour les amener dans l'espace FLAIR.

même transformation pour recalcr les cartes de probabilités en tissus sur l'image FLAIR. Nous obtenons de cette manière SG^{FLAIR}_{sujet} , SB^{FLAIR}_{sujet} , LCS^{FLAIR}_{sujet} et SB^{FLAIR}_{MNI} (figure 3.3).

3.2.3 Correction de biais de l'image FLAIR

L'image FLAIR, tout comme l'image T1, est affectée par les inhomogénéités de champ, ce qui peut fortement pénaliser les procédures automatiques de segmentation. Le logiciel SPM permet également de corriger ce biais de manière satisfaisante.

Pour cela, nous avons utilisé à nouveau le module *New Segment* disponible dans la version SPM8. Ce module permet également une prise en compte de plusieurs modalités, permettant ainsi de corriger le biais de l'image FLAIR. Nous n'avons pas utilisé précédemment la version multimodale pour la correction de biais de l'image T1 et la segmentation en tissus à cause de l'importante différence d'épaisseur de coupe entre l'image T1 et l'image FLAIR. L'image FLAIR corrigée de biais est notée *mFLAIR*.

3.2.4 Masque de la région intracrânienne

Un masque de la région intracrânienne dans l'espace de l'image FLAIR peut être directement dérivé de la segmentation en tissus obtenue par la segmentation unifiée en sommant les cartes de probabilités SB_{sujet}^{FLAIR} , SG_{sujet}^{FLAIR} et LCS_{sujet}^{FLAIR} , puis en seuillant la somme à 0,5. L'application d'une fermeture morphologique, suivi d'une ouverture morphologique, toutes deux ayant un disque de 2 voxels de rayon comme élément structurant, permet d'éliminer les groupes isolés de voxels et de combler les trous pouvant exister dans le masque.

3.2.5 Bilan des prétraitements

En résumé, l'ensemble des prétraitements présentés nous permet d'obtenir, dans le même espace que l'image FLAIR originale :

- une image *mFLAIR* corrigée du biais, plus adaptée à une segmentation utilisant des critères d'intensité ;
- des cartes des probabilité de tissus SG_{sujet}^{FLAIR} , SB_{sujet}^{FLAIR} , LCS_{sujet}^{FLAIR} et SB_{MNI}^{FLAIR} . En particulier, un masque de la région intracrânienne est dérivé directement des trois premières cartes de probabilités.

L'ensemble des opérations de segmentation des HSB étant effectué dans l'espace de l'image FLAIR, nous omettrons dans la suite du manuscrit la mention *FLAIR* afin d'alléger les notations.

Il est à noter que le modèle de segmentation unifiée ne permet pas d'intégrer de manière simple la segmentation d'anomalies telles que les HSB. En effet, il apparaît difficile d'ajouter au modèle de la segmentation unifiée une classe aussi variable que celle des HSB. Par ailleurs, la majorité des HSB est difficilement visible sur les images T1 et la définition d'un modèle multimodal n'apporte pas de résultats probants pour la segmentation des HSB [Smart et al., 2011].

3.3 Segmentation de l'image *mFLAIR*

Le FLAIR est une séquence IRM permettant de s'affranchir des effets de volume partiel causés par l'hypersignal du LCS, particulièrement autour des ventricules ou au niveau du cortex cérébral (voir section 1.4). Elle est donc particulièrement adaptée pour la détection des HSB et leur

différentiation avec les tissus environnants. Anbeek et al. [2005] ont également montré, en étudiant de manière exhaustive les combinaisons de séquences IRM, que le FLAIR fournissait la plus grande valeur ajoutée pour la segmentation automatique des HSB. Cependant, un traitement automatisé des images FLAIR est rendu difficile par leur caractère bruité, leur forte sensibilité au mouvement, la grande variabilité présente dans la dynamique des intensités en fonction des acquisitions et les effets de volume partiel engendrés par l'importante épaisseur de coupe³ de cette séquence.

Une représentation simplifiée de l'image *mFLAIR* est obtenue en faisant évoluer l'image *mFLAIR* suivant un processus de diffusion non linéaire. Souvent utilisée dans le cadre du débruitage d'image, la diffusion non linéaire dépend principalement d'un paramètre, dit *paramètre de contraste*. Nous proposons d'estimer ce paramètre de manière à rehausser le contraste des HSB avec les tissus environnants. De plus, nous couplons la diffusion non linéaire avec un algorithme de segmentation par ligne de partage des eaux (LPE), qui nous permet de déterminer un critère d'arrêt au processus de diffusion.

3.3.1 Diffusion non linéaire

La diffusion non linéaire permet de lisser une image en préservant les sauts d'intensité les plus forts. Elle a été introduite dans le domaine du traitement d'images par Perona & Malik [1990]. L'idée principale est de considérer l'image I comme la condition initiale de l'équation aux dérivées partielles suivante :

$$\frac{\partial I}{\partial t} = \text{div} (g(\|\nabla I\|) \nabla I) \quad (3.1)$$

où g est une fonction monotone décroissante, appelée fonction de diffusivité.

3.3.1.1 Fonction de diffusivité

Dans leur article original, Perona & Malik [1990] introduisaient deux fonctions de diffusivité g_1 et g_2 , permettant de rendre le processus de diffusion sensible soit au contraste des bords, soit à la taille des régions.

3. selon les protocoles, entre 2,5 et 8 mm

Par la suite, Black et al. [1998] ont montré le lien entre la diffusion non linéaire et les statistiques robustes et ont introduit une nouvelle fonction de diffusivité g_3 , dérivée de la fonction bicarrée de Tukey (figure 3.4).

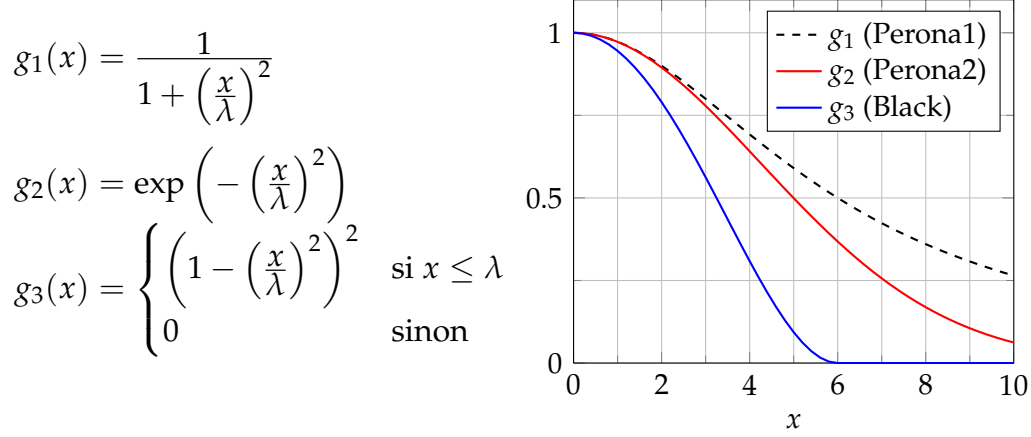


FIGURE 3.4 – Définition et graphes des fonctions de diffusivité g_1 , g_2 et g_3 . (Dans les trois cas, $\lambda = 6$)

Les trois fonctions font intervenir un paramètre noté λ . Pour expliciter le rôle de ce paramètre dans le processus de diffusion, nous nous restreignons dans le paragraphe suivant au cas unidimensionnel. Dans ce cas, l'équation 3.1 devient :

$$\frac{\partial I}{\partial t} = \frac{\partial}{\partial x} \left(g \left(\frac{\partial I}{\partial x} \right) \frac{\partial I}{\partial x} \right)$$

qui met en évidence le rôle de la fonction $\Phi : x \mapsto xg(x)$, appelée *fonction de flux*. En effet, en poursuivant la dérivation, on obtient :

$$\begin{aligned} \frac{\partial I}{\partial t} &= \frac{\partial}{\partial x} \left(\Phi \left(\frac{\partial I}{\partial x} \right) \right) \\ &= \Phi' \left(\frac{\partial I}{\partial x} \right) \frac{\partial^2 I}{\partial x^2} \end{aligned}$$

Lorsque la dérivée Φ' est positive, on retrouve l'équation de la chaleur classique (*forward parabolic*), les contrastes sont atténués lorsque t croît. En revanche, lorsque la dérivée Φ' est négative, les contrastes sont accentués (*backward parabolic*). L'évolution de l'image I au cours de la diffusion dépend donc du signe de la dérivée Φ' . Pour la première fonction g_1

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

proposée dans [Perona & Malik, 1990], le calcul de la dérivée de la fonction de flux $\Phi_1 : x \mapsto xg_1(x)$ montre que Φ_1' s'annule en λ , qui sépare donc les régions où les contrastes seront accentués des régions où les contrastes seront atténués. Pour cette raison, le paramètre λ est souvent appelé également *paramètre de contraste*.

Il faut cependant remarquer que la signification du paramètre λ n'est pas identique selon les fonctions de diffusivité. En effet, le calcul des dérivées des fonctions $\Phi_2 : x \mapsto xg_2(x)$ et $\Phi_3 : x \mapsto xg_3(x)$ montre que leurs maxima respectifs sont atteints pour $x = \frac{\lambda}{\sqrt{2}}$ et $x = \frac{\lambda}{\sqrt{5}}$. Afin de pouvoir comparer l'effet des différentes fonctions de diffusivité, il est donc nécessaire de normaliser les paramètres de diffusion afin que les maxima des fonctions de flux Φ_1 , Φ_2 et Φ_3 soient alignés. Dans la figure 3.5, les pics ont ainsi été alignés sur la valeur $x = 4$ en utilisant $\lambda_1 = 4$, $\lambda_2 = 4\sqrt{2}$ et $\lambda_3 = 4\sqrt{5}$.

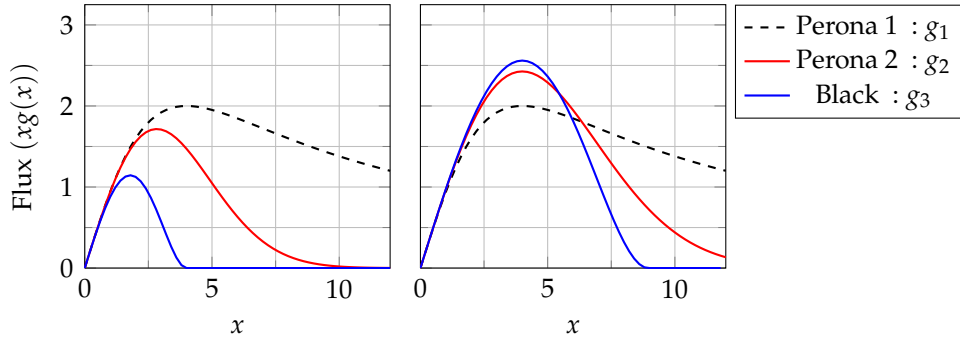


FIGURE 3.5 – Représentation graphique des fonctions de flux Φ_1 , Φ_2 et Φ_3 .

Gauche : dans les trois cas, $\lambda = 4$. **Droite** : la valeur du paramètre de contraste est adapté de manière à aligner les maxima de Φ_1 , Φ_2 et Φ_3 . ($\lambda_1 = 4$, $\lambda_2 = 4\sqrt{2}$ et $\lambda_3 = 4\sqrt{5}$).

La particularité de la fonction g_3 est que la fonction de flux est certes maximale en $\frac{\lambda}{\sqrt{5}}$ mais qu'elle s'annule également pour les valeurs supérieures à λ . Or, en dimension 2, si l'on se place dans un référentiel local tel que $\eta = \frac{\nabla I}{\|\nabla I\|}$ est la direction parallèle au gradient et ξ la direction perpendiculaire, Alvarez et al. [1993] ont montré que l'équation 3.1 se réécrivait :

$$\frac{\partial I}{\partial t} = \Phi'(\|\nabla I\|) I_{\eta\eta} + g(\|\nabla I\|) I_{\xi\xi} \quad (3.2)$$

$I_{\eta\eta}$ (respectivement $I_{\xi\xi}$) désigne la dérivée seconde directionnelle de I dans la direction η (respectivement ξ). L'équation 3.2 signifie que la dif-

fusion le long des lignes de niveau dépend du signe de g alors que la diffusion dans la direction du gradient dépend du signe de Φ' . Ainsi, pour les fonctions g_1 et g_2 , qui sont strictement positives, il existe toujours une diffusion dans la direction ξ , même pour les très grandes valeurs de gradient, c'est-à-dire même entre des régions très contrastées. En revanche, pour la fonction g_3 , lorsque la valeur du gradient est supérieure à λ , la dérivée de la fonction de flux Φ' et la fonction de diffusivité g_3 sont nulles. Le processus de diffusion cesse donc totalement pour les valeurs de gradient supérieures à λ .

3.3.1.2 Discrétisation et résolution de l'équation de diffusion non linéaire

A cause du caractère non linéaire de la fonction de diffusivité g , la résolution de l'équation de diffusion non linéaire (équation 3.1) est effectuée numériquement par un schéma itératif. La discrétisation spatiale est déjà obtenue puisque les images sont décomposées en pixels. La discrétisation temporelle est la plupart du temps obtenue à l'aide d'un schéma explicite, c'est-à-dire que l'expression de I au temps $t + \Delta t$ ne dépend que de la valeur de I lors des itérations précédentes.

$$I(t + \Delta t) = I(t) + \Delta t \sum_{l=1}^m A_l(I(t))$$

où Δt est le pas de temps, m le nombre de dimensions de l'image et A_l la discrétisation par différences finies de la dérivée de I dans la direction l . En particulier, ce type de schéma explicite est stable uniquement si $\Delta t < \frac{h^2}{2m}$, où h est le pas de discrétisation spatiale. Dans le cas de coupes IRM de dimension 2, il est donc nécessaire que $\Delta t < 0,25$.

Plusieurs autres schémas de discrétisation ont été proposés [Weickert et al., 1998 ; Mikula & Ramarosy, 2001 ; Duarte-Carvajalino et al., 2007]. Une revue particulièrement complète concernant les aspects numériques de la diffusion non linéaire pourra être trouvée dans Weickert [1998]. Nous nous sommes restreints au schéma explicite présenté ci-dessus et nous sommes appuyés sur une toolbox disponible librement sur Internet⁴. Cette toolbox ne proposant initialement que les fonctions de diffusivité g_1

4. <http://www.mathworks.com/matlabcentral/fileexchange/3710-nonlinear-diffusion-toolbox>

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

et g_2 , nous y avons ajouté la fonction de diffusivité g_3 . En ce qui concerne la discrétisation temporelle, nous utiliserons par la suite $\Delta t = 0, 1$.

3.3.1.3 Evolution de la diffusion non linéaire selon la fonction de diffusivité

La figure 3.6 montre l'évolution d'une coupe IRM pour les trois fonctions de diffusivité où les paramètres de diffusion ont été ajustés de manière à ce que les maxima des fonctions de diffusivité soient alignés. On voit que la fonction g_1 lisse les discontinuités au bout d'un plus petit nombre d'itérations que g_2 et g_3 . Le résultat de la diffusion non linéaire obtenu avec g_1 sera donc très sensible au nombre d'itérations. En particulier, les petites lésions sont effacées au bout d'un petit nombre d'itérations. L'utilisation des fonctions de diffusivité g_2 et g_3 , en revanche, permet de rendre le processus de diffusion relativement robuste au nombre d'itérations.

En ce qui concerne les fonctions de diffusivité g_2 et g_3 , le comportement apparaît visuellement similaire même si la fonction g_3 semble conserver plus de détails et produire des bords plus francs. Par ailleurs, le comportement de la fonction g_3 est également plus facile à interpréter puisque la valeur du paramètre λ utilisé dans l'équation indique la valeur seuil pour lequel le processus de diffusion cesse complètement. C'est pourquoi nous utiliserons la fonction g_3 par la suite.

3.3.2 Paramètre de contraste

Comme le souligne la section précédente, un paramètre clé de la diffusion non linéaire avec la fonction de diffusivité de Black g_3 est le paramètre de contraste λ . Il doit refléter la valeur maximale des discontinuités que l'on souhaite effacer. Dans notre cas, nous cherchons à renforcer le contraste entre les HSB et les tissus environnants et à affaiblir les contrastes entre la substance grise et la substance blanche. Il apparaît donc nécessaire de fixer le contraste moyen entre la substance grise et la substance blanche comme valeur pour λ .

Afin de définir ce contraste moyen, nous utilisons les segmentations en tissus obtenues au cours des prétraitements. Pour chaque sujet, dans l'espace FLAIR, les cartes de probabilité de substance blanche SB_{sujet} et de substance grise SG_{sujet} sont ainsi seuillées à 0,5. Les masques binaires

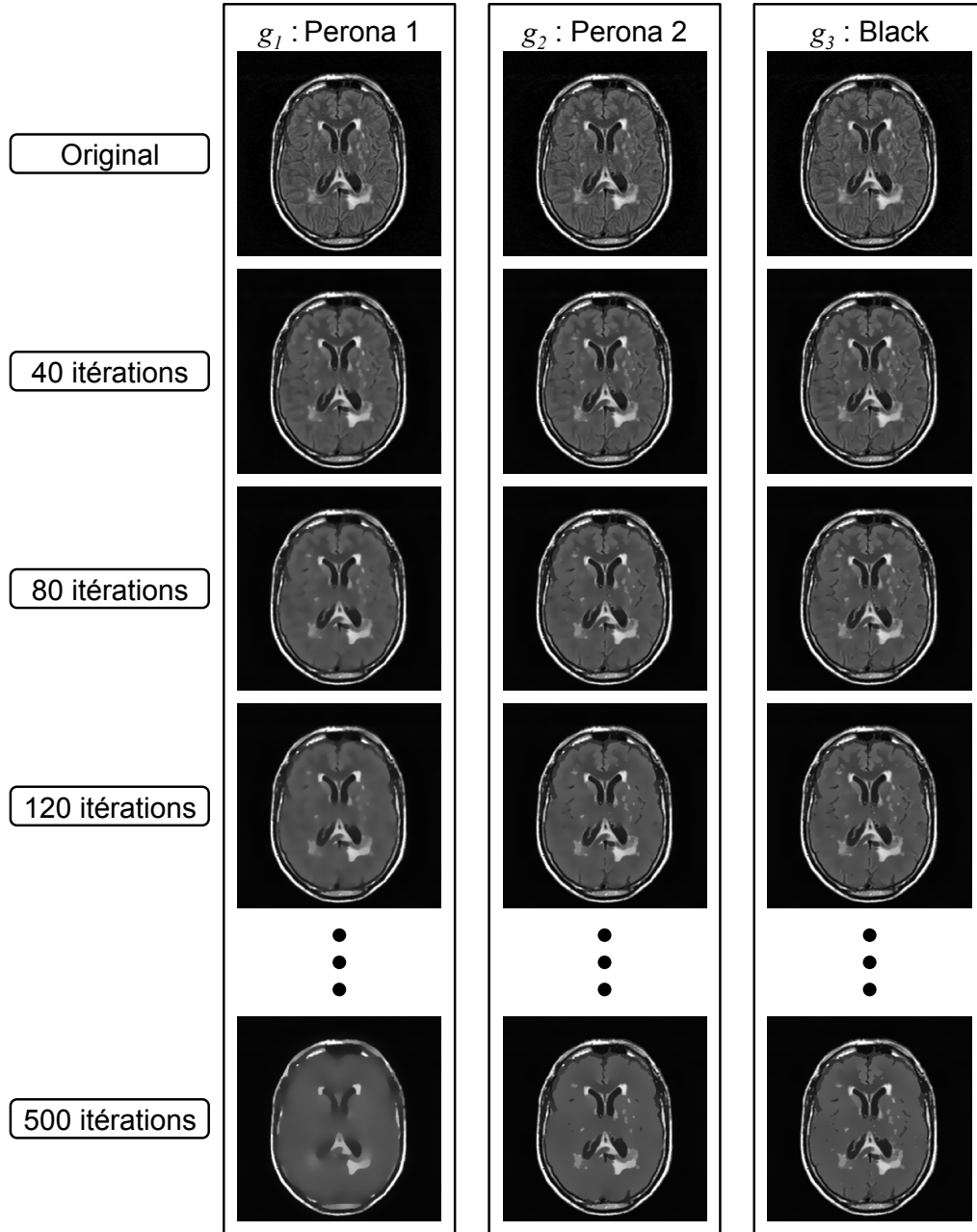


FIGURE 3.6 – Comparaison du processus de diffusion non linéaire pour les trois fonctions de diffusivité g_1 , g_2 et g_3 pour différents nombre d'itérations.

Les paramètres utilisés sont $\lambda_1 = 10$, $\lambda_2 = 10\sqrt{2}$ et $\lambda_3 = 10\sqrt{5}$ et $\Delta t = 0,1$.

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

ainsi obtenus sont dilatés par un élément structurant 2D de rayon 1 voxel (connectivité 4). L'intersection entre les deux masques dilatés fournit une estimation de l'interface entre la substance grise et la substance blanche $SB|SG$ (figure 3.7). Pour chaque sujet, nous utilisons alors comme valeur du paramètre de contraste la moyenne des valeurs du gradient de l'image $mFLAIR$ sur $SB|SG$. Nous notons cette valeur moyenne λ_{est} .

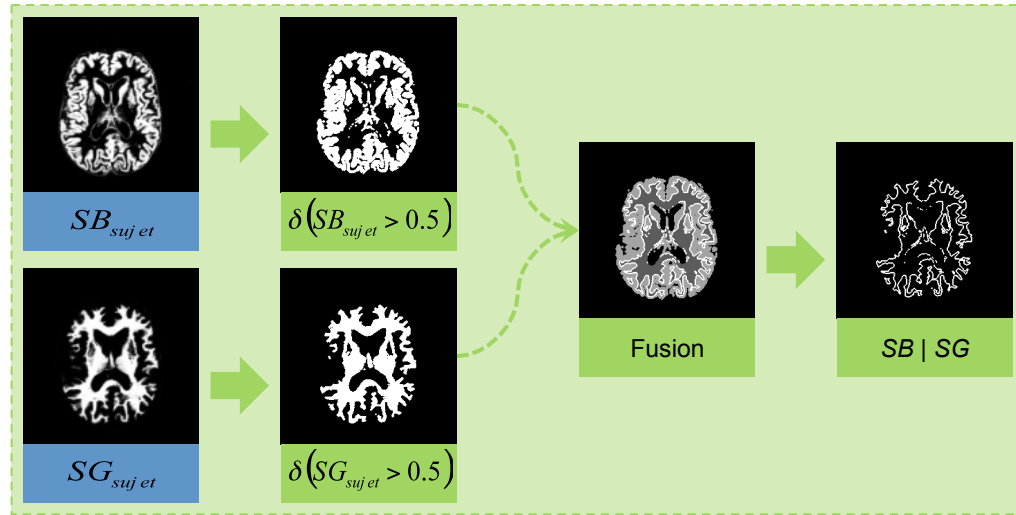


FIGURE 3.7 – Détermination de l'interface substance blanche/substance grise ($SB|SG$) dans l'espace FLAIR.

La figure 3.8 présente l'effet de la diffusion non linéaire pour différentes valeurs du paramètre de contraste. La valeur λ_{est} permet de renforcer le contraste à la fois des grosses lésions périventriculaires et des plus petites lésions isolées, tout en éliminant la majeure partie du contraste entre substance blanche et substance grise. Une valeur plus faible conserverait certains contrastes entre substance blanche et substance grise alors qu'une valeur plus importante aurait tendance à lisser de manière trop importante et donc à faire disparaître de nombreuses HSB.

3.3.3 Ligne de partage des eaux

Le processus de diffusion non linéaire décrit précédemment conduit à lisser de manière importante les zones où le contraste est inférieur à λ_{est} et à renforcer le contraste des interfaces. Ce type de résultats est une étape préliminaire particulièrement adaptée pour une segmentation à l'aide de l'algorithme de ligne de partage des eaux (LPE). Elle fait en

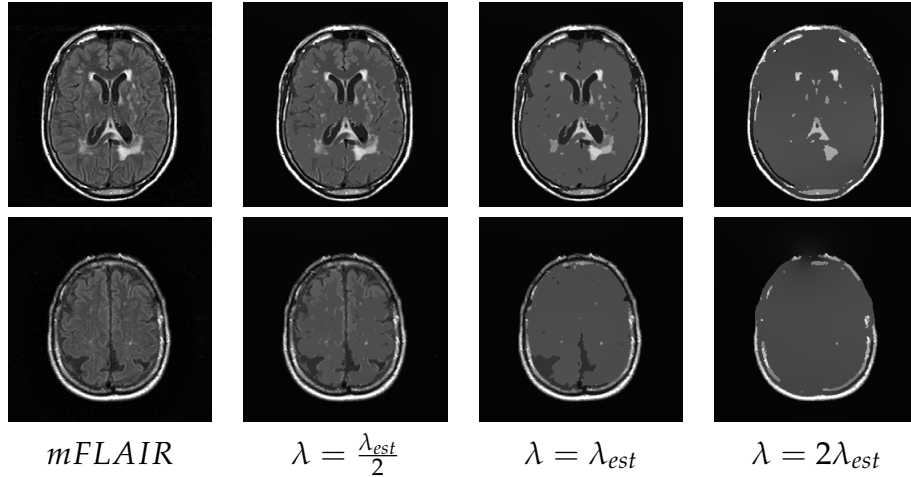


FIGURE 3.8 – Effet de la diffusion non linéaire sur une image *mFLAIR* pour différentes valeurs du paramètre de contraste ($\Delta t = 0, 1, 10000$ itérations).

effet disparaître des minima locaux, qui engendreraient autrement une segmentation excessivement bruitée, comme détaillé ci-dessous.

La LPE est un outil classique issu de la morphologie mathématique, introduit dans [Beucher & Lantuejoul, 1979]. L'idée sur laquelle repose la LPE est de considérer une image en niveaux de gris, c'est-à-dire une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$, comme un relief topographique. Il s'agit ensuite d'imaginer que l'on immerge ce relief en autorisant l'eau à pénétrer uniquement par les minima du relief, chaque minimum fournissant ainsi un bassin. Au fur et à mesure de l'immersion, les bassins vont se remplir jusqu'à rencontrer les bassins voisins. Les points où deux bassins se rejoignent forment alors la ligne de partage des eaux (figure 3.9). Une fois que l'ensemble du relief a été immergé, on obtient alors une segmentation de l'image en régions. Deux points sont à noter concernant la LPE :

- les points appartenant à une ligne de partage des eaux ne sont pas attribués à une région, ils forment les frontières de la segmentation. Une solution simple concernant ces points est de les affecter à la région dont l'intensité moyenne est la plus proche de la leur.
- la localisation des lignes de partage des eaux dépendant de la distribution de l'ensemble des minima de l'image, la LPE n'est pas un concept local.

Classiquement, la LPE est calculée sur le gradient de l'image que l'on souhaite segmenter.

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

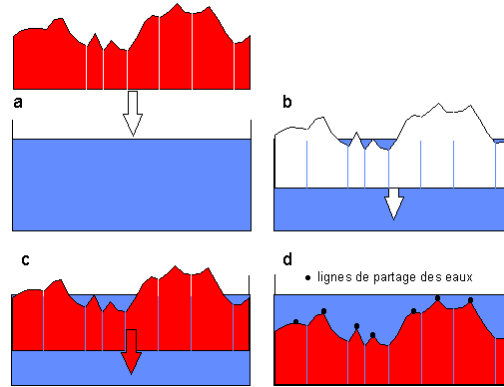


FIGURE 3.9 – Principe de la ligne de partage des eaux (Source : <http://www.tsi.telecom-paristech.fr/pages/enseignement/ressources/beti/lpe2/lpe/lpe.htm>)

La figure 3.10 montre l'effet de la LPE sur une région agrandie de l'image *mFLAIR*. On observe un résultat classique de sursegmentation due à la présence d'un grand nombre de minima locaux lié à la texture et au bruit de l'image. L'effet de la diffusion non linéaire permet d'éliminer la plupart de ces minima tout en conservant les transitions importantes concernant les HSB. La figure 3.11 montre le résultat de la LPE sur l'image diffusée ($\Delta t = 0, 1, 10000$ itérations). Le problème de sursegmentation est ainsi résolu.

3.3.4 Critère d'arrêt

L'algorithme de ligne de partage des eaux permet également de régler le temps d'arrêt de la diffusion, c'est-à-dire le nombre d'itérations au bout duquel le processus de diffusion est arrêté. En effet, comme le montre la figure 3.12, la segmentation obtenue par LPE se stabilise avec la disparition des minima locaux. Il est donc possible d'instaurer un critère d'arrêt basé sur la stabilité de la segmentation obtenue par LPE.

Fixons $\Delta t = 0, 1$ le pas de temps pour la résolution de l'équation de diffusion non linéaire et n le nombre d'itérations effectuées. L'image diffusée à l'itération n est donc notée $I_{n\Delta t}$. $I_{n\Delta t}$ est segmentée par LPE, et le résultat de la segmentation est notée LPE_n . A l'itération suivante, la segmentation de $I_{(n+1)\Delta t}$ fournit donc LPE_{n+1} . Tant que la segmentation en régions LPE_n est différente de LPE_{n+1} , $I_{n\Delta t}$ est diffusée pour une itération supplémentaire. Le processus de diffusion s'arrête lorsque $LPE_n = LPE_{n+1}$. En pratique, il est inutile d'effectuer la comparaison entre les LPE à toutes

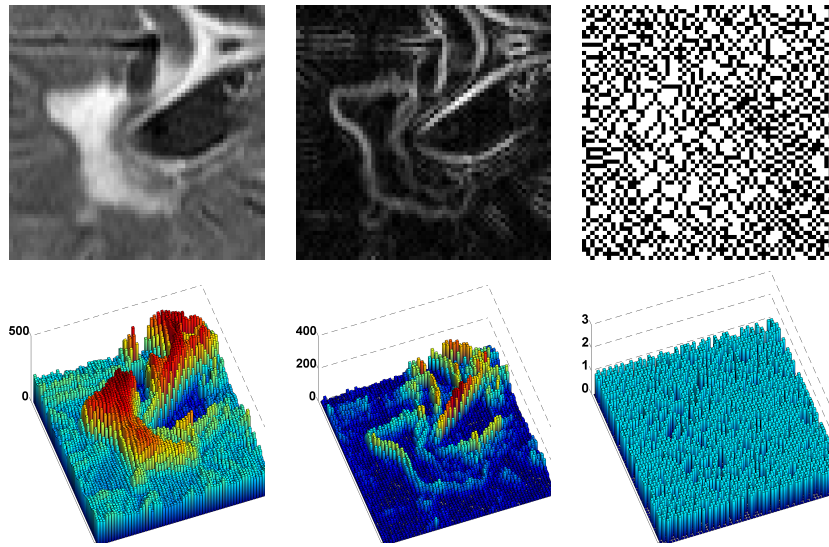


FIGURE 3.10 – Illustration de l'effet de la LPE sur une image *mFLAIR* non diffusée.

Ligne du haut : Agrandissement d'une partie d'une image *mFLAIR*. Norme du gradient de l'image *mFLAIR*. Résultat de la LPE effectuée sur l'image Gradient. **Ligne du bas** : Similaire à la ligne du haut, mais visualisées en relief. La hauteur du pic représente l'intensité du pixel.

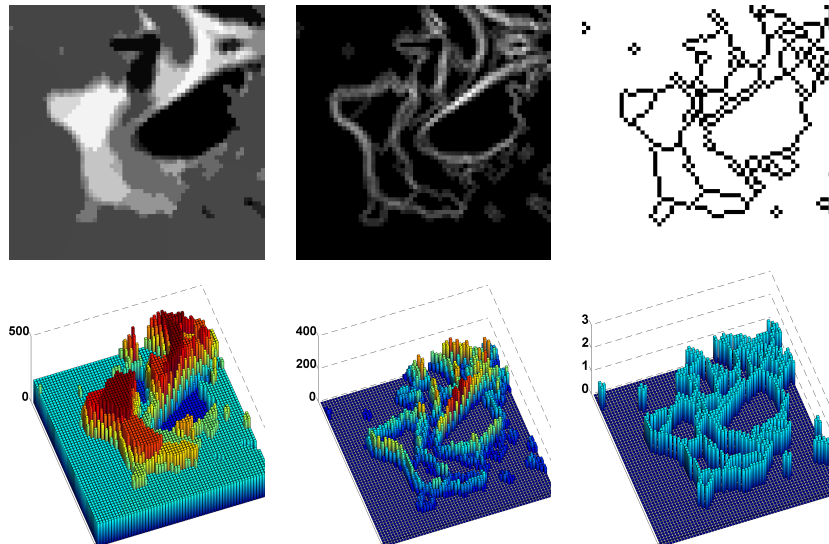


FIGURE 3.11 – Illustration de l'effet de la LPE sur une image diffusée ($\Delta t = 0, 1, 10000$ itérations).

Ligne du haut : Agrandissement d'une partie d'une image *mFLAIR* diffusée ($\Delta t = 0, 1, 10000$ itérations). Norme du gradient de l'image diffusée. Résultat de la LPE effectuée sur l'image Gradient. **Ligne du bas** : Similaire à la ligne du haut, mais visualisées en relief. La hauteur du pic représente l'intensité du pixel.

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

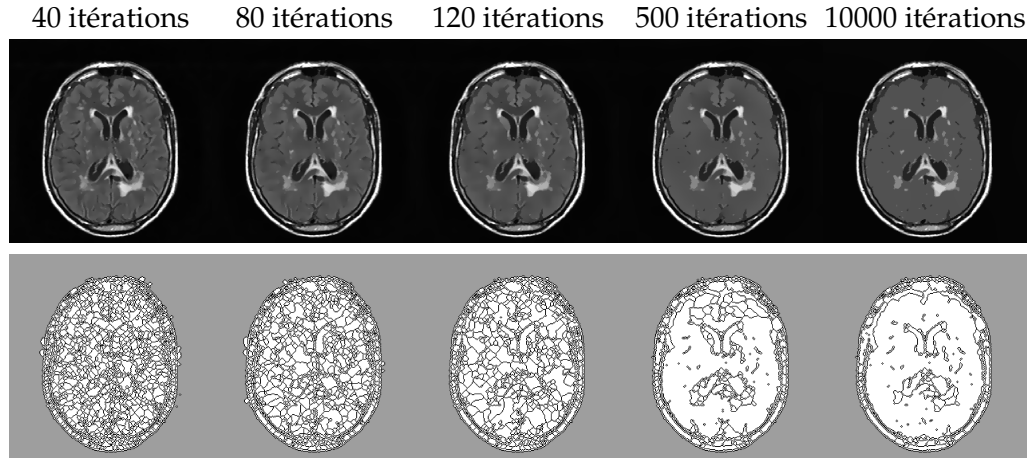


FIGURE 3.12 – Résultat de la LPE sur une image diffusée avec différents nombres d'itérations ($\Delta t = 0, 1$).

les itérations, ce qui augmente le temps de calcul alors que l'évolution de l'image est minime entre deux itérations. La comparaison des segmentations obtenues par LPE est donc effectuée toutes les 100 itérations (figure 3.13).

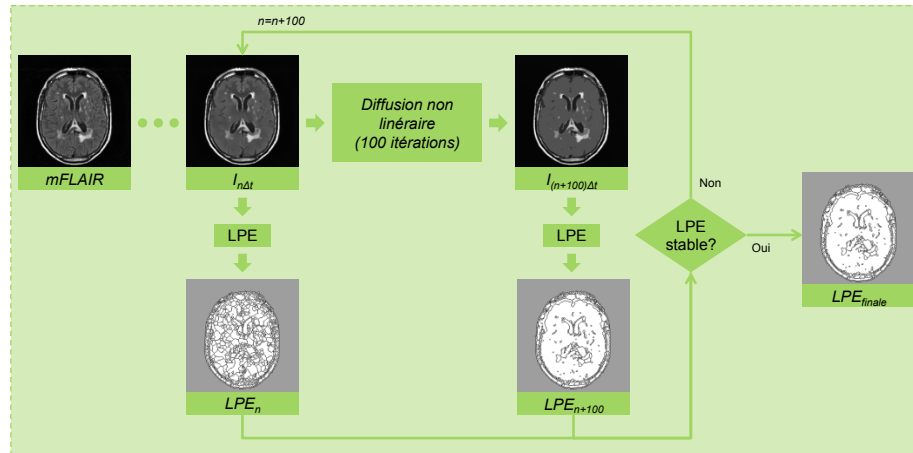


FIGURE 3.13 – Critère d'arrêt pour le processus de segmentation de l'image *mFLAIR*.

L'image *mFLAIR* est la condition initiale de l'équation de diffusion non linéaire. Le processus de segmentation fait alterner une étape de diffusion non linéaire de 100 itérations (avec $\Delta t = 0, 1$) avec une étape de LPE. A l'itération n , l'image $I_{n\Delta t}$ est segmentée par LPE, le résultat étant LPE_n . Puis $I_{n\Delta t}$ est diffusée pendant 100 itérations supplémentaires pour obtenir $I_{n+100\Delta t}$. La segmentation par LPE de $I_{n+100\Delta t}$, notée LPE_{n+100} , est alors comparée à LPE_n . Si elles sont identiques, le processus de segmentation s'arrête et la segmentation finale est obtenue par $LPE_{finale} = LPE_n = LPE_{n+100}$. Autrement, $I_{(n+100)\Delta t}$ est considérée comme la nouvelle image courante et est diffusée pour 100 itérations supplémentaires.

3.3.5 Fusion de régions adjacentes

Bien que la diffusion non linéaire permette de résoudre en grande partie le problème de sursegmentation, nous effectuons une étape supplémentaire de fusion des régions pour augmenter la robustesse. L'intensité *mFLAIR* moyenne sur les régions issues de la LPE est comparée à celle des régions adjacentes. Si la différence est inférieure à λ_{est} , ces régions sont alors fusionnées.

3.3.6 Segmentation détaillée de l'image *mFLAIR*

Le résultat de l'ensemble des étapes amenant à LPE_{finale} est essentiellement déterminé par le paramètre de contraste utilisé dans la diffusion non linéaire. En faisant varier la valeur du paramètre de contraste, il est ainsi possible de créer une segmentation adaptée à différents niveaux de finesse, de plus en plus détaillée lorsque λ diminue. La valeur λ_{est} a été calculée afin d'obtenir un niveau de détail correspondant au niveau de contraste des HSB. Cependant, nous verrons par la suite qu'il peut être utile de disposer d'une segmentation plus détaillée de l'image *mFLAIR* (section 3.5).

Pour ce faire, nous relançons la procédure de segmentation de l'image *mFLAIR* en prenant cette fois pour valeur du paramètre de contraste $\lambda_{détails} = \frac{\lambda_{est}}{2}$. Tous les autres paramètres sont identiques ($\Delta t = 0,1$, alternance de 100 itérations entre les LPE pour déterminer le critère d'arrêt, fusion des régions adjacentes...). Nous notons $LPE_{grossière}$ le résultat de la segmentation de l'image *mFLAIR* avec le paramètre λ_{est} et $LPE_{détaillée}$ celui obtenu avec le paramètre $\frac{\lambda_{est}}{2}$. La figure 3.14 montre les résultats ainsi obtenus. Pour $LPE_{détaillée}$, les régions sont plus petites et la délimitation des HSB plus précise. En revanche, certains détails inutiles sont conservés.

3.3.7 Bilan de la segmentation de l'image *mFLAIR*

Le processus de segmentation de l'image *mFLAIR* prend donc comme entrées l'image corrigée du biais *mFLAIR* et les cartes de probabilité SB_{sujet} et SG_{sujet} . Ces deux cartes permettent d'estimer une valeur du paramètre de contraste λ_{est} . Des étapes de diffusion non linéaire sont alternées toutes les 100 itérations avec une segmentation par LPE. Ce processus est arrêté lorsque deux LPE consécutives sont identiques. En

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

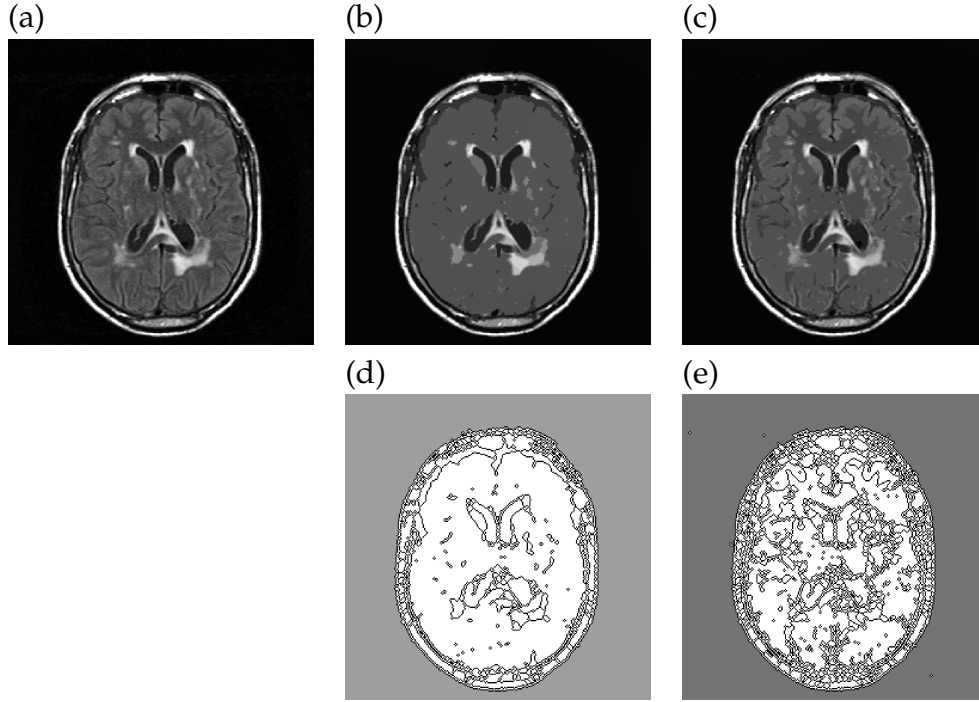


FIGURE 3.14 – Comparaison de $LPE_{grossière}$ et de $LPE_{détaillée}$. (a) image *mFLAIR*. (b) image diffusée au moment de l’arrêt de la procédure de segmentation ($\Delta t = 0, 1, \lambda = \lambda_{est}$). (c) image diffusée au moment de l’arrêt de la procédure de segmentation ($\Delta t = 0, 1, \lambda = \lambda_{est}/2$). (d) $LPE_{grossière}$. (e) $LPE_{détaillée}$.

effectuant l’ensemble de ces étapes pour deux valeurs du paramètre de contraste (λ_{est} et $\lambda_{est}/2$), nous obtenons deux segmentations de l’image *mFLAIR* avec différents niveaux de détails :

- $LPE_{grossière}$: une segmentation de l’image *mFLAIR* peu détaillée mais où le contraste des HSB avec les tissus environnants a été fortement réhaussé tout en supprimant le contraste entre substance blanche et substance grise.
- $LPE_{détaillée}$: une segmentation de l’image *mFLAIR* plus détaillée, mais où certains détails inutiles sont conservés et où le contraste des HSB avec les tissus environnants est moins fort.

3.4 Sélection des régions sur $LPE_{grossière}$

La segmentation $LPE_{grossière}$ obtenue à la fin l’étape précédente est une image constante par morceaux, séparée en régions. La différence entre les intensités moyennes de deux régions adjacentes est toujours supérieure à λ_{est} , l’estimation du contraste moyen SB/SG. L’objectif de la section

suivante est de parvenir à sélectionner parmi ces régions uniquement celles qui correspondent aux hyperintensités de la substance blanche. Pour chaque région, il est alors possible de calculer plusieurs caractéristiques, telles que l'intensité moyenne, la distribution des intensités, la localisation, la taille... Ces caractéristiques n'étaient soit pas disponibles au niveau du voxel (par exemple, la taille), soit moins informatives (l'intensité moyenne). Pour ce dernier cas, il aurait évidemment été possible de considérer la valeur moyenne de l'intensité dans un voisinage, mais cela aurait conduit à mélanger des informations provenant de tissus sains et de tissus lésés. Nous décrivons trois critères simples utilisés pour la sélection des régions HSB.

3.4.1 Critère d'intensité

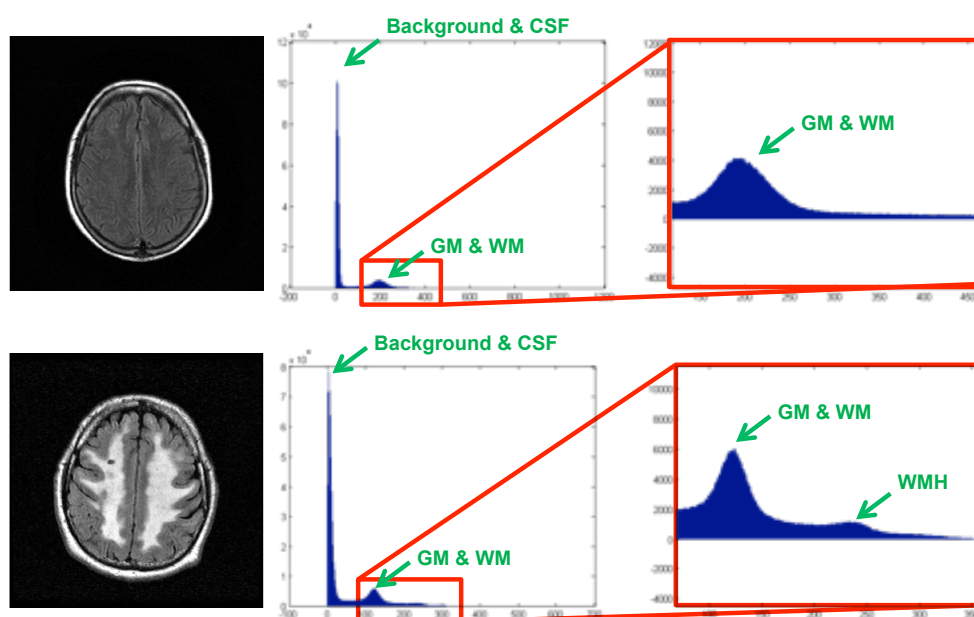


FIGURE 3.15 – Histogramme des intensités d’images *mFLAIR* avec différentes charges lésionnelles.

Ligne du haut : image *mFLAIR*, histogramme des intensités et zoom de l’histogramme sur le mode représentant les tissus cérébraux d’un sujet ne présentant aucune HSB. **Ligne du bas :** image *mFLAIR*, histogramme des intensités et zoom de l’histogramme sur le mode représentant les tissus cérébraux d’un sujet présentant de très importantes HSB.

Le premier critère est celui de l’intensité. Nous voulons ici déterminer un seuil permettant d’indiquer quelle région peut être considérée comme hyperintense par rapport aux tissus normaux, que sont la substance blanche et la substance grise. Ces deux tissus ont été la plupart du

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

temps fusionnés par le processus de diffusion non linéaire. A la fin de ce processus, les régions correspondantes ont donc une intensité moyenne égale à la moyenne des intensités de ces deux types de tissus.

Il est possible d'estimer grossièrement cette valeur d'après l'histogramme de l'image *mFLAIR*. En effet, l'histogramme des intensités d'une image *mFLAIR* (lorsque les valeurs faibles, correspondant au fond et au liquide cébrospinal, ne sont pas considérées) est constitué d'un seul mode principal. Deux modes distincts pour la substance grise et la substance blanche sont habituellement observés pour les images pondérées en T1. La valeur du mode principal de l'histogramme nous fournit ainsi rapidement une estimation de l'intensité moyenne des tissus normaux, que nous notons $\tilde{i}$. La figure 3.15 montre également que la détermination de ce mode est robuste à la présence, même importante, d'HSB.

Une fois $\tilde{i}$ déterminé, il nous faut indiquer quel écart par rapport à cette valeur caractérise les régions hyperintenses. Grâce à l'étape de fusion des régions, nous savons que les intensités moyennes des régions adjacentes sont séparées d'au moins λ_{est} . Une région comprenant des tissus normaux aura son intensité moyenne à peu près égale à $\tilde{i}$. Les régions un peu plus hyperintenses auront donc une intensité moyenne d'au moins $\tilde{i} + \lambda_{est}$ mais cela peut correspondre à des régions composées uniquement de substance grise par exemple plutôt qu'à des régions réellement hyperintenses. Nous définissons alors le seuil permettant de séparer les régions hyperintenses comme :

$$s = \tilde{i} + 2\lambda_{est}$$

Bien que ce critère ait été choisi de manière empirique, il faut noter que la diffusion non linéaire a permis de renforcer les contrastes entre régions. $LPE_{grossière}$ est ainsi beaucoup moins sensible au choix de ce seuil sur les intensités que l'image *mFLAIR* originale (voir section 4.5.5).

3.4.2 Critère de localisation

Le second critère pour déterminer les régions correspondant aux HSB est celui de la localisation. Nous voulons ici retenir les régions situées principalement dans la substance blanche. Cependant, les cartes de probabilité en tissus SB_{sujet} , obtenues à partir de SPM, sont parfois imprécises, notamment au niveau des HSB. En effet, la segmentation unifiée s'effec-

tuant sur l'image pondérée en T1, les HSB apparaissent hypointenses et peuvent être fréquemment affectées à la substance grise plutôt qu'à la substance blanche.

Afin de pallier à cet inconvénient, nous utilisons plutôt l'*a priori* de substance blanche SB_{MNI} , obtenu à partir du template ICBM452 recalé dans l'espace de l'image FLAIR du sujet. A chaque région, nous associons cette fois la probabilité moyenne d'appartenance à la substance blanche. La figure 3.16 montre l'effet des différentes cartes de probabilité de substance blanche sur les probabilités moyennes associées aux régions. Suite à la mauvaise classification sur l'image T1, une partie importante des HSB n'est pas comprise dans SB_{sujet} alors que SB_{MNI} la couvre partiellement.

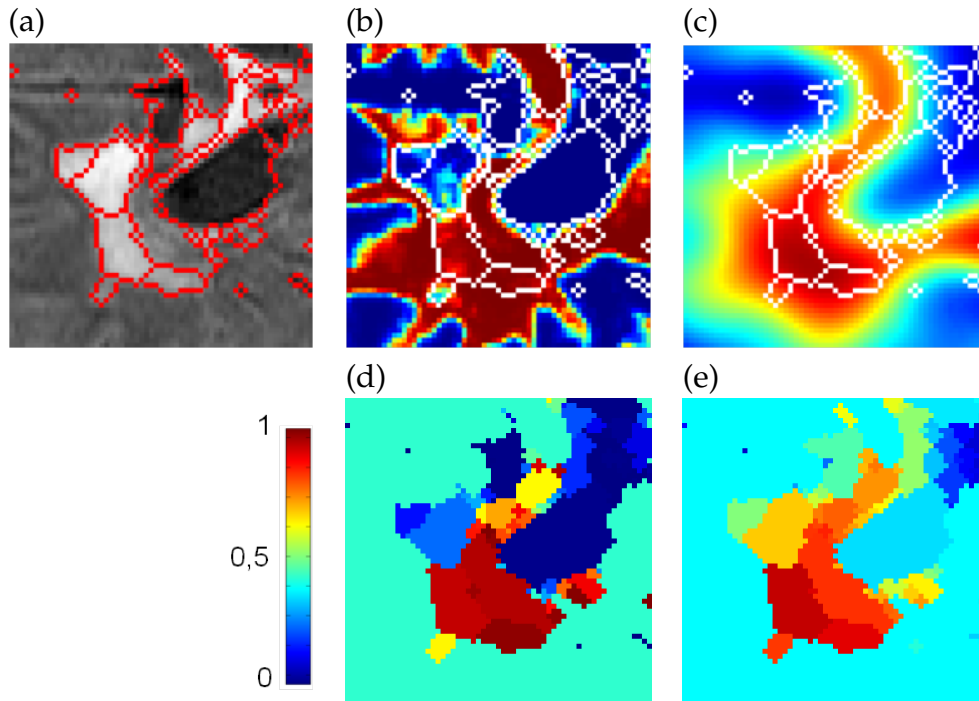


FIGURE 3.16 – Effet de la carte de probabilité utilisée pour déterminer la probabilité moyenne d'une région.

(a) Agrandissement de l'image $mFLAIR$ non diffusée. En rouge sont superposés les bords de $LPE_{grossière}$. (b) Superposition de SB_{sujet} et des bords de $LPE_{grossière}$ (en blanc). (c) Superposition de SB_{MNI} et des bords de $LPE_{grossière}$ (en blanc). (d) A chacune des régions de $LPE_{grossière}$ est attribuée la probabilité moyenne de SB_{sujet} sur la région. (e) A chacune des régions de $LPE_{grossière}$ est attribuée la probabilité moyenne de SB_{MNI} sur la région.

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

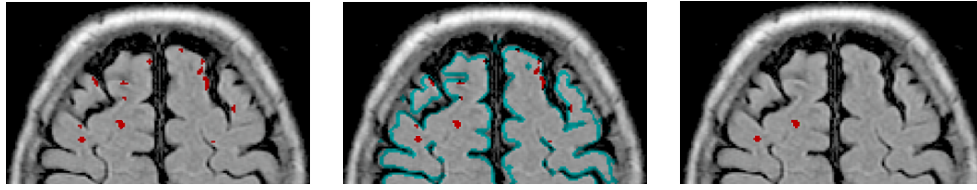


FIGURE 3.17 – Procédure d'élimination des régions corticales.

Gauche : Sur certaines coupes, un grand nombre de régions corticales ne correspondant pas à des HSB peuvent être détectées. **Milieu** : Les régions que l'on souhaite éliminer sont presque toutes connexes à l'interface $(SG|LCS)_{cortex}$ (en bleu cyan). **Droite** : Seules les régions non connexes à l'interface $SG|LCS$ sont conservées.

3.4.3 Elimination de régions corticales

L'utilisation des deux critères d'intensité et de localisation nous permet d'obtenir une première estimation des HSB. Cependant, un certain nombre de régions ne correspondant pas à des HSB peuvent également vérifier ces deux critères. Il s'agit en particulier de régions corticales qui peuvent être hyperintenses en raison des plissements du cortex qui entraînent des effets de volume partiel (figure 3.17).

Le troisième critère utilisé est donc un critère d'exclusion. Nous imposons que les régions sélectionnées ne soient pas connexes à l'interface $SG|LCS$ au niveau des circonvolutions du cortex. Ce critère est particulièrement adapté pour le cas des HSB, puisque celles-ci épargnent de manière systématique les fibres en U sous-corticales et ne sont donc jamais localisées près du ruban cortical [Pantoni & Garcia, 1997; Young et al., 2008].

L'interface $SG|LCS$ est définie de la même manière que l'interface $SB|SG$ (section 3.3.2) : SG_{sujet} et LCS_{sujet} sont seuillées à 0,5, les masques binaires ainsi obtenus sont dilatés par un élément structurant 2D de rayon 1 voxel (connectivité 4) et l'intersection entre les deux masques dilatés fournit une estimation de l'interface $SG|LCS$. Afin de nous restreindre uniquement au niveau des circonvolutions du cortex, une procédure particulière a été implémentée à partir du masque cérébral et de l'interface $SG|LCS$.

Tout d'abord, une couronne indiquant grossièrement la zone corticale est définie à partir du masque cérébral. Celui-ci est érodé par un élément structurant circulaire (érosion 2D) de rayon 10 voxels. L'érodé est soustrait du masque cérébral initial afin d'obtenir une couronne de 10 voxels indiquant grossièrement la zone corticale. Nous sélectionnons ensuite

toutes les composantes connexes de l'interface $SG|LCS$ qui intersectent cette couronne (figure 3.18).

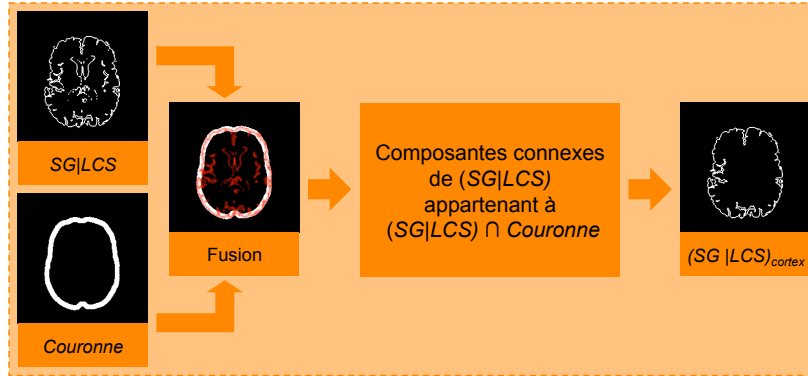


FIGURE 3.18 – Détermination de l'interface $(SG|LCS)_{cortex}$

3.4.4 Bilan de la sélection des régions

En utilisant les trois critères présentés ci-dessus, nous sélectionnons alors les régions :

- dont l'intensité moyenne est supérieure au critère d'intensité s ,
- dont la probabilité moyenne d'appartenance à la substance blanche sur la région est supérieure à 0,5,
- et qui ne sont pas connectés à l'interface $(SG|LCS)_{cortex}$.

Nous obtenons ainsi une carte binaire des HSB, que nous appellerons HSB_{init} .

3.5 Etape de croissance

3.5.1 Insuffisances de la segmentation HSB_{init}

Les critères détaillés dans la section précédente pour la sélection des régions à partir de $LPE_{grossière}$ peuvent néanmoins dans certains cas aboutir à ne pas détecter certaines HSB. Deux types de problèmes principaux ont été identifiés :

- sur certaines coupes, notamment dans les cas de très importantes HSB, le paramètre de diffusion ne semble pas optimal. Le critère d'élimination des lésions corticales aboutit ainsi à éliminer une HSB qui a été fusionnée, suite à une diffusion trop forte, avec une circonvolution cérébrale (figure 3.19).

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

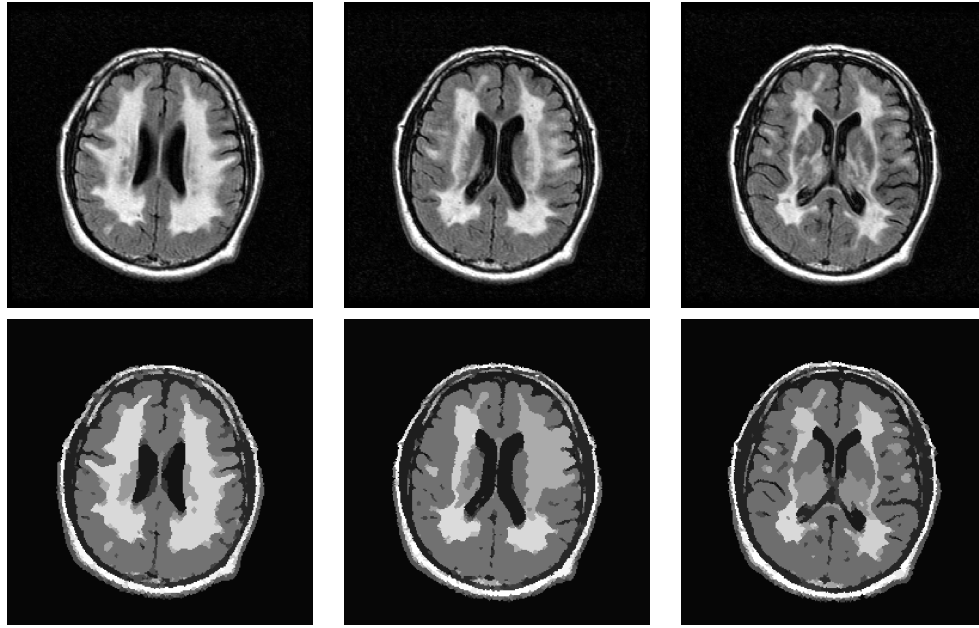


FIGURE 3.19 – Echec du processus de diffusion sur certaines coupes.
Haut : Trois coupes consécutives de l'image *mFLAIR* non diffusée d'un patient CADA-SIL. **Bas** : $LPE_{grossière}$ pour ces trois mêmes coupes. Pour une meilleure visualisation, nous montrons l'intensité moyenne associée à chaque région. On remarque sur la coupe centrale une homogénéisation trop importante de plusieurs HSB confluentes. La région ainsi obtenue mélange les HSB avec du cortex sain et sera éliminée par la procédure d'élimination des régions corticales.

- dans le cas de lésions temporales (particulièrement présentes chez les patients atteints de CADASIL), les cartes de probabilités de substance blanche ne sont pas suffisamment précises avec le recalage utilisé. Le critère de localisation empêche donc de sélectionner des régions comme HSB, alors qu'elles étaient correctement individualisées par la diffusion (figure 3.20).

Il faut ici remarquer que ces problèmes se posent dans le cas d'HSB importantes. En effet, les zones proches des circonvolutions et les zones temporales ne sont très souvent affectées que lorsque les HSB sont déjà très importantes et confluentes. La majeure partie de la charge lésionnelle est ainsi correctement détectée dans HSB_{init} et les HSB non détectées sont pour leur très grande majorité connexes à HSB_{init} . Il est alors possible de les intégrer à la segmentation en utilisant une dilatation géodésique.

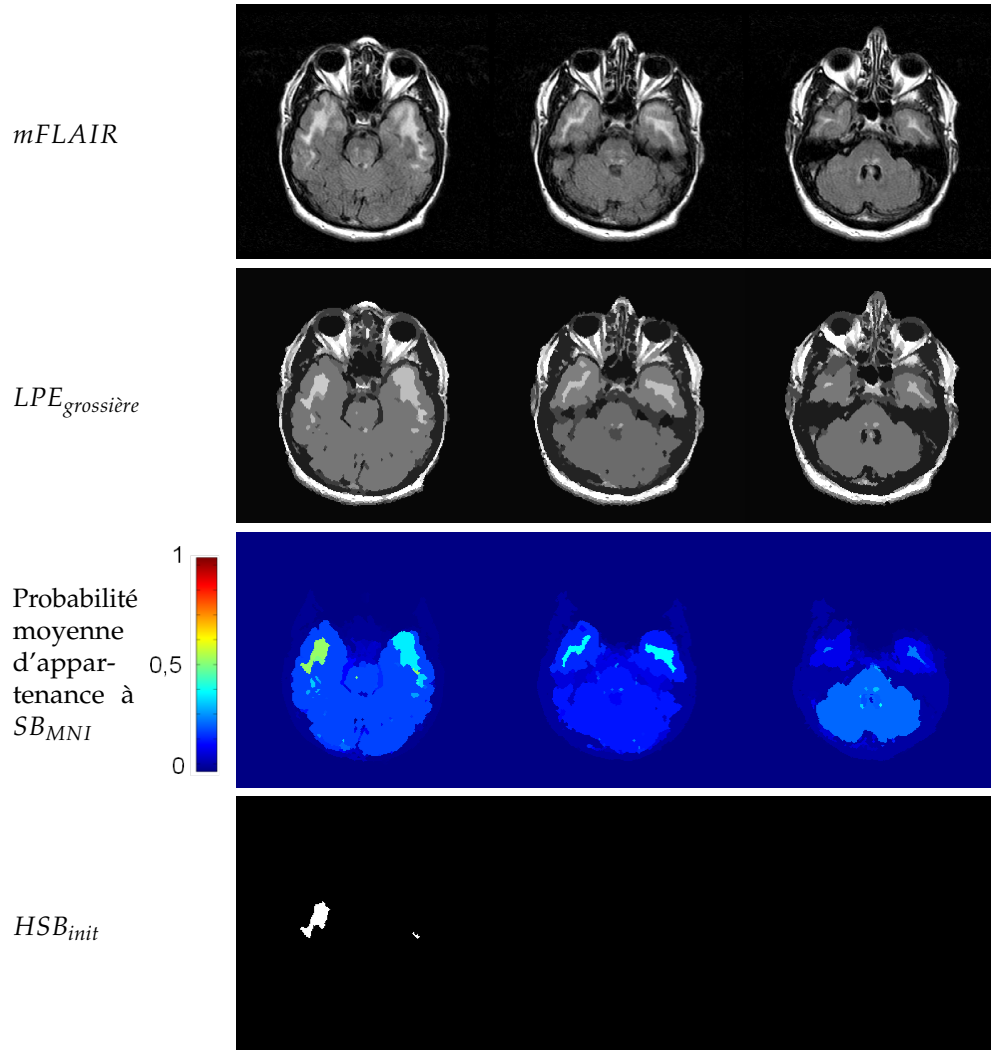


FIGURE 3.20 – Mauvaise détection des HSB temporales.

Ligne du haut : Trois coupes consécutives de l'image *mFLAIR* non diffusée d'un patient CADASIL. **2^{ème} ligne :** $LPE_{grossière}$ pour ces trois mêmes coupes. **3^{ème} ligne :** Probabilité moyenne de SB_{MNI} sur les régions de $LPE_{grossière}$. **Ligne du bas :** HSB_{init} pour ces trois mêmes coupes. La probabilité moyenne étant inférieure à 0,5 dans les extrémités des zones temporales, les HSB localisées à cet endroit ne sont pas retenues dans HSB_{init} .

3.5.2 Dilatation géodésique

La dilatation géodésique est une opération de morphologie mathématique, nécessitant un élément structurant et un ensemble X , dit *ensemble marqueur*. A partir d'un ensemble binaire initial, on effectue la dilatation par l'élément structurant puis on ne conserve que l'intersection du dilaté avec le marqueur X . Cette procédure est itérée jusqu'à ce que l'ensemble obtenu soit stable (figure 3.21).

Il nous faut donc définir l'ensemble X dans lequel effectuer la dilatation

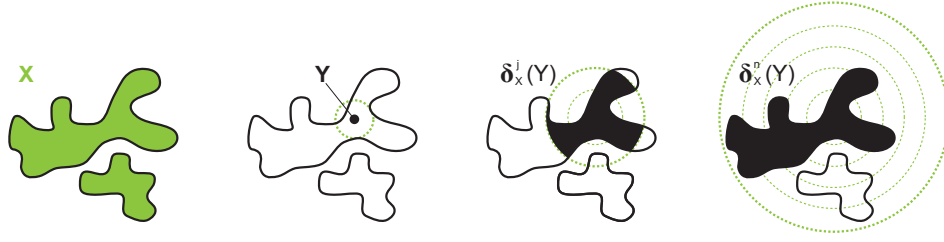


FIGURE 3.21 – Principe de la dilatation géodésique.

On considère un ensemble marqueur X , qui peut être constitué de plusieurs composantes connexes. On dispose d'un élément initial Y . Cet élément est dilaté de manière itérative (ici, la dilatation est effectuée radialement). Après chaque itération, on ne conserve que l'intersection du dilaté et de l'ensemble X . Cette procédure est répétée jusqu'à convergence. La procédure permet de récupérer la forme exacte des éléments de l'ensemble marqueur X , qui contenaient l'élément initial Y .

géodésique. Cet ensemble sera défini à partir de $LPE_{détailée}$ pour récupérer les coupes manquées et d'une couronne corticale pour récupérer les HSB temporales.

3.5.3 Ensemble marqueur pour la dilatation géodésique

Nous avons expliqué (section 3.3.6) comment obtenir une segmentation plus détaillée de l'image *mFLAIR* en utilisant comme paramètre de contraste $\lambda_{est}/2$. La figure 3.22 illustre le fait que $LPE_{détailée}$ permet de résoudre le premier type de problème mentionné. Par rapport à la figure 3.19, sur la même image, les HSB sont cette fois correctement individualisées et ne sont plus fusionnées avec les circonvolutions.

En ce qui concerne les HSB temporales, nous pouvons nous permettre d'avoir une contrainte spatiale moins forte sur $LPE_{détailée}$ que sur $LPE_{grossière}$ puisque la segmentation est désormais contrainte par HSB_{init} et que les régions de $LPE_{détailée}$ sont plus petites.

Nous allons donc considérer comme ensemble marqueur X les régions de $LPE_{détailée}$ qui vérifient les critères d'intensité et de localisation suivants :

- intensité moyenne de la région supérieure au seuil $s = \tilde{i} + 2\lambda_{est}$,
- probabilité moyenne SB_{MNI} de la région supérieure à 0,5 OU régions situées dans une couronne corticale de largeur 40 voxels (obtenue de manière similaire à la couronne de largeur 10 voxels présentée à la figure 3.18).

Les figures 3.23 et 3.24 illustrent l'ensemble marqueur X .

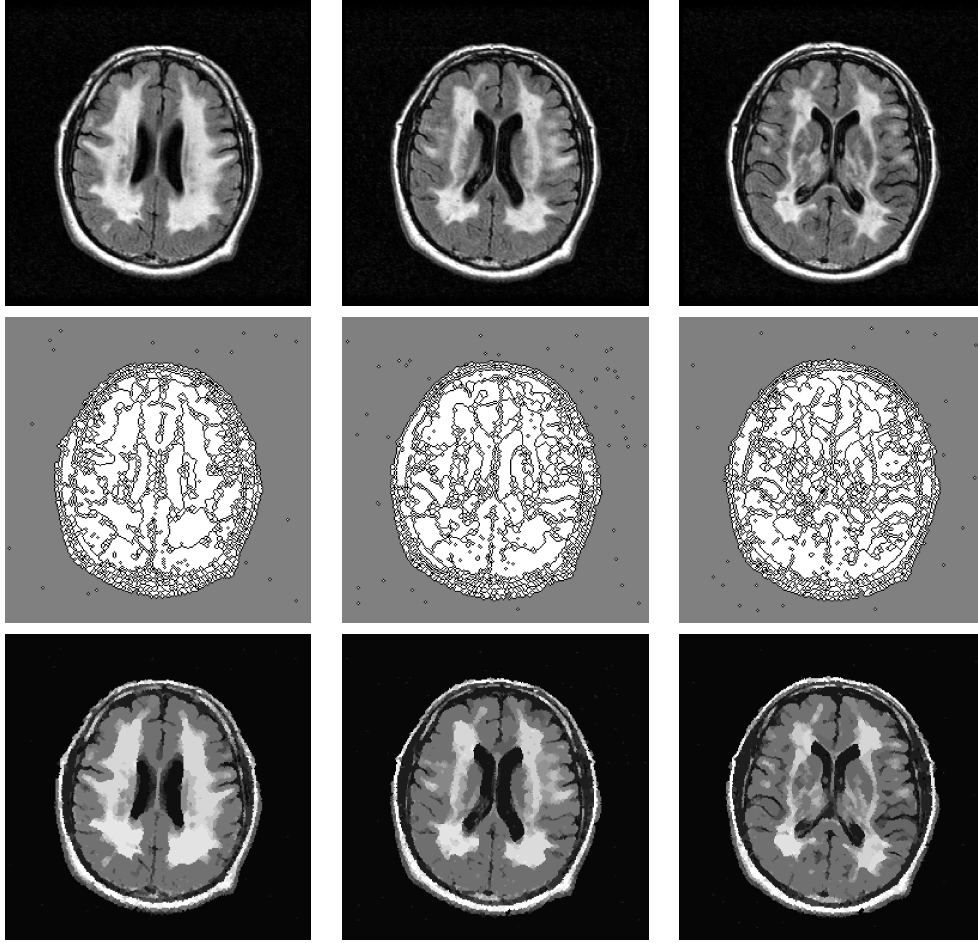


FIGURE 3.22 – Apport de $LPE_{détaillée}$.

Haut : Trois coupes consécutives de l'image $mFLAIR$ non diffusée d'un patient CADA-SIL. Ces coupes sont identiques à celles de la figure 3.19. **Milieu** : $LPE_{détaillée}$ pour ces trois mêmes coupes. **Bas** : Pour une meilleure visualisation, nous montrons l'intensité moyenne associée à chaque région de $LPE_{détaillée}$. Par rapport à la figure 3.19, des régions moins contrastées sont conservées. En particulier, la coupe médiane délimite désormais les HSB de manière correcte.

3.5.4 Obtention de HSB_{final}

Nous partons de la segmentation binaire HSB_{init} obtenue précédemment. Comme précisé précédemment, les anomalies observées se produisent uniquement pour des HSB importantes. Nous ne retenons donc que les composantes connexes dont le volume est supérieur à 50 voxels. La dilatation géodésique de ces régions conditionnellement à l'ensemble X des voxels candidats est ensuite effectuée avec un élément structurant 3D (6-connectivité) afin d'obtenir la segmentation finale HSB_{final} . La dilatation est ici effectuée en 3D de manière à pouvoir réintégrer dans la segmentation des coupes où la diffusion était trop importante.

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

Les figures 3.23 et 3.24 montrent l'effet de cette étape de croissance sur les mêmes images que celles utilisées pour illustrer les insuffisances de HSB_{init} . Grâce à cette étape de croissance, les HSB initialement non détectées sont désormais correctement segmentées.

3.6 Conclusion

Ce chapitre a présenté la chaîne de traitements WHASA, dédiée à la segmentation des hyperintensités de la substance blanche. Les étapes de prétraitement usuelles ont d'abord été réalisées à l'aide du logiciel SPM. Cela nous permet par la suite d'estimer, pour chaque sujet, la valeur moyenne du contraste entre la substance blanche et la substance grise sur l'image FLAIR. Un processus de diffusion non linéaire, dépendant de ce paramètre, est ensuite alterné avec un algorithme de ligne de partage des eaux afin d'obtenir deux segmentations de l'image FLAIR en régions :

- $LPE_{grossière}$, peu détaillée, permet néanmoins le renforcement du contraste de la quasi totalité des HSB, facilitant ainsi leur détection ;
- $LPE_{détaillée}$, plus précise, permet une meilleure délimitation des HSB.

Trois critères (intensité, localisation dans la substance blanche et non connexité à l'interface $SG|LCS$) sont alors utilisés pour obtenir à partir de $LPE_{grossière}$ une première segmentation des HSB (HSB_{init}). Les plus grosses composantes connexes sont finalement affinées par l'intermédiaire d'une dilatation géodésique s'appuyant sur les régions plus détaillées de $LPE_{détaillée}$.

L'avantage principal de cette méthode est d'être basée sur le contraste propre à chaque sujet. Elle peut donc être appliquée directement sur des données multicentriques, y compris pour des IRM ayant différentes intensités de champ magnétique (1T, 1.5T ou 3T). Elle repose également sur peu de paramètres et est de plus complètement automatique. Le chapitre suivant sera consacré à l'évaluation de cette chaîne de traitements sur différents jeux de données afin d'analyser les performances et la robustesse de la méthode.

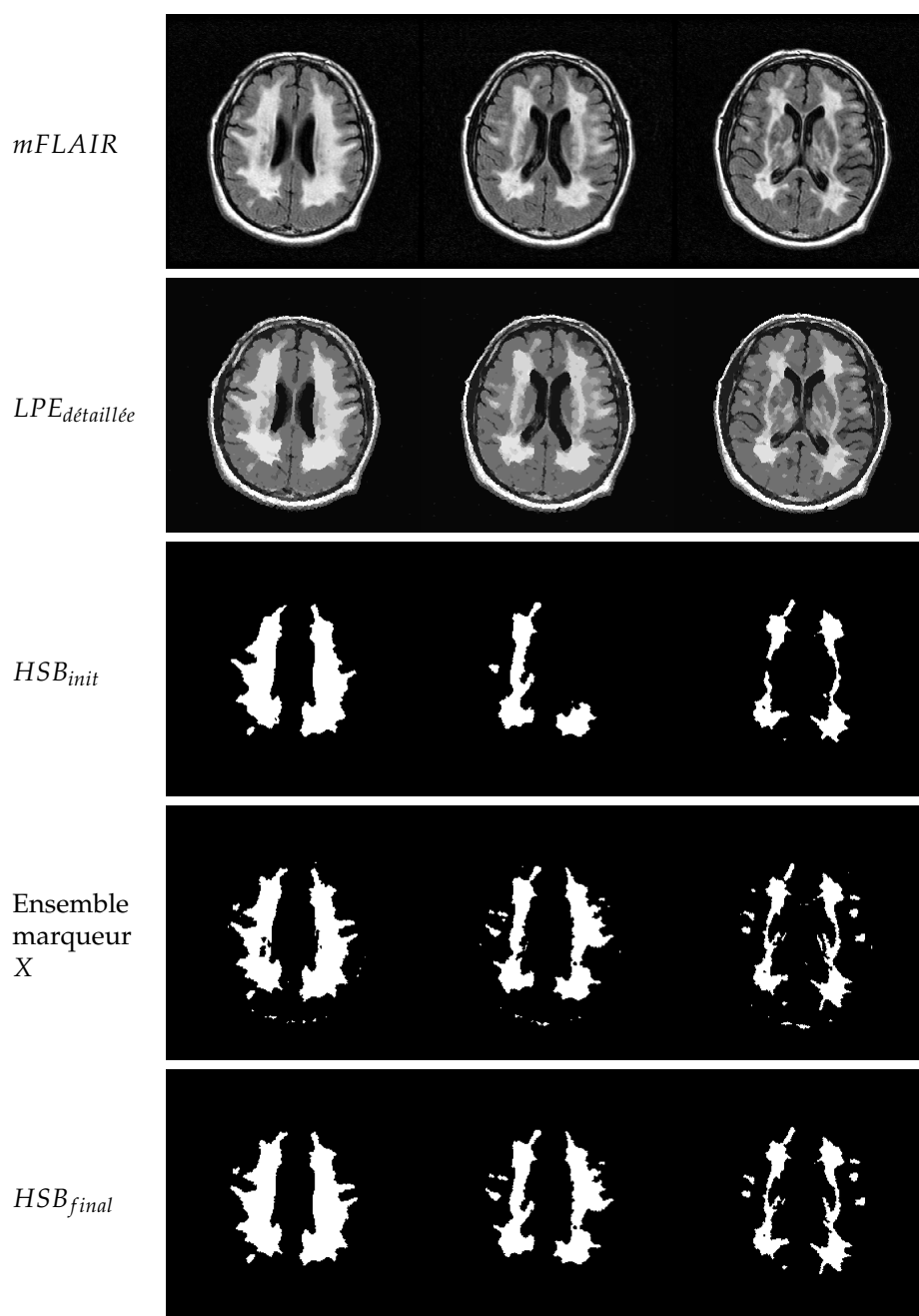


FIGURE 3.23 – Illustration de l'étape de croissance dans le cas de l'échec du processus de diffusion sur certaines coupes.

Ligne du haut : trois coupes consécutives d'une image $mFLAIR$ d'un patient CADASIL. **2^{ème} ligne :** $LPE_{détailée}$ pour ces trois mêmes coupes. **3^{ème} ligne :** HSB_{init} pour ces trois mêmes coupes. On remarque qu'une HSB importante est manquante sur la coupe médiane (voir figure 3.19). **4^{ème} ligne :** Ensemble marqueur pour la dilatation géodésique défini à partir de $LPE_{détailée}$. La contrainte spatiale étant moins forte, on remarque un certain nombre de petites régions au niveau cortical. **Ligne du bas :** HSB_{final} obtenu par dilatation géodésique de HSB_{init} dans l'ensemble marqueur X. La zone manquante est bien récupérée.

3. MÉTHODE DE SEGMENTATION DES HYPERINTENSITÉS DE LA SUBSTANCE BLANCHE (WHASA)

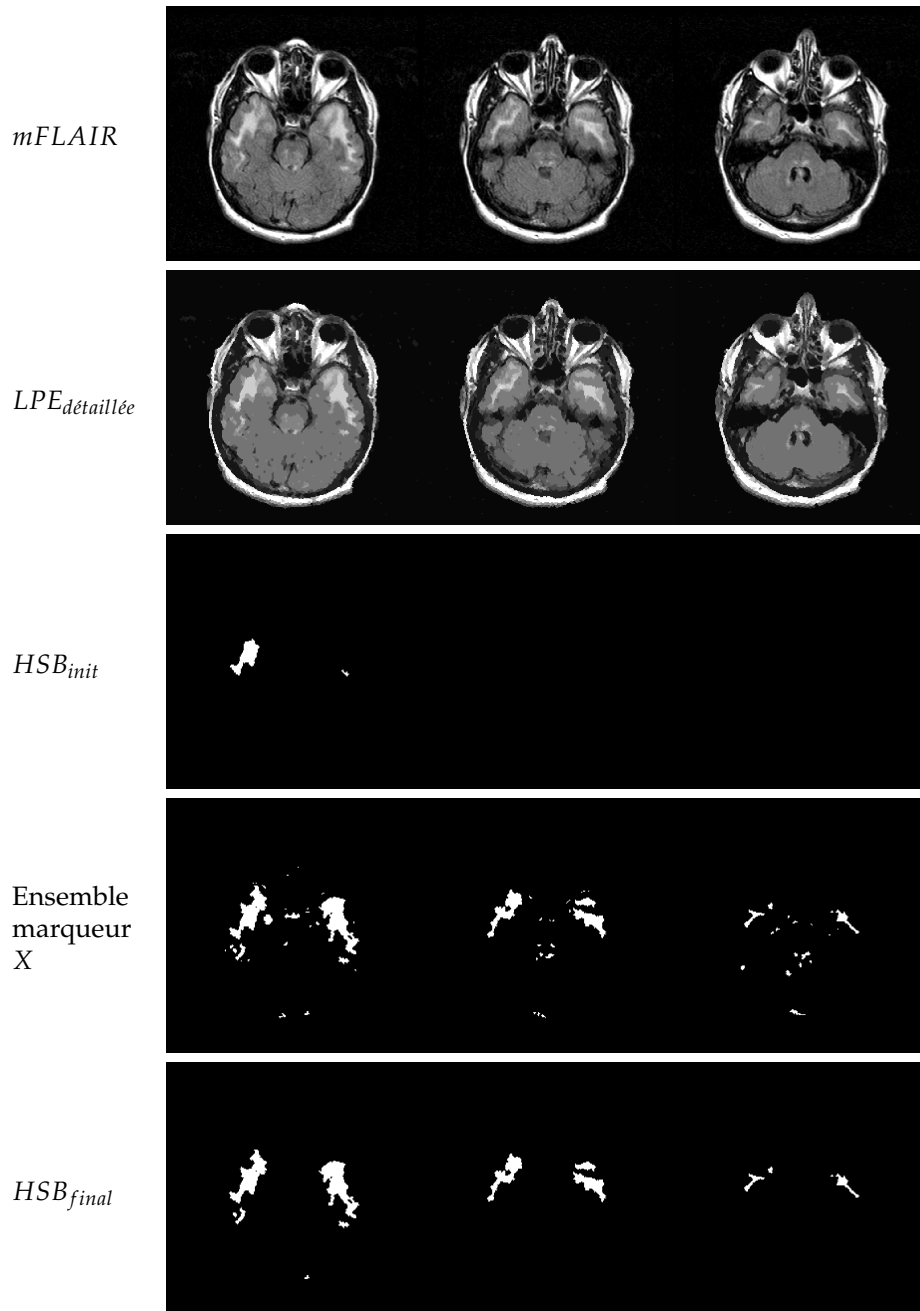


FIGURE 3.24 – Illustration de l'étape de croissance dans le cas d'une mauvaise détection des HSB temporales.

Ligne du haut : trois coupes consécutives d'une image *mFLAIR* d'un patient CADASIL. **2^{ème} ligne :** *LPE_{détaillée}* pour ces trois mêmes coupes. **3^{ème} ligne :** *HSB_{init}* pour ces trois mêmes coupes. On remarque que les HSB temporales ne sont pas détectées (voir figure 3.20). **4^{ème} ligne :** Ensemble marqueur pour la dilatation géodésique défini à partir de *LPE_{détaillée}*. **Ligne du bas :** *HSB_{final}* obtenu par dilatation géodésique de *HSB_{init}* dans l'ensemble marqueur X. La zone temporale est bien récupérée.

Chapitre 4

Evaluation de la méthode de segmentation

Le chapitre précédent a présenté la chaîne de traitements WHASA, permettant d'obtenir une segmentation des hyperintensités de la substance blanche (HSB) à partir d'une image T1 et d'une image FLAIR. Dans ce chapitre, nous évaluons la qualité de la segmentation obtenue par WHASA. Le problème de l'évaluation d'une méthode de segmentation des HSB est difficile car celles-ci ne sont pas toujours bien définies, particulièrement dans le cas des HSB confluentes. De plus, la grande hétérogénéité des lésions sous-jacentes rend difficile l'établissement de critères clairs pour la prise en compte des HSB. La plupart du temps, l'évaluation d'une méthode de segmentation automatique se fait donc par comparaison avec une segmentation manuelle servant de référence, approche qui a été utilisée dans cette thèse. De plus, il est nécessaire de disposer d'images provenant de différents centres d'acquisition et présentant une large variabilité de charge lésionnelle pour s'assurer que la méthode de segmentation est suffisamment robuste.

Ce chapitre est organisé de la manière suivante. Nous présenterons dans un premier temps deux jeux de données qui seront utilisés pour l'évaluation et les segmentations manuelles utilisées comme référence. Nous introduirons ensuite les outils permettant la comparaison entre les résultats de WHASA et les références manuelles. Cette comparaison sera effectuée au niveau volumique et topographique. Enfin, nous évaluerons la robustesse des résultats obtenus par WHASA lorsque certains paramètres importants de la chaîne de traitements varient. Ce chapitre se

concentre sur l'évaluation des performances de WHASA, la comparaison avec d'autres méthodes de segmentation des HSB issues de la littérature étant présentée dans le chapitre 6.

4.1 Données

Deux jeux de données différents, pour lesquels les paramètres d'acquisition IRM sont résumés dans le tableau 4.1, ont été utilisés pour l'évaluation.

	Machine	Sequence	TR (ms)	TE (ms)	TI (ms)	FA (°)	Taille voxel (mm)
Centre 1 (N=6)	Philips	3DT1	7,9	3,7	672	8	$1 \times 1 \times 1,3$
	1,5T	FLAIR	10000	140	2200	90	$0,94 \times 0,94 \times 5,5$
Centre 2 (N=1)	Philips	3DT1	7,9	3,7	674	8	$1 \times 1 \times 1,3$
	1,5T	FLAIR	10000	140	2200	90	$0,94 \times 0,94 \times 5,5$
Centre 3 (N=7)	Philips	3DT1	7,9	3,7	665	8	$1 \times 1 \times 1,3$
	1,5T	FLAIR	11000	140	2200	90	$0,9 \times 0,9 \times 6,2$
Centre 4 (N=3)	Siemens	3DT1	11,2	3,5	1100	10	$0,94 \times 0,94 \times 1,3$
	1,5T	FLAIR	8280	111	2500	90	$0,9 \times 0,9 \times 4$
Centre 5 (N=7)	GE	3DT1	11,2	3,5	600	10	$0,94 \times 0,94 \times 1,3$
	1,5T	FLAIR	10002	145	2200	90	$0,94 \times 0,94 \times 5,5$
Centre 6 (N=43)	GE	3DT1	8,6	1,9	-	20	$1,02 \times 1,02 \times 0,8$
	1,5T	FLAIR	8402	161	2002	90	$0,94 \times 0,94 \times 5,5$

TABLEAU 4.1 – Principaux paramètres d'acquisition pour les différents centres des jeux de données *Hippocampe* et *Cadasil*.

Les données *Hippocampe* ont été acquises dans les centres 1 à 5, les données *Cadasil* dans le centre 6.

4.1.1 Jeu de données *Hippocampe*

Ce jeu de données correspond aux centres 1 à 5 dans le tableau 4.1. Il est composé de 24 patients ayant passé un examen IRM dans le cadre de l'étude *Hippocampe*. Cette étude multicentrique visait à évaluer l'effet du Donépézil sur des patients MCI amnésiques [Dubois et al., 2009]. Les 24 patients retenus ont été sélectionnés de manière à être représentatifs des conditions d'acquisition (cinq centres d'acquisition différents) ainsi que de la variabilité de la charge lésionnelle tout en répondant à certains critères de qualité sur les images FLAIR (absence de gros artefacts de mouvement notamment). Les patients ont été choisis parmi l'ensemble

des participants, ils pouvaient donc présenter des atteintes vasculaires supérieures aux critères d'exclusion de l'étude.

4.1.2 Jeu de données *Cadasil*

Ce jeu de données correspond au centre 6 dans le tableau 4.1. Il comporte 43 patients atteints de CADASIL (*Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarct and Leukoaraiosis*), qui est un type de démence vasculaire héréditaire (voir sec 1.2.3). Ces données ont été analysées dans le cadre d'une collaboration avec le Professeur Hugues Chabriat et le Docteur Eric Jouvent, du service de Neurologie de l'hôpital Lariboisière à Paris. Ces 43 patients sont un sous-échantillon, sélectionné aléatoirement, d'une étude prospective plus large [Viswanathan et al., 2006]. Tous les patients ont eu leur diagnostic confirmé par l'identification d'une mutation sur le gène *Notch3*. Les patients CADASIL présentent pour une grande partie d'entre eux des HSB importantes.

4.2 Référence manuelle

Pour les données *Hippocampe*, j'ai effectué une segmentation manuelle sous la supervision du Dr Claire Boutet, neuroradiologue à l'hôpital de la Pitié-Salpêtrière. Cette segmentation a été réalisée à l'aide du logiciel *Anatomist*¹ en utilisant l'image FLAIR. Un seuil sur l'intensité a permis de guider la segmentation mais chaque composante connexe pouvait être modifiée manuellement afin d'ajuster les bords. De plus, ce seuil sur l'intensité pouvait être modifié entre les coupes d'un même sujet, ainsi que sur une même coupe, si des HSB présentaient des valeurs d'intensités différentes.

En ce qui concerne les données *Cadasil*, les HSB ont été segmentées en analysant les coupes axiales de l'image FLAIR de la base du cervelet jusqu'au sommet de la tête. Après application d'un masque de la cavité intra-crânienne, les HSB étaient détectées en appliquant sur l'intensité de l'image FLAIR un seuil déduit de l'histogramme. Deux neurologues experts ont ensuite validé et, si nécessaire, corrigé les contours des HSB pour chaque lésion.

1. <http://brainvisa.info/doc/axon/en/help/aboutAnatomist.html>

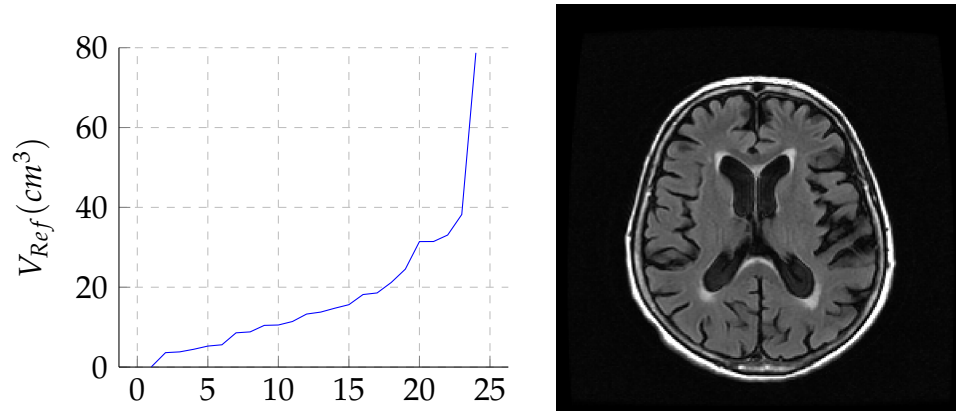


FIGURE 4.1 – Répartition des volumes manuels pour les 24 sujets des données *Hippocampe*.

La coupe présentée est issue d'un sujet dont le volume des HSB est proche du volume médian des patients *Hippocampe* (volume du sujet : $13,3 \text{ cm}^3$, volume médian *Hippocampe* : $13,5 \text{ cm}^3$).

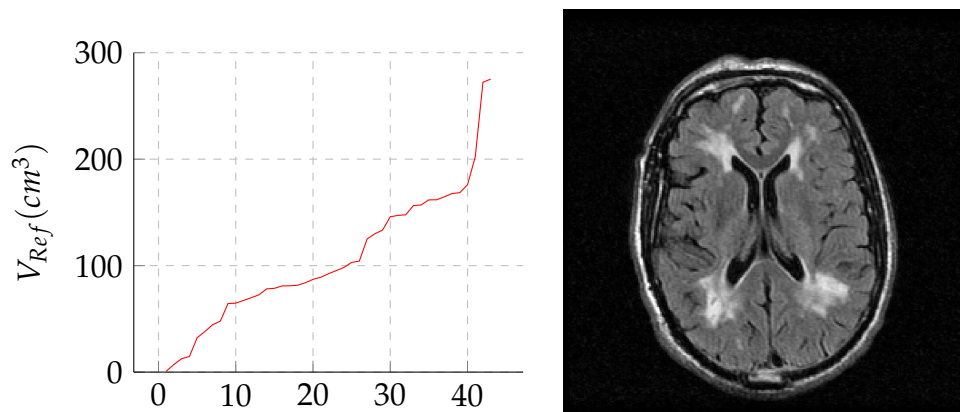


FIGURE 4.2 – Répartition des volumes manuels pour les 43 sujets des données *Cadasil*.

La coupe présentée est issue d'un sujet dont le volume des HSB est proche du volume médian des patients *Cadasil* (volume du sujet : $89,1 \text{ cm}^3$, volume médian *Cadasil* : $92,4 \text{ cm}^3$).

Une segmentation binaire de référence est ainsi obtenue pour les 67 images IRM des deux jeux de données. Les volumes médians, quartiles et volumes minima et maxima pour chacun des jeux de données sont résumés dans le tableau 4.2 et des exemples sont donnés dans les figures 4.1 et 4.2.

Jeu de données	Volume médian (cm^3)	[Q1 Q3] (cm^3)	[Min ; Max] (cm^3)
<i>Hippocampe</i>	13,5	[7,1 22,9]	[0,1 ; 78,7]
<i>Cadasil</i>	92,4	[67,9 154,2]	[0,8 ; 275,2]

TABLEAU 4.2 – Répartition des volumes manuels de référence dans les deux jeux de données.

Q1 : 1^{er} quartile, Q3 : 3^{ème} quartile.

4.3 Outils d'évaluation

Nous disposons désormais d'une segmentation de référence, notée *Ref*, à laquelle nous pouvons comparer la segmentation obtenue par la chaîne de traitements WHASA, notée *Seg*. Cette section présente les méthodes statistiques utilisées pour comparer *Seg* et *Ref*, aussi bien au niveau de leur volume que de leur topographie. Le volume des HSB est calculé en multipliant le nombre total des voxels appartenant à *Ref* (respectivement *Seg*) par le volume d'un voxel et est noté V_{Ref} (respectivement V_{Seg}). Toutes les analyses statistiques présentées ici ont été réalisées avec Matlab.

4.3.1 Coefficient de corrélation intraclasse (ICC)

Le coefficient de corrélation intraclasse (ICC) est une statistique descriptive permettant de quantifier la concordance entre plusieurs observateurs. Plusieurs variantes existent pour les ICC, selon le nombre d'observateurs et le nombre de mesures effectuées par chaque observateur. Nous sommes ici dans le cas où nous disposons d'une seule valeur par deux observateurs différents. Nous utilisons donc l'ICC(3,1) selon la terminologie employée par Shrout & Fleiss [1979] (analyse de variance à deux facteurs, modèle mixte, mesure unique).

4.3.2 Régression lineaire

L'ICC nous fournit une quantification de l'accord entre V_{Ref} et V_{Seg} , il ne permet néanmoins pas de savoir, par exemple, si l'une des observations est systématiquement plus faible que l'autre.

Pour cela, nous effectuons une régression linéaire de V_{seg} par rapport à V_{Ref} . Nous reportons l'équation de régression obtenue, le coefficient de corrélation élevé au carré R^2 et la p-value associée au test de Fischer (H_0 :

nullité du coefficient de la droite de régression). La comparaison de la droite de régression avec la droite identité permet de détecter les valeurs pour lesquels V_{Seg} sous-estime ou surestime V_{Ref} .

4.3.3 Graphe de Bland-Altman

La méthode de régression linéaire ne permet toutefois pas de quantifier la différence entre V_{Ref} et V_{Seg} . Pour cela, nous utilisons la méthode d'évaluation proposée par Bland & Altman [1986] permettant de comparer deux méthodes de mesure d'une même grandeur physique.

Dans un graphe de Bland-Altman, les différences entre les valeurs des deux méthodes de mesure ($D = V_{Seg} - V_{Ref}$) sont représentées en fonction de leur moyenne ($A = \frac{V_{Seg} + V_{Ref}}{2}$). L'analyse de ces différences permet d'estimer la cohérence entre les deux méthodes de mesure.

4.3.3.1 Cas où les différences sont normalement distribuées

Dans le cas où les différences sont distribuées selon une loi normale, la moyenne de D , notée $\bar{D}$, fournit alors une estimation du biais entre les deux méthodes. Cependant, cette valeur du biais sera dépendante de l'échantillon considéré. Un intervalle de confiance à 95% pour l'estimation du biais peut être obtenu comme $\bar{D} \pm t \frac{sd(D)}{\sqrt{n}}$, où $sd(D)$ est l'écart-type des différences, t est le quantile correspondant à une valeur de risque bilatérale de 5% pour une loi de Student à $n - 1$ degrés de liberté et n le nombre d'observations.

De plus, toujours sous l'hypothèse que les différences sont distribuées normalement, il est également possible de calculer l'intervalle contenant 95% des valeurs des différences. Cet intervalle s'écrit $[\bar{D} - 1,96sd(D) ; \bar{D} + 1,96sd(D)]$ et ses bornes sont appelées *limits of agreement*.

4.3.3.2 Cas où les différences ne sont pas normalement distribuées

Dans le cas où les différences ne sont plus normalement distribuées, les estimations du biais et des *limits of agreement* ne peuvent plus être obtenues de la même manière. C'est en particulier le cas si la variance des différences augmente avec la valeur mesurée. Dans ce cas, les *limits of agreement* calculées selon la formule présentée au paragraphe précédent

seront trop larges pour les faibles valeurs mesurées et trop étroites pour les valeurs importantes.

Deux variantes ont donc été introduites pour traiter de tels cas [Bland & Altman, 1999]. La première modélise les variations en fonction de la valeur mesurée par des régressions linéaires et permet d'interpréter les différences en conservant l'échelle initiale de mesure (dans notre cas, en volume). La seconde s'appuie sur une transformation logarithmique, exprimant ainsi l'accord entre les méthodes en termes de ratio.

Modélisation par régression La première méthode proposée effectue la régression linéaire des différences D par rapport à la moyenne A . Cela aboutit à l'équation de régression

$$\hat{D} = b_0 + b_1 A \quad (4.1)$$

Si le coefficient b_1 est significativement différent de 0 (nous considérons par la suite que ce sera le cas si la p – *value* associée au test de Fischer est inférieure à 0,05), l'équation 4.1 fournit une estimation du biais pour les différentes valeurs de l'intervalle de mesure. Si ce n'est pas le cas, l'estimateur du biais demeure $\bar{D}$.

En ce qui concerne la variabilité des différences, il est nécessaire d'étudier les résidus R du modèle de régression 4.1. En effectuant la régression de la valeur absolue de ces résidus, noté $|R|$, en fonction de la valeur mesurée (estimée par A), il est possible d'analyser la relation entre la variabilité des mesures et la valeur mesurée. Cette relation est alors décrite par l'équation de régression :

$$|\widehat{R}| = c_0 + c_1 A \quad (4.2)$$

Si le coefficient c_1 n'est pas significativement différent de 0 (p – *value* associée au test de Fischer supérieure à 0,05), la variabilité des différences ne dépend pas de la valeur mesurée, les *limits of agreement* sont calculées comme $\hat{D} \pm 1.96sd(R)$.

Dans le cas où c_1 est significativement différent de 0, il est possible d'exprimer $sd(R)$ en fonction de $|\widehat{R}|$. En effet, en supposant que les résidus R soient normalement distribués, la valeur absolue des résidus suit alors une loi demi-normale dont la moyenne vaut $\sqrt{\frac{2}{\pi}}sd(R)$. En combinant cette relation et l'équation 4.2, les *limits of agreement* peuvent ainsi être

obtenues par :

$$\widehat{D} \pm 1,96\sqrt{\frac{\pi}{2}}|\widehat{R}| = (b_0 + b_1A) \pm 1,96\sqrt{\frac{\pi}{2}}(c_0 + c_1A)$$

Transformation logarithmique Dans le cas où la variabilité des différences augmente avec la valeur mesurée, Bland & Altman [1999] ont également proposé d'analyser les logarithmes des valeurs afin de se ramener au cas de différences normalement distribuées.

Dans ce cas, on analyse la différence $D = \log V_{Seg} - \log V_{Ref}$ en fonction de la moyenne $A = \frac{1}{2}(\log V_{Ref} + \log V_{Seg})$. Le biais est alors à nouveau estimé comme $\overline{D}$ et les *limits of agreement* par $[\overline{D} - 1,96sd(D) ; \overline{D} + 1,96sd(D)]$. Néanmoins, ces valeurs sont obtenues pour les logarithmes des quantités initiales. L'interprétation en fonction des logarithmes étant moins aisée, il faut utiliser la transformation réciproque (en l'occurrence, l'exponentielle) qui permet de se ramener à l'étude du ratio des quantités initiales. Le graphe final représentera donc $\exp(\log V_{Seg} - \log V_{Ref}) = \frac{V_{Seg}}{V_{Ref}}$ en fonction, non plus de la moyenne arithmétique, mais de la moyenne géométrique $\sqrt{V_{Seg}V_{Ref}}$.

4.3.4 Validation spatiale

La comparaison volumique n'apporte cependant aucune information sur la précision spatiale. Il est ainsi possible que deux segmentations aient le même volume mais soient complètement disjointes. En plus de l'évaluation volumique présentée ci-dessus, il est donc nécessaire d'évaluer la précision spatiale de la segmentation automatique, notamment dans le cadre de caractérisation de topographie des lésions.

Lors de la comparaison entre *Ref* et *Seg*, quatre types de voxels peuvent être distingués (figure 4.3) :

- les *Vrais Positifs* (VP) : les voxels appartenant à la fois à *Seg* et à *Ref* ;
- les *Faux Négatifs* (FN) : les voxels appartenant à *Ref* mais n'étant pas dans *Seg* ;
- les *Faux Positifs* (FP) : les voxels appartenant à *Seg* mais n'étant pas dans *Ref* ;
- les *Vrais Négatifs* (VN) : les voxels n'appartenant ni à *Ref* ni à *Seg*.

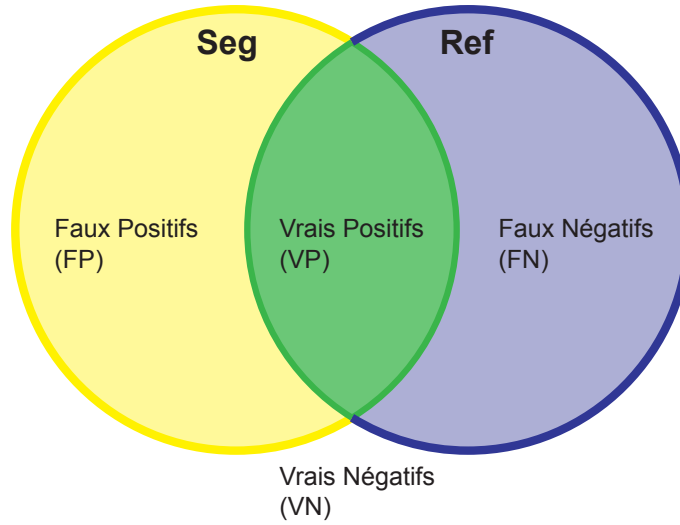


FIGURE 4.3 – Schéma présentant les différentes catégories de voxels lors d'une comparaison entre *Ref* et *Seg*.

En ce qui concerne les HSB, les VN représentent la quasi totalité du volume de l'image initiale et ne sont donc pas une grandeur informative. Pour les autres catégories, les VP nous indiquent les voxels qui sont corrects, les FN ceux qui manquent (cas de sous-segmentation) et les FP ceux qui sont en trop (cas de sursegmentation). Un indice composite permet de résumer ces informations en une seule grandeur indiquant la superposition spatiale des deux segmentations, il s'agit de l'indice de similarité (SI), aussi appelé coefficient de Dice [Dice, 1945]. Le SI est défini par :

$$SI = 2 \frac{Card(Ref \cap Seg)}{Card(Ref) + Card(Seg)} = 2 \frac{VP}{2VP + FP + FN}$$

Ainsi, une valeur de 0 indique que *Ref* et *Seg* n'ont aucune partie commune alors qu'une valeur de 1 indique une superposition parfaite des deux segmentations.

4.4 Résultats

Nous présentons ci-dessous les résultats de l'évaluation quantitative entre Ref et Seg selon les indicateurs exposés précédemment et les illustrerons par quelques exemples qualitatifs.

4.4.1 ICC et régression linéaire

La valeur de l'ICC entre V_{Ref} et V_{Seg} est de 0,97 indiquant une très bonne cohérence entre les volumes de référence et les volumes obtenus par WHASA.

La régression linéaire de V_{Seg} par rapport à V_{Ref} est présentée dans la figure 4.4. La droite de régression d'équation $y = 0,95x + 5,3$ est très proche de l'identité, le coefficient de corrélation élevé ($R^2 = 0,93$) et la p-value inférieure à 0,01.

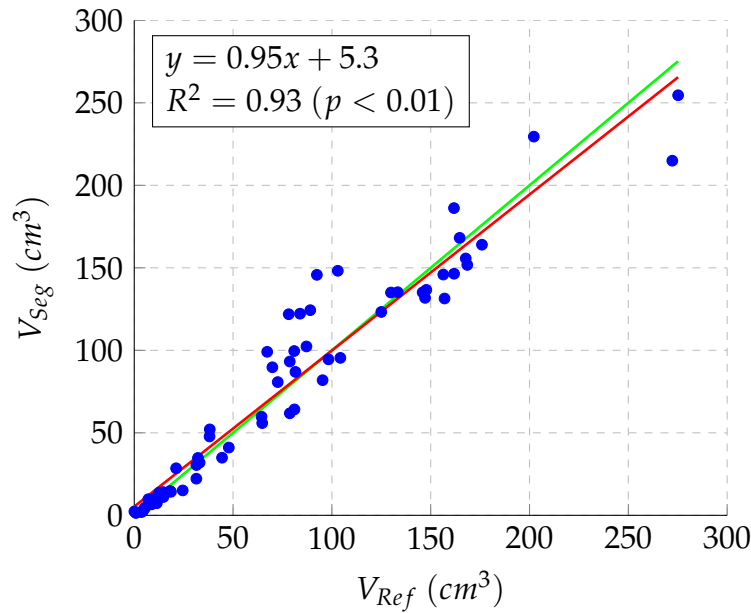


FIGURE 4.4 – Régression linéaire de V_{Seg} par rapport à V_{Ref} .
En rouge : Droite de régression. En vert : Droite identité.

On remarque toutefois que la variance semble augmenter avec la valeur de V_{Ref} et que la droite est principalement influencée par les volumes les plus importants, tous les points ayant un faible volume ($V_{Ref} < 50cm^3$) étant très proches. On effectue donc également la régression de $\log V_{Seg}$ par rapport à $\log V_{Ref}$ (figure 4.5).

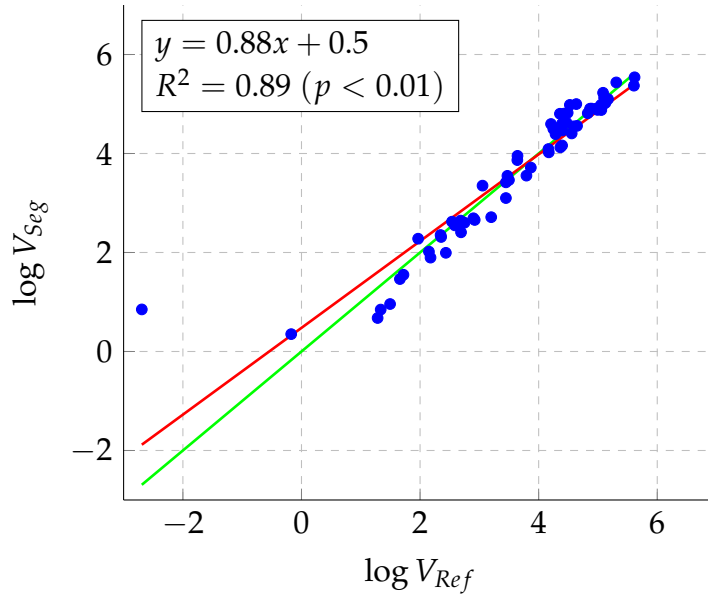


FIGURE 4.5 – Régression linéaire de $\log V_{Seg}$ par rapport à $\log V_{Ref}$.
En rouge : Droite de régression. En vert : Droite identité.

Après transformation logarithmique, la corrélation est toujours très bonne mais deux points sont des *outliers* du modèle. Il s'agit des deux sujets ayant les plus faibles volumes V_{Ref} . En effet, plus la valeur de V_{Ref} est faible, plus une faible erreur sur V_{Seg} sera proportionnellement importante. Précisons cependant que, pour ces deux sujets, le volume V_{Seg} déterminé automatiquement figure également parmi les plus faibles. Ainsi, les volumes de références pour ces deux sujets sont de $0,07 \text{ cm}^3$ et de $0,84 \text{ cm}^3$. Les volumes obtenus par WHASA sont respectivement de $2,3 \text{ cm}^3$ et de $1,4 \text{ cm}^3$. Les figures 4.6 et 4.7 montrent les images correspondantes à ces deux sujets.

Les erreurs de segmentation sont relativement faibles. Dans les deux cas, les problèmes de sursegmentation viennent en particulier d'effets de volume partiel ou sont des points isolés aux bords des ventricules. La segmentation n'a en tout cas pas complètement échoué. Il apparaît donc justifié de considérer ces deux sujets comme non représentatifs et de les exclure de l'analyse lorsque la transformation logarithmique est utilisée. La figure 4.8 montre les résultats de la régression après transformation logarithmique en excluant les deux *outliers*. La corrélation demeure excellente et la droite de régression est quasiment confondue avec la droite identité.

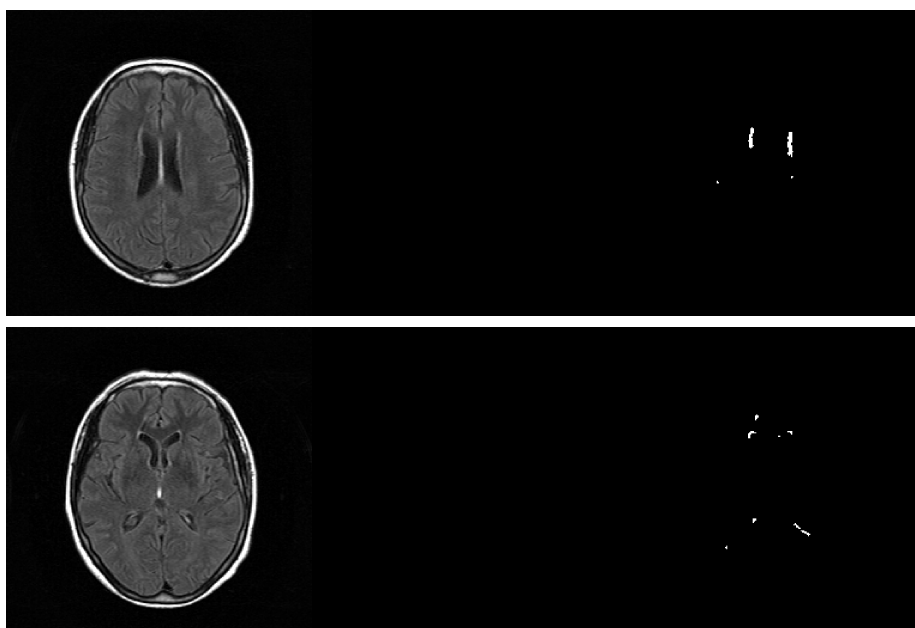


FIGURE 4.6 – Coupes illustratives du premier sujet *outlier*.
 Pour ce sujet, $V_{Ref} = 0,07 \text{ cm}^3$ et $V_{Seg} = 2,3 \text{ cm}^3$. Les coupes choisies sont celles présentant le plus d'hypersignaux. **1^{ère} colonne** : coupes de l'images *mFLAIR*. **2^{ème} colonne** : *Ref* pour les mêmes coupes. **3^{ème} colonne** : *Seg* pour les mêmes coupes.

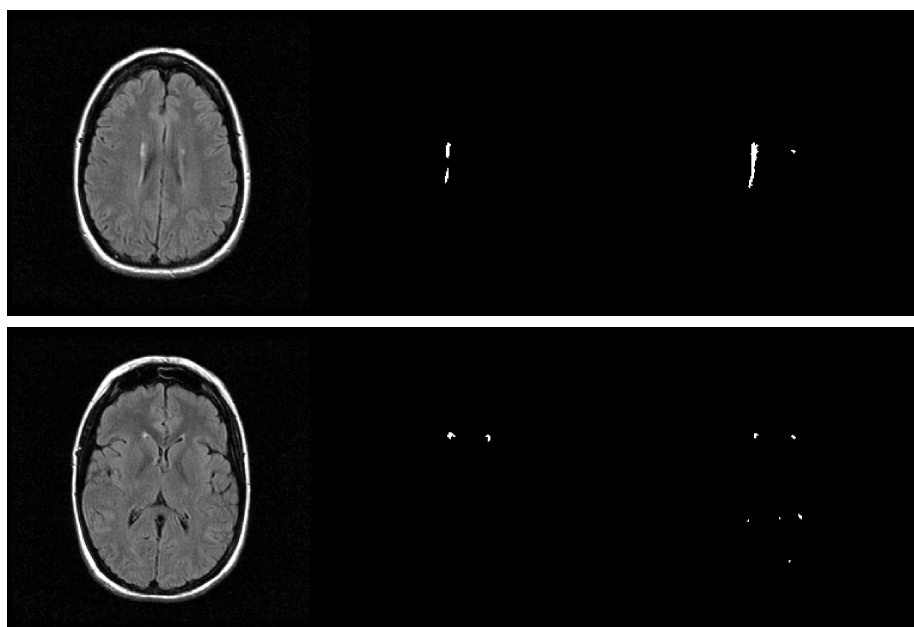


FIGURE 4.7 – Coupes illustratives du second sujet *outlier*.
 Pour ce sujet, $V_{Ref} = 0,84 \text{ cm}^3$ et $V_{Seg} = 1,4 \text{ cm}^3$. Les coupes choisies sont celles présentant le plus d'hypersignaux. **1^{ère} colonne** : coupes de l'images *mFLAIR*. **2^{ème} colonne** : *Ref* pour les mêmes coupes. **3^{ème} colonne** : *Seg* pour les mêmes coupes.

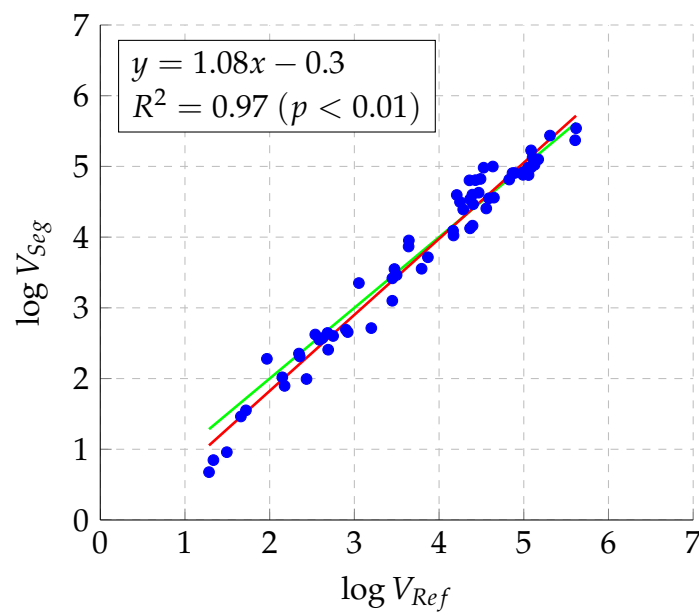


FIGURE 4.8 – Régression linéaire de $\log V_{Seg}$ par rapport à $\log V_{Ref}$ en excluant les sujets *outliers*.

En rouge : Droite de régression. En vert : Droite identité.

4.4.2 Graphes de Bland-Altman

Le graphe de Bland-Altman est présenté dans la figure 4.9. Le biais ne varie pas avec l'amplitude du volume et est estimé à $1,2 \text{ cm}^3$. Le calcul de l'intervalle de confiance à 95% indique que cette estimation du biais est comprise entre $-3,1$ et $5,5 \text{ cm}^3$. Il n'y a donc, en moyenne, ni surestimation ni sous-estimation du volume. Les *limits of agreement*, obtenues par la modélisation par régression, sont indiquées sur la figure 4.9.

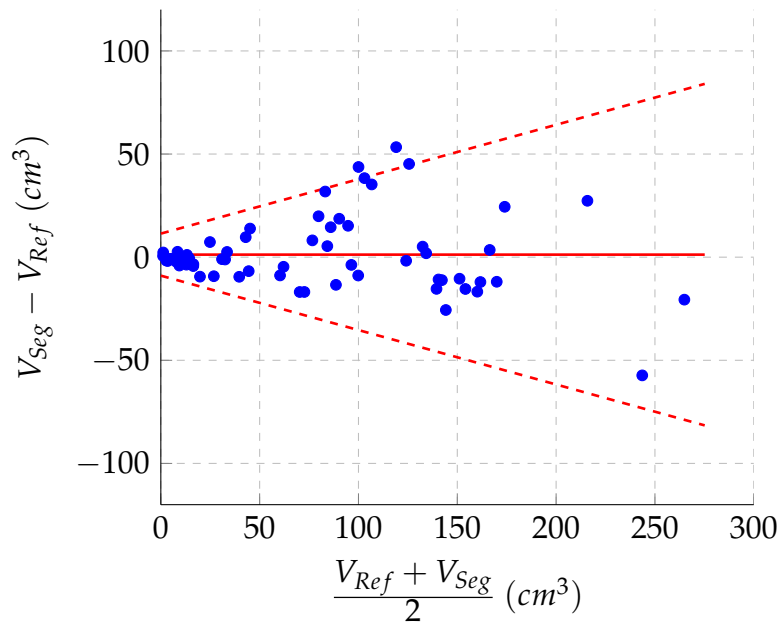


FIGURE 4.9 – Graphe de Bland-Altman entre V_{Seg} et V_{Ref} . La ligne rouge continue indique l'estimation du biais, les lignes pointillées les *limits of agreement* comprenant 95% des points.

La graphe de Bland-Altman après transformation logarithmique est montré sur la figure 4.10. Les deux outliers présentés précédemment sont exclus des analyses. Après transformation logarithmique, les différences sont distribuées normalement (test de Jacques-Berra, $p\text{-value} > 0,5$). Le biais est croissant avec la valeur moyenne du volume considéré, la variance des différences est par contre constante.

L'interprétation des logarithmes étant peu aisée, les équations du biais et des *limits of agreement* sont transformées par l'exponentielle. La figure 4.11 présente la rapport des volumes en fonction de la moyenne géométrique de V_{Ref} et V_{Seg} , avec la courbe d'estimation du biais et les courbes d'estimation des *limits of agreement*. Pour les faibles volumes (moyenne géométrique inférieure à 30 cm^3), la valeur de V_{Seg} est sous-

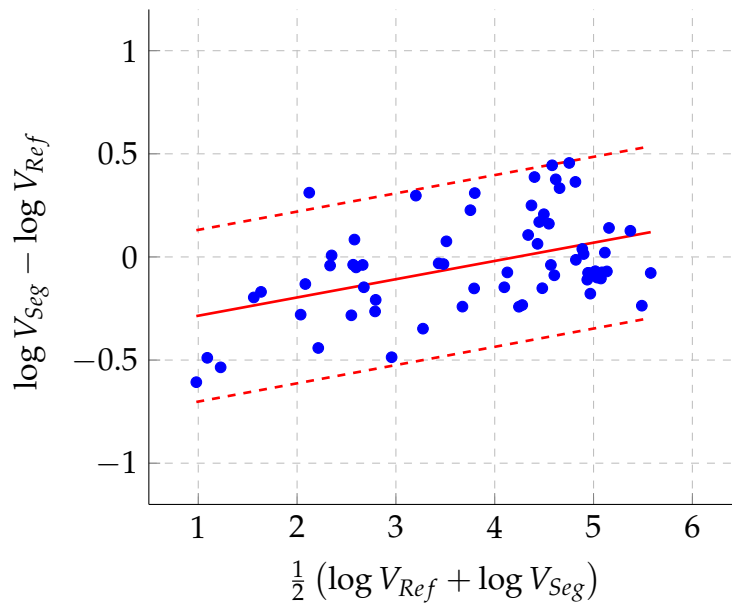


FIGURE 4.10 – Graphe de Bland-Altman entre $\log V_{Seg}$ et $\log V_{Ref}$. La ligne rouge continue indique l'estimation du biais, les lignes pointillées les *limits of agreement* comprenant 95% des points. Les deux sujets outliers sont exclus de l'analyse.

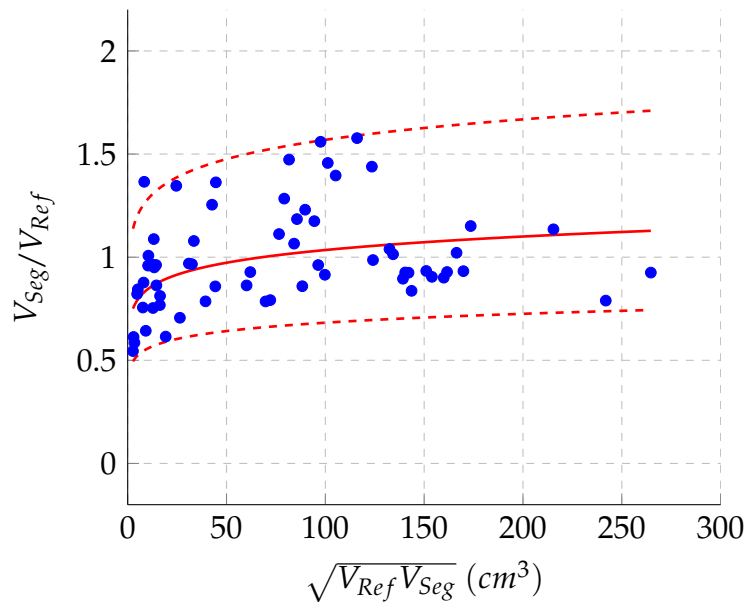


FIGURE 4.11 – Graphe de Bland-Altman obtenu après transformation réciproque de l'analyse des logarithmes.

Ce graphe est obtenu en passant à l'exponentielle les résultats présentés à la figure 4.10. Le passage à l'exponentielle permet une interprétation plus facile des résultats. La ligne rouge continue indique l'estimation du biais, les lignes pointillées les *limits of agreement* comprenant 95% des points. Les deux sujets outliers sont exclus de l'analyse.

estimée d'environ 20% par rapport à celle de V_{Ref} (dans 95% des cas, elle sera comprise entre une sous-estimation de 40% et une surestimation de 20% environ). Pour les plus gros volumes, le biais moyen s'approche de 1,1 avec des *limits of agreement* entre 0,8 et 1,6.

4.4.3 Validation spatiale

Sur l'ensemble des 67 sujets, la valeur médiane du SI est de 0,79. La valeur du premier quartile est de 0,70 et celle du 3^{ème} quartile de 0,83 (écart interquartile $\simeq 0,14$), indiquant un accord spatial globalement bon entre les segmentations automatiques et les segmentations manuelles.

La figure 4.12 montre la répartition du SI en fonction du volume. On remarque en particulier que le SI croît avec l'amplitude du volume. Cet effet est attendu et classique avec l'utilisation de cet indice. En effet, le SI est calculé proportionnellement au volume mesuré. Si celui-ci est très faible, il suffit de quelques voxels mal classés pour obtenir une valeur très basse. Néanmoins, les valeurs du SI croissent très rapidement et atteignent des valeurs supérieures à 0,6 y compris pour des faibles volumes de HSB (à partir d'environ 5 cm³).

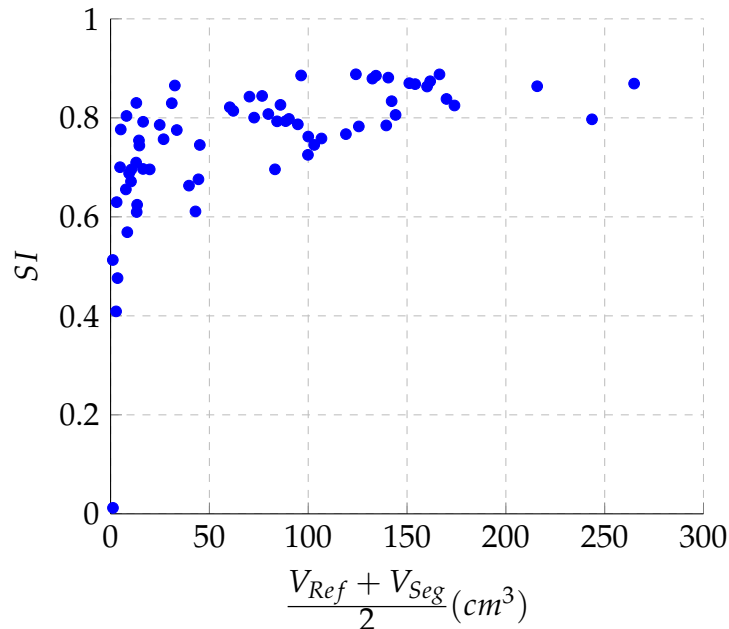


FIGURE 4.12 – Répartition des SI en fonction du volume.

4.5 Discussion

La section précédente a évalué la chaîne de traitements WHASA sur 67 patients, dont les images avaient été acquises dans six centres différents et présentaient une variabilité importante de charge lésionnelle (V_{Ref} compris entre 0,07 et 275 cm³). L'évaluation a été effectuée par comparaison de la segmentation obtenue automatiquement avec la segmentation de référence. Au niveau volumique, le biais moyen de la segmentation automatique est de 1,2 cm³, la surestimation maximale est de 53 cm³ et la sous-estimation maximale est de -57 cm³. Au niveau spatial, le SI médian sur les 67 sujets est de 0,79 et l'écart interquartile de 0,14.

4.5.1 Sujets ayant un mauvais accord spatial

Les sujets présentant un faible SI sont également les sujets ayant très peu de HSB. Le SI étant exprimé en proportion du volume moyen, même un très léger désaccord entre *Seg* et *Ref* entraînera une valeur très faible. Les figures 4.6 et 4.7 présentaient quelques coupes des deux sujets ayant les plus mauvais accords volumiques et qui présentent également les SI les plus faibles. Nous avons vu que la segmentation automatique demeurerait de qualité acceptable pour ces sujets.

4.5.2 Sujets ayant un mauvais accord volumique

Les figures 4.9 et 4.11 montrent la concordance entre le volume mesuré automatiquement et le volume manuel en fonction de la moyenne de ces deux volumes. En particulier, on s'aperçoit que les *limits of agreement* sur ces deux graphiques sont en partie influencées par un groupe de points dont le volume moyen est aux environs de 100 cm³ et le ratio proche de 1,5. Les figures 4.13 et 4.14 montrent plusieurs coupes des deux sujets présentant le ratio $\frac{V_{Seg}}{V_{Ref}}$ le plus important ainsi que la comparaison de *Ref* et de *Seg*.

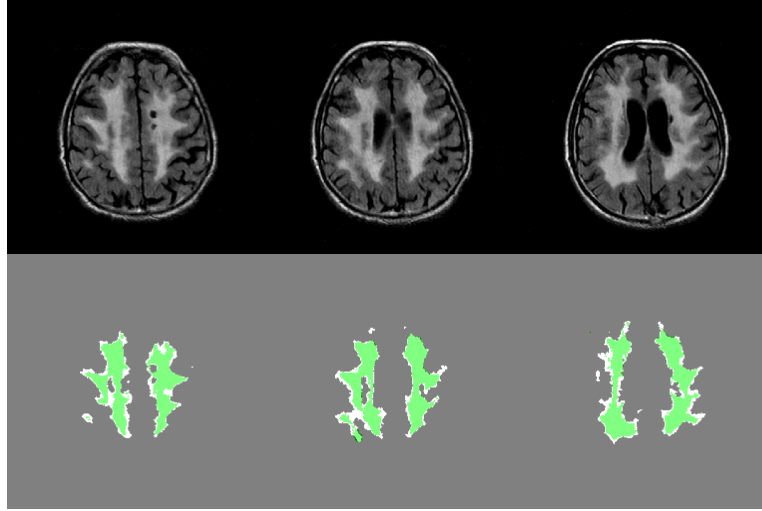


FIGURE 4.13 – Coupes illustratives d'un sujet ayant un mauvais accord volumique.

Pour ce sujet, $V_{Ref} = 92,4 \text{ cm}^3$, $V_{Seg} = 145,8 \text{ cm}^3$, $\sqrt{V_{Seg}V_{Ref}} = 116,1 \text{ cm}^3$, $V_{Seg}/V_{Ref} \simeq 1,58$. **Ligne du haut** : Coupes issues de l'image *mFLAIR*. **Ligne du bas** : Comparaison de *Ref* et *Seg* pour ces mêmes coupes. En vert clair sont montrés les voxels correctement segmentés (Vrais Positifs), en blanc ceux segmentés en trop (Faux Positifs) et en vert foncé ceux manqués par la segmentation automatique (Faux Négatifs).

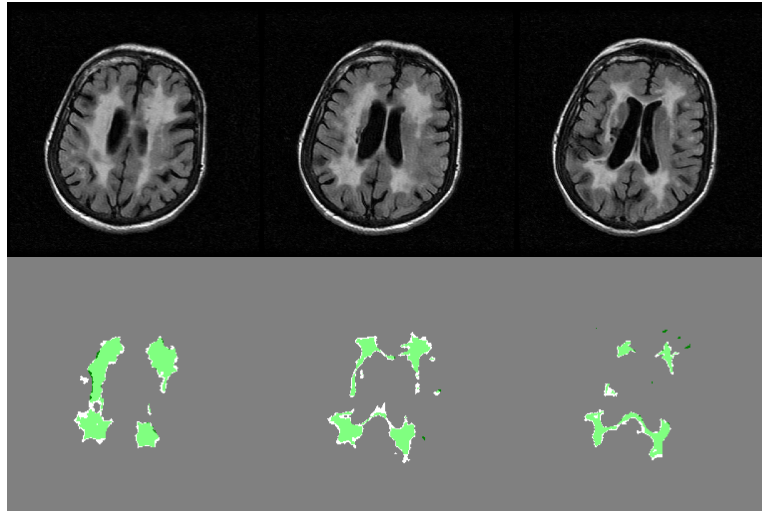


FIGURE 4.14 – Coupes illustratives d'un sujet ayant un mauvais accord volumique.

Pour ce sujet, $V_{Ref} = 78,2 \text{ cm}^3$, $V_{Seg} = 121,9 \text{ cm}^3$, $\sqrt{V_{Seg}V_{Ref}} = 97,6 \text{ cm}^3$, $V_{Seg}/V_{Ref} \simeq 1,56$. **Ligne du haut** : Coupes issues de l'image *mFLAIR*. **Ligne du bas** : Comparaison de *Ref* et *Seg* pour ces mêmes coupes. En vert clair sont montrés les voxels correctement segmentés (Vrais Positifs), en blanc ceux segmentés en trop (Faux Positifs) et en vert foncé ceux manqués par la segmentation automatique (Faux Négatifs).

La qualité de la segmentation demeure correcte. La segmentation automatique surestime effectivement globalement les HSB par rapport à Ref en épaississant tous les bords des lésions. Le volume total étant déjà important, le volume obtenu en ajoutant quelques rangées de voxels au bord des segmentations entraîne un ratio $\frac{V_{Seg}}{V_{Ref}}$ important. De plus, la segmentation automatique a également tendance à inclure l'ensemble des HSB confluentes alors que la segmentation manuelle tend dans ces cas-là à ne conserver que le coeur en hypersignal des HSB. Il est difficile dans ce genre de situations d'établir laquelle des deux méthodes est la plus exacte.

4.5.3 Interêt de l'étape de croissance

L'étape de croissance pourrait être la cause de la sursegmentation observée dans le cas des HSB confluentes. Nous avons donc comparé les résultats obtenus avant l'étape de croissance (HSB_{init} selon les notations du chapitre 3) avec ceux obtenus après l'étape de croissance.

La figure 4.15 montre les analyses par régression linéaire, d'abord pour les volumes segmentés puis pour leurs logarithmes. La droite de régression est beaucoup plus éloignée de la droite identité avant l'étape de croissance ($y = 0,70x + 5$ au lieu de $y = 0,95x + 5,3$). Pour l'analyse utilisant les logarithmes, les deux sujets outliers ont été exclus comme précédemment. La droite de régression met alors en évidence une sous-segmentation quasi systématique ($y = 1,1x - 0,57$).

La figure 4.16 présente, elle, les graphes de Bland-Altman avant et après l'étape de croissance. L'analyse des différences de volume montre que le biais avant l'étape de croissance décroît avec le volume mesuré ($y = -0,31x + 3,0$, $R^2 = 0,49$) alors qu'après croissance, le biais est constant et proche de 0 ($y = 1,23$, $R^2 = 0,32$). L'analyse en utilisant d'abord une transformation logarithmique confirme cette tendance. Les courbes d'estimation du ratio $\frac{V_{Seg}}{V_{Ref}}$ en fonction de la moyenne géométrique $\sqrt{(V_{Seg} V_{Ref})}$ ont la même allure avant et après croissance. Cependant, avant croissance, le biais est aux environs de 0,7 et est systématiquement inférieur à 1. Après croissance, en revanche, le biais varie entre 0,9 et 1,1. Les *limits of agreement* sont légèrement plus larges après croissance. Néanmoins, cet élargissement est essentiellement dû aux sujets ayant un ratio de 1,5. Comme nous l'avons vu précédemment, ces sujets présentent

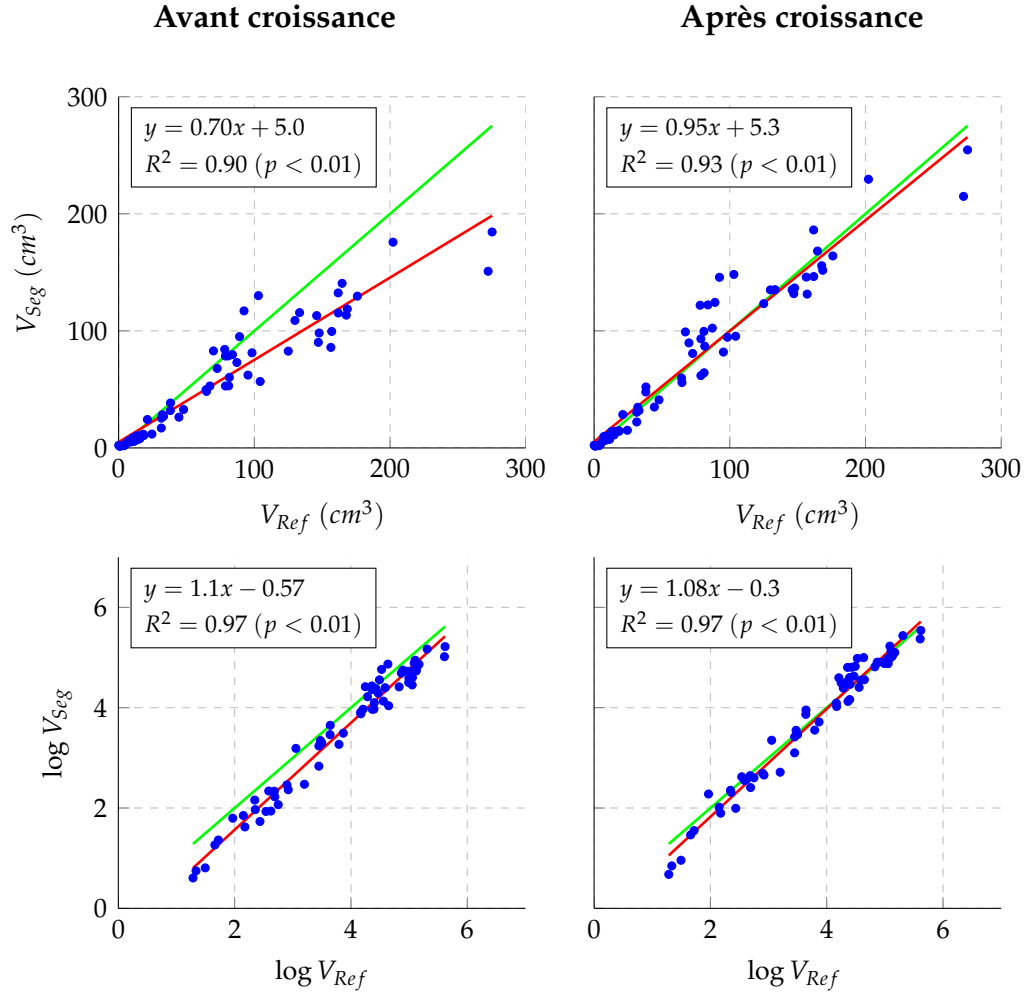


FIGURE 4.15 – Analyses par régression de la segmentation avant et après l'étape de croissance.

Ligne du haut : régression linéaire de V_{Seg} avant croissance par rapport à V_{Ref} . **Ligne du bas :** régression linéaire de $\log V_{Seg}$ avant croissance par rapport à $\log V_{Ref}$. Pour tous les graphiques, la droite identité est tracée en vert et la droite de régression en rouge.

la plupart du temps des HSB confluentes, dont le volume exact est difficile à déterminer.

Enfin, la figure 4.17 compare les distributions des SI avant et après l'étape de croissance. Le SI médian avant croissance est de 0,74, la valeur du premier quartile est de 0,63 et celle du troisième quartile de 0,79 (écart interquartile : 0,16). Après croissance, la valeur médiane du SI atteint 0,79, le premier quartile 0,70 et le troisième 0,83 (écart interquartile : 0,14). La différence est statistiquement significative (test de Wilcoxon pour échantillons appariés, $p < 0,05$).

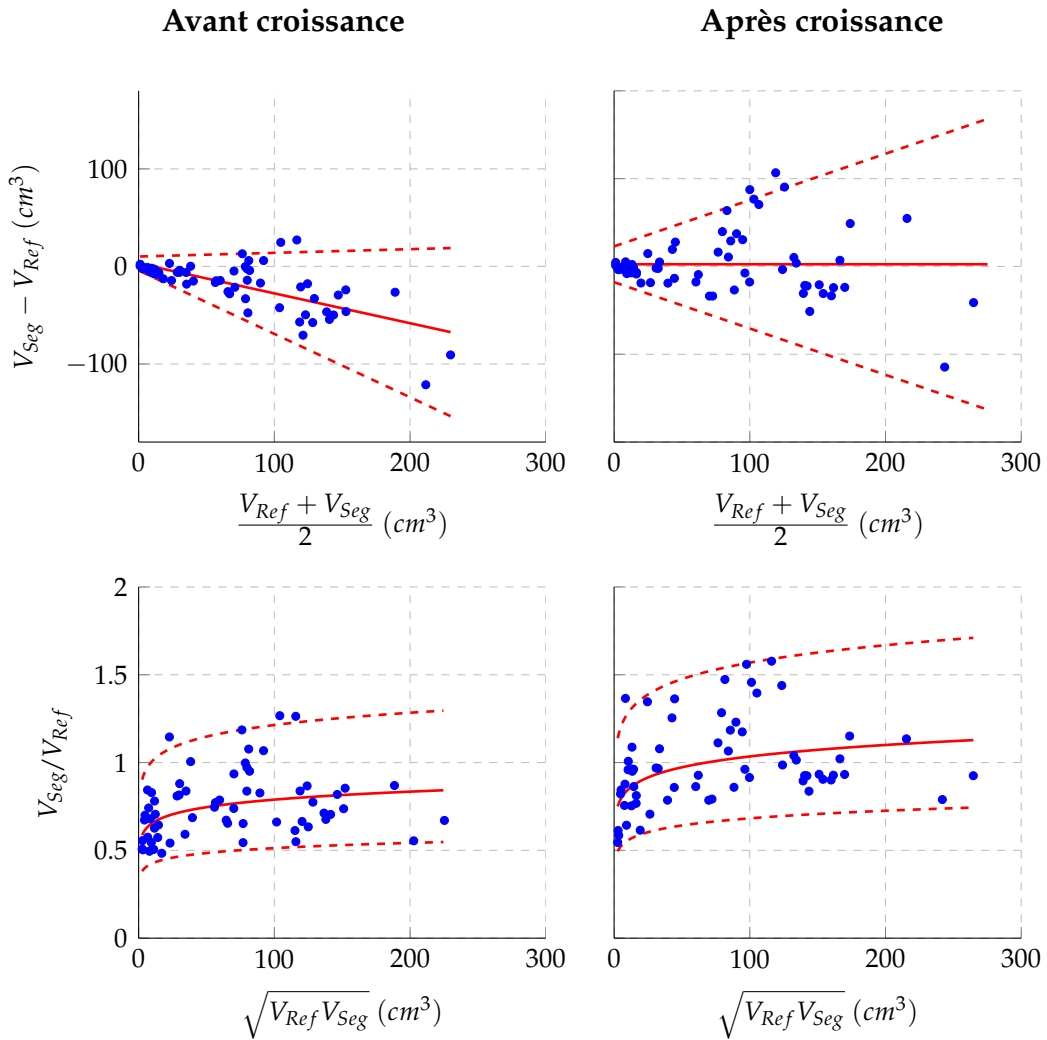


FIGURE 4.16 – Graphes de Bland-Altman avant et après l’étape de croissance.

Ligne du haut : Modélisation par régression linéaire de la variabilité en fonction du volume. **Ligne du bas :** Graphe de Bland-Altman obtenu après transformation réciproque de l’analyse des logarithmes (Les deux patients *outliers* sont exclus de l’analyse). Pour tous les graphiques, la ligne rouge continue indique l’estimation du biais, les lignes pointillées les *limits of agreement* comprenant 95% des points.

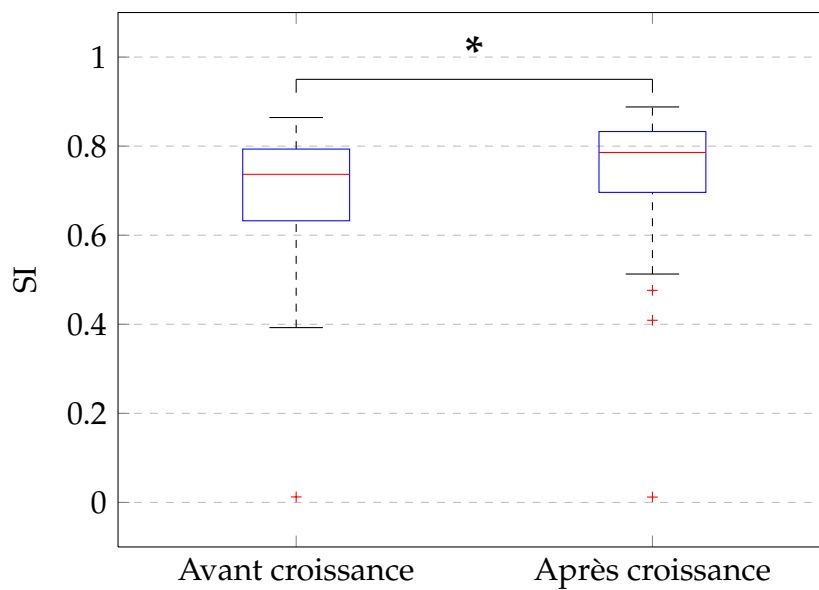


FIGURE 4.17 – Comparaison des distributions des SI pour les 67 sujets avant et après l'étape de croissance.

* : différence statistiquement significative (test de Wilcoxon pour échantillons appariés, $p < 0,05$)

4.5.4 Sensibilité à l'estimation du paramètre de contraste

Le paramètre de contraste utilisé pour obtenir une segmentation de l'image FLAIR est estimé automatiquement pour chaque sujet comme étant la moyenne du gradient de l'image FLAIR sur l'interface entre la substance blanche et la substance grise. Cependant, la présence d'importantes HSB, une mauvaise qualité des images, l'existence d'artefacts de mouvements, un effet de volume partiel important peuvent altérer la qualité de la segmentation en tissus et donc la précision de l'interface. Il est donc important de quantifier la robustesse des résultats de l'algorithme aux variations du paramètre de contraste.

Nous avons donc fait varier la valeur de λ entre $0,7\lambda_{est}$ et $1,3\lambda_{est}$ avec un pas $\Delta\lambda = 0,05\lambda_{est}$. La figure 4.18 montre les distributions des SI obtenus à travers les 67 sujets pour ces différentes valeurs du paramètre de contraste. La valeur médiane des SI varie peu : elle est de 0,73 pour $0,7\lambda_{est}$, de 0,79 pour λ_{est} et de 0,78 pour $1,3\lambda_{est}$. Statistiquement, les distributions des SI sont significativement différentes de celle obtenue avec λ_{est} uniquement pour $\lambda < 0,8\lambda_{est}$ ou $\lambda > 1,25\lambda_{est}$ (test de Friedman, corrigé pour les comparaisons multiples par le HSD de Tukey, $p < 0,05$). Des variations de l'ordre de 20% sur l'estimation de λ_{est} n'ont donc pas d'effet significatif sur les résultats obtenus.

4.5.5 Sensibilité au critère d'intensité

Un autre paramètre fixé empiriquement dans la chaîne de traitements WHASA est le seuil sur les intensités $s = \tilde{i} + 2\lambda_{est}$. Afin d'évaluer son influence sur les résultats, nous l'avons fait varier entre $s - 2\lambda_{est}$ et $s + 2\lambda_{est}$ avec un pas de $0,5\lambda_{est}$.

La figure 4.19 montre les distributions des SI obtenues à travers les 67 sujets en fonction de la valeur du seuil. Les variations sont importantes. L'évolution de la médiane des SI évolue en forme de cloche avec un maximum de 0,79, obtenu pour le seuil s et $s + 0,5\lambda_{est}$. Statistiquement, les distributions des SI sont significativement différentes de celle obtenue avec s lorsque le seuil est inférieur à $s - 0,5\lambda_{est}$ ou supérieur à $s + \lambda_{est}$. La valeur optimale pour le critère d'intensité est donc bien située aux environs de s et de faibles variations autour de cette valeur n'affecteront pas de manière significative les résultats obtenus.

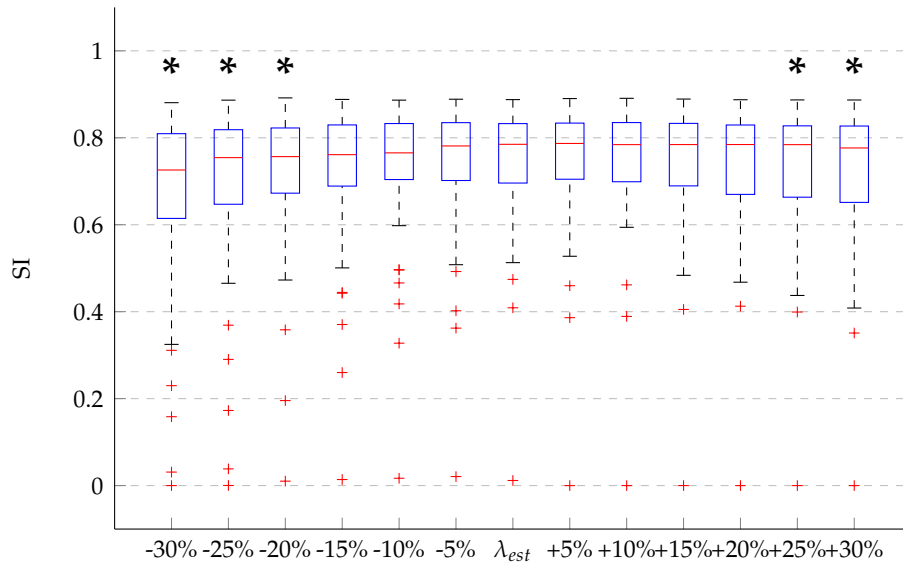


FIGURE 4.18 – Distribution des SI sur les 67 sujets en fonction des variations du paramètre de contraste.

* : différence statistiquement significative par rapport à la distribution des SI obtenus avec λ_{est} . (test de Friedman, corrigé pour les comparaisons multiples par le HSD de Tukey, $p < 0,05$).

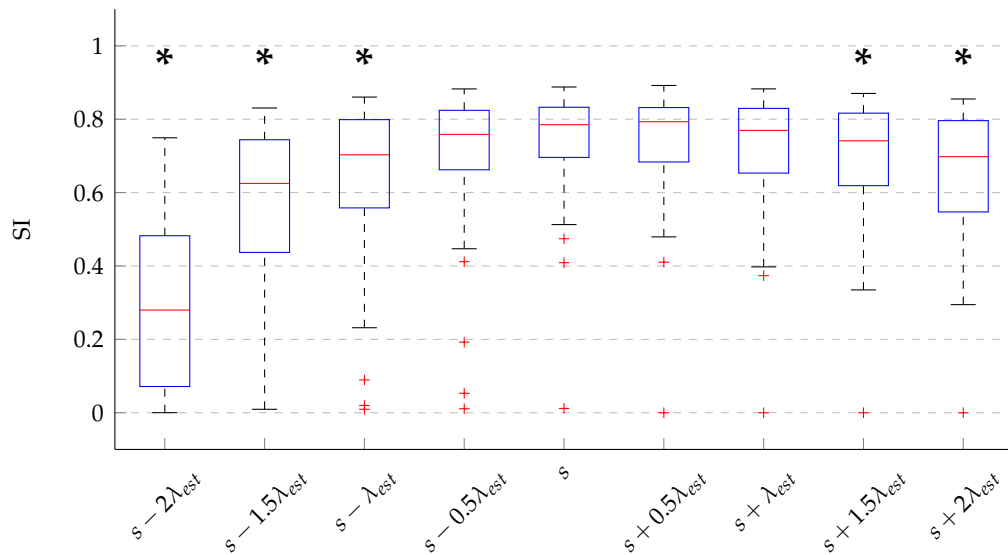


FIGURE 4.19 – Distribution des SI sur les 67 sujets en fonction des variations du seuil sur les intensités.

* : différence statistiquement significative par rapport à la distribution des SI obtenus avec le seuil s . (test de Friedman, corrigé pour les comparaisons multiples par le HSD de Tukey, $p < 0,05$).

4.6 Conclusion

Dans ce chapitre, nous avons évalué les performances de WHASA par comparaison avec une segmentation manuelle de référence sur 67 sujets, dont les images IRM ont été acquises dans six centres différents et présentaient une importante variabilité de charge lésionnelle. En moyenne, le volume des segmentations automatiques est équivalent aux volumes manuels (biais de $1,2 \text{ cm}^3$, non significativement différent de 0). Un fort recouvrement spatial est également observé sur l'ensemble des sujets avec un SI médian de 0,79 et un écart interquartile de 0,14.

L'analyse des sujets présentant un mauvais accord spatial ou un mauvais accord volumique montre que la qualité de la segmentation demeure toutefois acceptable. Pour ces sujets, la discordance entre la méthode automatique et la référence manuelle est souvent dû, soit à de très faibles HSB, soit à la présence d'HSB confluentes. Dans le premier cas, les indicateurs sont fortement affectés même par de très faibles erreurs de segmentation. Dans le deuxième, la délimitation exacte des HSB est impossible à définir, la segmentation de référence tendant à ne conserver que le coeur des HSB alors que la segmentation automatique prend également en compte la plage confluyente.

Par ailleurs, nous avons confirmé l'importance de l'étape de croissance, qui améliore de manière significative l'accord volumique et spatial. La robustesse de la chaîne de traitements WHASA à ses principaux paramètres a également été démontrée. Une variation de l'ordre de 20% sur le paramètre de contraste n'affecte pas de manière significative les résultats obtenus. De même, la valeur optimale pour le seuil des intensités est bien aux environs de la valeur s retenue et les résultats ne seront pas significativement différents pour de faibles variations de ce seuil.

Une évaluation complète de la méthode de segmentation nécessiterait enfin de comparer les résultats à ceux des autres méthodes de la littérature afin de situer ses performances. Cette comparaison est l'objet du chapitre suivant.

Chapitre 5

Comparaison aux autres méthodes de la littérature

Le chapitre 4 montre que les performances de WHASA sont satisfaisantes et robustes par rapport aux conditions d'acquisition et par rapport à la charge lésionnelle. L'état de l'art présenté au chapitre 2 montre néanmoins que de nombreuses approches avaient été précédemment proposées concernant la segmentation des HSB. Malheureusement, ces approches ont toujours été évaluées sur des données différentes, empêchant une réelle comparaison non biaisée de leurs performances, de leurs atouts et faiblesses.

Dans ce chapitre, nous avons donc choisi d'évaluer un large panel de méthodes, essayant de couvrir la plupart des approches proposées et de les évaluer sur les mêmes données. A notre connaissance, une seule étude a auparavant évalué différentes méthodes de segmentation des HSB sur un même jeu de données [Klöppel et al., 2011]. Celui-ci n'était cependant composé que de 20 sujets, dont les images avaient été acquises dans un seul centre. Nous nous sommes appuyés sur ce travail pour effectuer une comparaison sur une plus grande base de données et avons également comparé des méthodes supplémentaires. Le choix des méthodes présentées ici a été guidé par leur disponibilité ou la possibilité de les réimplémenter d'après les détails fournis dans les publications.

Les résultats présentés dans ce chapitre ont été obtenus en collaboration avec Ludovic Fillon, dans le cadre de son stage de Master 2, que j'ai co-encadré avec Marie Chupin.

5.1 Choix des méthodes

5.1.1 Méthodes non supervisées

Le logiciel Freesurfer est l'une des suites logicielles les plus utilisées actuellement en neuroimagerie. Elle est historiquement centrée sur l'analyse des surfaces corticales, qu'il s'agisse d'études morphométriques ou fonctionnelles. Une partie des outils développés dans Freesurfer permet la segmentation des structures sous-corticales ainsi que des HSB les plus importantes. Bien qu'une segmentation précise des HSB n'ait jamais été un objectif de Freesurfer, ce logiciel a été utilisé dans certaines publications récentes pour l'étude des HSB [Smith et al., 2011 ; Durgerian et al., 2011].

Plusieurs méthodes basées sur le principe du seuillage ont été développées pour la segmentation des HSB. Elles s'appuient essentiellement sur l'image FLAIR. La problématique est alors de déterminer le seuil optimal pour la segmentation des HSB. Jack et al. [2001] ont proposé dès 2001 de calculer ce seuil par des analyses statistiques de la forme de l'histogramme. Par la suite, Wen & Sachdev [2004] ont proposé de le déterminer en fonction de la moyenne et de l'écart-type des intensités de la substance blanche et de la substance grise. de Boer et al. [2009] ont repris cette approche mais en utilisant un prétraitement différent. Nous avons donc décidé de réimplémenter la méthode proposée par Wen & Sachdev [2004] puisqu'elle permet d'utiliser les mêmes prétraitements que les autres méthodes et semble plus robuste que l'approche proposée par Jack et al. [2001].

Trois articles décrivent une méthode s'appuyant sur les principes de logique floue et l'algorithme des *fuzzy C-means* [Admiraal-Behloul et al., 2005 ; Wu et al., 2006 ; Gibson et al., 2010]. La description de la méthode par Admiraal-Behloul et al. [2005] n'était pas suffisamment détaillée pour permettre une réimplémentation fidèle de la méthode. Wu et al. [2006] faisaient intervenir de nombreux pré et post-traitements, qui rendent la comparaison avec d'autres méthodes difficiles. Nous avons donc choisi d'utiliser la méthode proposée par Gibson et al. [2010], qui propose une implémentation disponible en ligne.

Enfin, un certain nombre de méthodes reposaient également sur un modèle de mélange de gaussiennes [Khayati et al., 2008 ; Gicquel, 1998 ; Maillard et al., 2008 ; Van Leemput et al., 2001]. Malheureusement, ces mé-

thodes font généralement intervenir de très nombreux paramètres et les détails présentés dans les publications ne permettaient pas une réimplémentation fidèle. Une version de la méthode proposée par Van Leemput et al. [2001] est disponible en ligne¹. Néanmoins, elle n’a pas été mise à jour depuis 2001. Nous avons essayé cette implémentation sur nos données mais les résultats étaient complètement incohérents. Nous n’avons donc pas évalué plus avant ce type de méthodes sur nos données.

5.1.2 Méthodes supervisées

Les méthodes supervisées désignent l’ensemble des méthodes qui utilisent une base de données, dite *base d’apprentissage*, pour inférer automatiquement des règles permettant la classification de nouvelles données, non incluses dans la base d’apprentissage. Nous nous intéressons ici à deux méthodes supervisées qui ont été appliquées au problème de segmentation des HSB : la classification par k plus proches voisins et les machines à vecteurs de support.

Les réseaux de neurones (*Artificial Neural Networks*) sont un autre exemple classique d’apprentissage supervisé qui a été utilisé dans le cas de la Sclérose en plaques [Zijdenbos et al., 2002]. Une application aux HSB a été proposée dans [Dyrby et al., 2008]. Cependant les résultats présentés étaient relativement décevants. De plus, un très grand nombre d’implémentations différentes des réseaux de neurones est possible et les détails présentés dans [Dyrby et al., 2008 ; Zijdenbos et al., 2002] n’étaient pas suffisants pour permettre de réimplémenter les méthodes de manière similaire.

5.2 Implémentation des méthodes non supervisées

5.2.1 Freesurfer

Freesurfer (Massachussets General Hospital, Boston, MA) est disponible librement en ligne à l’adresse suivante : <http://surfer.nmr.mgh.harvard.edu/>. Nous avons utilisé la version 5.0.0 en conservant les para-

1. <http://mirc.uzleuven.be/MedicalImageComputing/downloads/ems.php?EMSsection=about>

mètres par défaut. Les détails techniques de ce logiciel sont décrits dans la documentation disponible sur le site ou, principalement, dans les articles suivants [Sled et al., 1998 ; Fischl et al., 1999, 2002, 2004 ; Dale et al., 1999 ; Ségonne et al., 2007].

Il est à noter que Freesurfer ne s'appuie que sur l'image pondérée en T1 et que son but premier est la segmentation des structures anatomiques (hippocampe, noyau caudé, putamen, parcellisation du cortex...) et l'analyse des surfaces corticales ainsi produites. La segmentation des HSB n'est qu'un dérivé des traitements effectués. La segmentation des HSB produite par Freesurfer a néanmoins été utilisée récemment dans plusieurs publications [Smith et al., 2011 ; Durgerian et al., 2011].

5.2.2 Seuillage

Wen & Sachdev [2004] segmentent les HSB en déterminant deux seuils à partir de l'image FLAIR et de l'image T1, suite à deux étapes de prétraitement.

- Un masque du cerveau est tout d'abord créé à partir du logiciel SPM, similaire à celui présenté dans la partie 3.2.4.
- Afin de renforcer l'importance des voxels situés dans la substance blanche, l'image $mFLAIR$ est pondérée de la manière suivante : $I_{new}^{FLAIR} = c I_{old}^{FLAIR} SB_{sujet}$, où SB_{sujet} désigne la carte de probabilité de substance blanche obtenue par SPM et c est une constante, dont il est seulement indiqué dans l'article que $c \geq 2$.

Un premier seuil est alors déterminé pour caractériser les voxels hyperintenses. Sur l'image FLAIR, la moyenne des intensités de la substance blanche (μ_{SB}^{FLAIR}) et leur écart-type (σ_{SB}^{FLAIR}) sont calculés en utilisant la segmentation de la substance blanche fournie par SPM, SB_{sujet} . Les voxels hyperintenses sont caractérisés par la relation $I_{new}^{FLAIR} \geq \mu_{SB}^{FLAIR} + 3\sigma_{SB}^{FLAIR}$.

Un second seuil permet d'éliminer les faux positifs présents dans les ventricules. A cette fin, la moyenne et l'écart-type des intensités du LCS sont calculés sur l'image T1 (μ_{LCS}^{T1} et σ_{LCS}^{T1}). Seuls les voxels vérifiant la relation $I^{T1} \geq \mu_{LCS}^{T1} - 3\sigma_{LCS}^{T1}$ sont conservés comme WMH.

La seule information manquante pour la réimplémentation de cette procédure concerne la valeur de c , qui n'est pas précisée. Nous avons donc fait varier la valeur de c dans l'intervalle $[1.5, 3.5]$ avec un pas de 0.1. La valeur optimale de c est définie comme celle permettant d'obtenir le

meilleur SI moyen sur l'ensemble des 67 sujets. La valeur de 2,1 fut ainsi retenue.

5.2.3 SIPS

Gibson et al. [2010] ont présenté une méthode s'appuyant sur les principes de logique floue à travers l'algorithme de classification *Fuzzy C-means*. Une implémentation de cette méthode est disponible librement en ligne². Nous avons donc utilisé cette implémentation avec les paramètres par défaut.

5.3 Implémentation des méthodes supervisées

Notre problème de segmentation peut également être vu comme un problème d'apprentissage automatique. En effet pour tout voxel x d'une image, nous cherchons à lui attribuer un label y indiquant s'il appartient ou non à une HSB, à partir de l'étude de certaines caractéristiques de ce voxel (par exemple, son intensité sur l'image FLAIR, sa probabilité d'appartenance à la substance blanche, etc...). Les méthodes d'apprentissage supervisé cherchent à déterminer automatiquement une décision de classification à partir d'un *ensemble d'apprentissage*.

Cet ensemble d'apprentissage est composé de N observations pour lesquelles le label y est connu, $\{(x_i, y_i)\}$. Nous détaillons sa construction dans le paragraphe 5.3.1. La décision de classification obtenue par les méthodes supervisées est construite uniquement sur les caractéristiques associées au voxel, qui représentent donc l'information fournie au classifieur. L'ensemble des caractéristiques considérées est rassemblé dans un vecteur de dimension d , appelé **vecteur de caractéristiques** (ou *feature vector* en anglais). Les vecteurs de caractéristiques considérés pour l'évaluation sont décrits dans le paragraphe 5.3.2. Enfin, il existe plusieurs méthodes de classification, nous détaillerons dans le paragraphe 5.3.3 l'algorithme de classification par k -plus proches voisins (ou kNN pour *k Nearest Neighbours*) et dans le paragraphe 5.3.4 les machines à vecteurs de support (ou SVM pour *Support Vector Machines*).

2. ftp://ftp.rotman-baycrest.on.ca/pub/Nancy/SunnybrookImageProcessingSoftware/sips_package_v1.0/

5.3.1 Définition de l'ensemble d'apprentissage

Les méthodes supervisées construisent leur fonction de décision à partir des données qui leur sont fournies dans l'ensemble d'apprentissage. Le choix de ce dernier est donc crucial. Il doit contenir suffisamment d'information pour que la fonction de décision obtenue par l'algorithme soit généralisable à toute nouvelle donnée présentée. Il est ainsi possible qu'une méthode supervisée soit très performante pour un jeu de données particulier, mais obtienne des résultats médiocres lorsque de nouvelles données lui sont présentées. On parle dans ce cas de *surapprentissage*. L'influence de la base d'apprentissage sur les performances de la méthode est rarement évaluée dans la littérature. Elle est pourtant intrinsèquement liée à la généralisabilité de ces méthodes.

5.3.1.1 Choix des sujets

Trois bases d'apprentissage ont été utilisées :

- base d'apprentissage 1 : 10 sujets issus de la base de données *Hippocampe*, choisis de manière à être représentatifs des différents centres d'acquisition ;
- base d'apprentissage 2 : 10 sujets issus de la base de données *Cadasil*, choisis de manière à refléter la variabilité des charges lésionnelles ;
- base d'apprentissage 3 : 5 sujets issus de la base de données *Hippocampe* et 5 sujets issus de la base de données *Cadasil*.

Dans chaque cas, nous définissons un ensemble de test composé des sujets qui n'appartiennent pas à la base d'apprentissage. L'évaluation des performances de segmentation est faite uniquement sur les sujets de l'ensemble de test, en utilisant les paramètres optimaux déterminés sur l'ensemble d'apprentissage.

5.3.1.2 Choix des voxels

Si l'on considère tous les voxels des images FLAIR des sujets de la base d'apprentissage, la classe représentant les HSB sera très fortement sous-représentée par rapport à l'autre classe. De plus, cela augmente le temps de calcul pour des voxels qui n'apportent aucune information tels que ceux du fond ou du crâne. Afin de rester le plus proche possible du travail de comparaison de Klöppel et al. [2011], nous avons adopté une définition similaire pour l'ensemble d'apprentissage.

500 voxels labellisés comme HSB dans la segmentation manuelle de référence ont ainsi été sélectionnés aléatoirement pour chaque patient de la base d'apprentissage. Les voxels "sains" regroupent des voxels très différents, qui peuvent être situés parfois loin des HSB. Une frontière de 5mm de large a alors été définie autour des HSB : 250 voxels "sains" ont été aléatoirement sélectionnés à l'intérieur de cette frontière et 250 l'ont été à l'extérieur. Pour chaque base d'apprentissage, le nombre total de voxels considérés est donc au final constitué de 10000 voxels (1000 voxels par patient).

5.3.2 Définition du vecteur de caractéristiques

Le vecteur de caractéristiques regroupe l'ensemble des informations à partir duquel la méthode de classification va séparer les classes. Dans le cas des HSB, il est possible de considérer les caractéristiques qui relèvent de l'intensité et les caractéristiques qui relèvent de la localisation.

5.3.2.1 Normalisation des intensités

Les méthodes supervisées classent de nouveaux voxels par comparaison de ses caractéristiques avec celles des voxels constituant l'ensemble d'apprentissage. Il est donc nécessaire que les caractéristiques utilisées soient comparables à travers les sujets. Les probabilités d'appartenance à un type de tissus satisfont cette condition. En revanche, les intensités des images IRM, dans le cas des séquences T1 ou FLAIR, ne correspondent pas à une grandeur physique. Cela est encore plus important pour des études multicentriques. Les intensités des images $mT1$ et $mFLAIR$ sont donc normalisées pour chaque sujet de manière à ce que la valeur médiane des intensités sur le masque du cerveau soit 0 et l'écart interquartile soit de 1.

5.3.2.2 Voisinage d'un voxel

Il semble également pertinent de prendre en compte les informations du voisinage du voxel. En effet, la distribution des intensités dans le voisinage d'un voxel situé au bord du ventricule sera différente de celle d'un voxel de la substance blanche profonde. Cette information supplémentaire entraîne l'augmentation de la dimension du vecteur de caractéristiques,

mais les SVM sont adaptés aux problèmes de grande dimension. Lao et al. [2008] utilisent comme voisinage un cube de 5mm de côté alors que Klöppel et al. [2011] s'appuient sur une sphère de rayon 8mm.

Dans notre cas, les données proviennent de centres différents avec des tailles de voxels différentes. L'épaisseur de coupe varie ainsi entre 4 et 6,2 mm. Or, il est nécessaire que le nombre de voxels considérés soit identique pour pouvoir comparer les vecteurs de caractéristiques. Nous avons privilégié le voisinage sphérique afin de ne pas donner trop de poids aux voxels situés dans les coupes supérieures et inférieures au voxel considéré. Afin d'obtenir un nombre de voxels inclus dans le voisinage identique, nous avons considéré la discrétisation d'une sphère de rayon 8 mm sur une image de taille $0,9375 \times 0,9375 \times 5,5 \text{ mm}^3$, qui correspond à la médiane de la taille des voxels à travers les centres. Cette discrétisation aboutit à un motif de 467 voxels³ (figure 5.1). Nous avons ensuite appliqué ce motif de voisinage pour toutes les images, indépendamment de la taille des voxels.

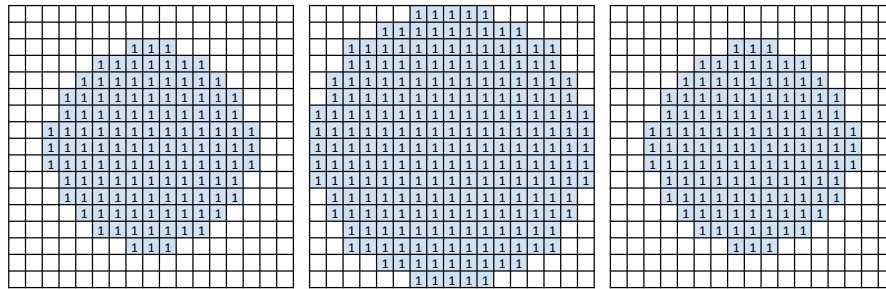


FIGURE 5.1 – Motif de voisinage obtenu par la discrétisation d'une sphère de rayon 8 mm sur une grille de taille $0,9375 \times 0,9375 \times 5,5 \text{ mm}^3$.
Gauche : coupe inférieure. **Centre** : coupe centrale. **Droite** : coupe supérieure.

5.3.2.3 Coordonnées dans l'espace MNI

Une manière différente d'inclure une information de localisation serait d'intégrer les coordonnées du voxel v . Pour pouvoir être comparées à travers les patients, ces coordonnées doivent être définies dans un espace de référence commun à tous les sujets, tel que l'espace MNI. Pour chaque voxel v , nous avons donc utilisé la transformation entre l'espace du sujet et l'espace MNI, calculée lors des prétraitements, pour déterminer les

3. le voisinage utilisé par Klöppel et al. [2011] comprenait 234 voxels, les images ayant été interpolées à une résolution dans le plan de $2 \times 2 \text{ mm}^2$.

coordonnées cartésiennes $x_{MNI}^v, y_{MNI}^v, z_{MNI}^v$ du voxel v dans l'espace MNI.

5.3.2.4 Vecteurs de caractéristiques utilisés

Quatre vecteurs de caractéristiques (VC) différents ont été évalués (figure 5.2) :

- VC1 : vecteur de dimension 3, regroupant l'intensité FLAIR normalisée, l'intensité T1 normalisée et la probabilité d'appartenance à SB_{MNI} du voxel considéré ;
- VC2 : vecteur de dimension 6, regroupant l'intensité FLAIR normalisée, l'intensité T1 normalisée, la probabilité d'appartenance à SB_{MNI} du voxel considéré et les 3 coordonnées cartésiennes de ce voxel dans l'espace MNI ;
- VC3 : vecteur de dimension 1401, regroupant l'intensité FLAIR normalisée du voxel considéré ainsi que celle des voxels du voisinage, l'intensité T1 normalisée du voxel et de son voisinage et la probabilité d'appartenance à SB_{MNI} du voxel et de son voisinage ;
- VC4 : vecteur de dimension 1404, regroupant l'intensité FLAIR normalisée du voxel considéré ainsi que celle des voxels du voisinage, l'intensité T1 normalisée du voxel et de son voisinage, la probabilité d'appartenance à SB_{MNI} du voxel et de son voisinage et les 3 coordonnées cartésiennes du voxel considéré dans l'espace MNI.

$$\begin{aligned}
 \text{VC1} &: (\underbrace{I_{T1}^v, I_{FLAIR}^v, p_{SB_{MNI}}^v}_{3 \text{ valeurs}}) \\
 \text{VC2} &: (\underbrace{I_{T1}^v, I_{FLAIR}^v, p_{SB_{MNI}}^v}_{3 \text{ valeurs}}, \underbrace{x_{MNI}^v, y_{MNI}^v, z_{MNI}^v}_{3 \text{ valeurs}}) \\
 \text{VC3} &: (\underbrace{I_{T1}^{v_1}, \dots, I_{T1}^{v_n}}_{467 \text{ valeurs}}, \underbrace{I_{FLAIR}^{v_1}, \dots, I_{FLAIR}^{v_n}}_{467 \text{ valeurs}}, \underbrace{p_{SB_{MNI}}^{v_1}, \dots, p_{SB_{MNI}}^{v_n}}_{467 \text{ valeurs}}) \\
 \text{VC4} &: (\underbrace{I_{T1}^{v_1}, \dots, I_{T1}^{v_n}}_{467 \text{ valeurs}}, \underbrace{I_{FLAIR}^{v_1}, \dots, I_{FLAIR}^{v_n}}_{467 \text{ valeurs}}, \underbrace{p_{SB_{MNI}}^{v_1}, \dots, p_{SB_{MNI}}^{v_n}}_{467 \text{ valeurs}}, x_{MNI}^v, y_{MNI}^v, z_{MNI}^v)
 \end{aligned}$$

FIGURE 5.2 – Construction des quatre vecteurs de caractéristiques (VC) évalués.

5.3.3 Classification par k plus proches voisins (kNN)

La méthode de classification des k plus proches voisins consiste, pour chaque nouvel élément que l'on souhaite classer, à regarder la classe des k points de l'ensemble d'apprentissage les plus proches et à affecter le nouvel élément à la classe la plus représentative (figure 5.3).

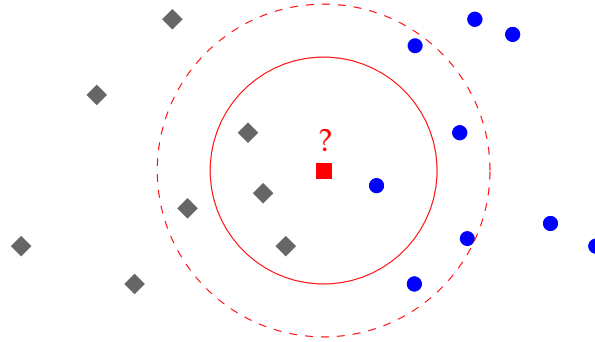


FIGURE 5.3 – Principe de la classification par k plus proches voisins (kNN).

L'ensemble d'apprentissage est constitué de deux classes, l'une symbolisée par des losanges gris, l'autre par des cercles bleus. On considère un nouvel élément (symbolisé par un carré rouge) et l'on souhaite savoir s'il appartient à la classe grise ou à la classe bleue. On regarde pour cela à quelle classe appartiennent ses k voisins les plus proches. Ici, on voit que, le nouvel élément étant situé à la frontière de deux classes proches, le résultat sera sensible au nombre de voisins k considéré. Si l'on prend en compte les 4 plus proches voisins (cercle rouge continu), le nouvel élément sera classé gris. En revanche, si l'on considère les 8 plus proches voisins, le nouvel élément sera classé bleu.

Cette approche repose donc sur peu de paramètres. Une fois l'ensemble d'apprentissage et la notion de distance utilisée (dans notre cas, la distance euclidienne classique) définis, il suffit de choisir le nombre de voisins à considérer. Anbeek et al. [2004] ont indiqué que, dans le cas des HSB, un nombre de voisins $k=100$, était un compromis satisfaisant entre performance et temps de calcul. Klöppel et al. [2011] utilisent également $k=100$ dans leur article de comparaison, valeur que nous reprenons donc dans notre implémentation.

Classiquement, pour le kNN, un nouvel élément est affecté à la classe majoritaire. Cependant, Anbeek et al. [2004] proposent, pour chaque voxel, de considérer la fraction de voisins HSB comme la probabilité pour ce voxel d'appartenir à la classe HSB. La segmentation finale est alors obtenue en seuillant cette carte de probabilité. En faisant varier ce seuil, Anbeek et al. [2004] construisent ainsi des courbes de sensibilité et spécificité et montrent que le seuil optimal pour leur données se situe aux environs de 0,3. Dans leur article de comparaison, Klöppel et al. [2011]

effectuent la même démarche. Le seuil optimal est déterminé comme étant celui fournissant le meilleur SI moyen sur l'ensemble d'apprentissage et est ensuite appliqué à l'ensemble test.

5.3.4 Classification par SVM

La classification par les machines à vecteurs de support (SVM) est une méthode de classification appartenant à la catégorie des classifieurs à large marge. Elle a été introduite par Vapnik [1995] et Cortes & Vapnik [1995]. Une excellente introduction aux SVM appliqués à la neuroimagerie peut être trouvée dans Cuingnet [2011]. Nous n'introduisons ici que l'interprétation géométrique permettant la définition de la fonction de classification ainsi que la notion de noyau.

Considérons l'ensemble d'apprentissage $\{(\mathbf{x}_i, y_i)\}_{i \in [1, N]}$ où N est le nombre de voxels de l'ensemble d'apprentissage, $\mathbf{x}_i$ le vecteur de caractéristiques associé à un voxel i et y_i est le label associé à ce voxel (+1 s'il appartient à une HSB, -1 sinon). La classification par SVM considère que les observations $\mathbf{x}_i$ appartiennent à un espace de Hilbert $\mathcal{H}$. Nous notons $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ le produit scalaire dans $\mathcal{H}$. Un hyperplan de $\mathcal{H}$ est alors défini par $\{\mathbf{x} | \langle \mathbf{w}, \mathbf{x} \rangle_{\mathcal{H}} + b = 0\}$, où $\mathbf{w}$ est un vecteur normal à l'hyperplan. De même, il est possible de définir deux hyperplans parallèles par $\{\mathbf{x} | \langle \mathbf{w}, \mathbf{x} \rangle_{\mathcal{H}} + b = 1\}$ et $\{\mathbf{x} | \langle \mathbf{w}, \mathbf{x} \rangle_{\mathcal{H}} + b = -1\}$. La distance entre ces deux hyperplans est alors appelé la **marge** m avec $m = \frac{2}{\|\mathbf{w}\|_{\mathcal{H}}}$.

Le principe du SVM est de trouver l'hyperplan, paramétrisé par $\mathbf{w}^{opt}$ et b^{opt} , qui permet de maximiser la marge, c'est-à-dire de séparer au mieux les vecteurs avec un label -1 de ceux avec un label +1 (figure 5.4). Cela est obtenu par l'intermédiaire du problème d'optimisation sous contraintes suivant :

$$(\mathbf{w}^{opt}, b^{opt}, \xi^{opt}) = \arg \min_{\mathbf{w} \in \mathcal{H}, b \in \mathbb{R}, \xi \in \mathbb{R}^N} \frac{1}{2} \|\mathbf{w}\|_{\mathcal{H}}^2 + C \sum_{i=1}^N \xi_i$$

$$\text{avec : } \begin{cases} y_i [\langle \mathbf{w}, \mathbf{x}_i \rangle_{\mathcal{H}} + b] \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases}$$

m étant inversement proportionnelle à $\|\mathbf{w}\|_{\mathcal{H}}$, la maximisation de la marge est obtenue par la minimisation de $\frac{1}{2} \|\mathbf{w}\|_{\mathcal{H}}^2$, le paramètre C est un paramètre de régularisation permettant de pénaliser de manière plus ou

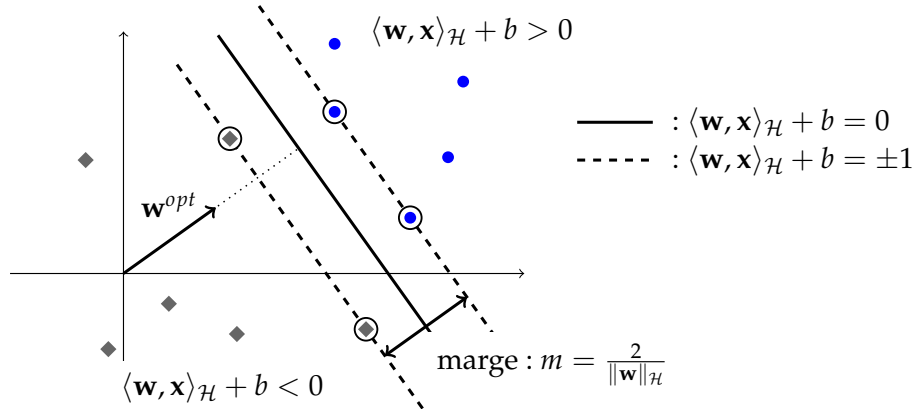


FIGURE 5.4 – Schéma de principe de la classification par machines à vecteurs de support (SVM) (Source : Cuingnet [2011]).

L'ensemble d'apprentissage est constitué de deux classes, l'une symbolisée par des losanges gris, l'autre par des cercles bleus. Le SVM construit l'hyperplan (dans cet exemple, nous sommes en dimension 2, un hyperplan est donc une droite) maximisant la marge, c'est-à-dire permettant de séparer au mieux les deux classes. La construction de cet hyperplan ne dépend au final que d'un petit nombre de vecteurs proches de l'hyperplan. Ces vecteurs sont appelés vecteurs supports et sont entourés en noir sur la figure.

moins importante les erreurs d'apprentissages et les variables ζ_i , appelées **variables ressort** (*slack variables*), permettent d'introduire une tolérance par rapport aux vecteurs mal classés ou situés près de l'hyperplan. La résolution de ce problème d'optimisation utilise les multiplicateurs de Lagrange et permet de montrer que seuls les vecteurs situés sur la marge ou les vecteurs mal classés contribuent à la définition de l'hyperplan. Ces vecteurs sont appelés **vecteurs support**.

Le formalisme sur lequel s'appuie le SVM repose sur la structure d'un espace de Hilbert pour construire un classifieur linéaire (l'hyperplan). Il est également possible d'adapter le SVM au cas non-linéaire en utilisant l'*astuce du noyau* [Aronszajn, 1950 ; Aizerman et al., 1964]. L'idée sous-jacente est de transformer, par l'intermédiaire d'une fonction ϕ , les données de l'espace de départ dans un espace de plus grande dimension, dans lequel il sera possible de déterminer un hyperplan séparant les données. En pratique, on considère une fonction K appartenant à $\mathbb{R}^{\mathcal{X} \times \mathcal{X}}$, où $\mathcal{X}$ est l'ensemble des observations. Si K est une fonction noyau définie positive, il existe alors une fonction ϕ telle que $K(\mathbf{x}, \mathbf{y}) = \langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle$. Tous les termes utilisés dans la fonctionnelle du SVM étant exprimés en fonction du produit scalaire des éléments, le calcul dans l'espace de grande dimension induit par ϕ peut s'exprimer uniquement en fonction de K .

Nous avons utilisé la version 3.1 de la bibliothèque libSVM disponible en ligne⁴ pour la résolution du SVM [Chang & Lin, 2011].

5.3.4.1 Choix du noyau

Il existe différents types de noyaux classiques pour la classification par SVM : noyau linéaire, noyau polynomial, noyau gaussien... Le noyau linéaire est préféré dans le cas des données linéairement séparables. Ce n'est pas généralement pas le cas en faible dimension, comme pour les vecteurs VC1 et VC2. Ce n'est pas non plus le cas pour VC4, qui utilise des coordonnées spatiales. En effet, un classifieur linéaire sur les coordonnées spatiales reviendrait à séparer l'image par un plan.

Nous utilisons donc le noyau gaussien (ou *RBF* pour *Radial Basis Functions*), qui est un choix classique pour le SVM et nécessite un paramètre de moins que les noyaux polynomiaux. Le noyau gaussien est également celui utilisé par Klöppel et al. [2011] et Lao et al. [2008]. Il est défini par :

$$K_{\gamma}(u, v) = \exp^{-\gamma \|u-v\|^2}$$

5.3.4.2 Optimisation des hyperparamètres

L'utilisation du SVM avec un noyau gaussien nécessite donc le réglage de deux hyperparamètres :

- le paramètre C , qui permet une régularisation de la fonction de classification ;
- le paramètre γ , qui intervient dans le noyau gaussien.

Les valeurs de ces hyperparamètres sont classiquement estimées par une recherche en grille sur l'ensemble d'apprentissage, de manière à fournir les meilleurs résultats de classification, estimés par validation croisée.

Nous avons procédé à une double recherche en grille (figure 5.5). Tout d'abord, une recherche grossière est effectuée, où C varie entre 2^{-5} et 2^{13} et γ entre 2^{-21} et 2^5 avec un pas de 2^2 . (C_{max}, γ_{max}) indique le couple obtenant la meilleure exactitude⁵ en validation croisée sur cette grille grossière. Une recherche en grille plus fine est ensuite conduite sur le sous-ensemble $[2^{C_{max}-2}, 2^{C_{max}+2}] \times [2^{\gamma_{max}-2}, 2^{\gamma_{max}+2}]$ avec un pas de $2^{0.25}$. Les

4. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

5. L'exactitude est définie comme la proportion des classifications correctes sur la totalité de l'ensemble (soit $\frac{TP+TN}{TP+FP+FN+TN}$)

5. COMPARAISON AUX AUTRES MÉTHODES DE LA LITTÉRATURE

hyperparamètres optimaux (C^{opt}, γ^{opt}) sont ceux obtenant la meilleure exactitude en validation croisée sur ce sous-ensemble. Le classifieur final est enfin obtenu en relançant la minimisation avec les hyperparamètres (C_{opt}, γ_{opt}) sur l'intégralité de l'ensemble d'apprentissage.

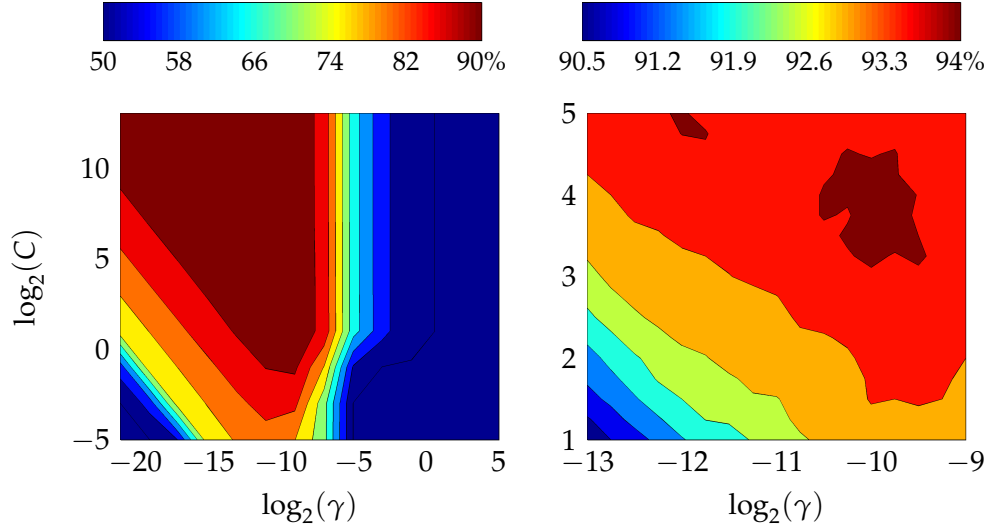


FIGURE 5.5 – Procédure de recherche des hyperparamètres du SVM.

Les deux hyperparamètres à optimiser pour le SVM sont le paramètre C de régularisation du SVM et le paramètre γ du noyau gaussien. La combinaison optimale de ces deux facteurs est déterminée par une double recherche en grille par validation croisée. L'ensemble d'apprentissage est divisé aléatoirement en deux parties. Pour chaque valeur de C et γ , on calcule l'hyperplan optimal sur la moitié de l'ensemble d'apprentissage et on classe la moitié restante grâce à l'hyperplan obtenu. On inverse ensuite les deux parties de manière à obtenir une classification pour tous les éléments de l'ensemble d'apprentissage. On répète ensuite cette procédure pour différentes valeurs de C et γ . Les figures représentent les courbes de niveau de l'exactitude de la classification obtenue pour chaque couple $(C; \gamma)$. **Gauche** : Recherche en grille grossière sur un large domaine et avec un pas important (C varie entre 2^{-5} et 2^{13} , γ varie entre 2^{-21} et 2^5 , avec un pas de 2^2 dans les deux cas). L'exactitude maximale est de 93,7% obtenue pour le couple $(C_{max}, \gamma_{max}) = (3; -11)$. **Droite** : Recherche en grille plus fine centrée autour du couple $(C_{max}; \gamma_{max})$. Cette fois, C varie entre 2^1 et 2^5 , γ varie entre 2^{-13} et 2^{-9} , avec un pas de $2^{0,25}$ dans les deux cas. L'exactitude maximale obtenue est de 94,1% pour le couple $(C_{opt}; \gamma_{opt}) = (3,5; -10)$.

5.3.4.3 Binarisation de la fonction de classification

Klöppel et al. [2011] suggèrent de plus d'analyser plusieurs seuils pour générer des courbes de sensibilité et de spécificité similaires à celles obtenues par le seuillage de la proportion de voisins HSB dans le cas du KNN. Nous avons donc également implémenté cette approche. Le seuil

optimal de la fonction de classification est défini comme étant celui qui permet d'obtenir le meilleur SI sur l'ensemble d'apprentissage.

5.4 Résultats

5.4.1 Comparaison avec les méthodes non supervisées

5.4.1.1 Analyses par régression linéaire

Les figures 5.6 et 5.7 montrent les analyses par régression linéaire pour les trois méthodes non supervisées (SIPS, Freesurfer et Wen) et leur comparaison avec les résultats obtenus par WHASA. La figure 5.6 montre l'analyse sur les volumes, la figure 5.7 celle sur les logarithmes des volumes. Pour cette dernière, les deux sujets outliers ayant un très faible volume de HSB n'ont pas été pris en compte.

L'analyse sur les volumes montre que les trois méthodes non supervisées sous-segmentent de manière importante les cas avec des HSB importantes. Cela se traduit par des coefficients directeurs de la droite de régression beaucoup plus faibles que pour WHASA ($y = 0,35x + 6,1$ pour SIPS, $y = 0,29x + 4,6$ pour Freesurfer, $y = 0,29x + 12,2$, $y = 0,95x + 5,3$ pour WHASA).

L'analyse sur les logarithmes des volumes permet de mieux mettre en évidence les différences de comportement entre les méthodes. Les méthodes SIPS et Freesurfer sont ainsi globalement parallèles à la droite identité, mais présentent une sous-segmentation systématique pour tous les volumes ($y = 0,90x - 0,4$ pour SIPS, $y = 0,86x - 0,5$ pour Freesurfer). De plus, dans le cas de la méthode SIPS, la variance des points augmente pour les faibles valeurs, alors qu'elle reste environ constante pour Freesurfer ($R^2 = 0,78$ pour SIPS, $R^2 = 0,90$ pour Freesurfer). Dans le cas de Wen, la droite de régression coupe la droite identité, indiquant une sursegmentation pour les faibles volumes et une sous-segmentation pour les volumes plus importants. La variance est à peu près constante ($R^2 = 0,87$). Par contraste, pour WHASA, la droite de régression est quasiment confondue avec la droite identité et la relation linéaire est très forte ($y = 1,08x - 0,3$, $R^2 = 0,97$).

5. COMPARAISON AUX AUTRES MÉTHODES DE LA LITTÉRATURE

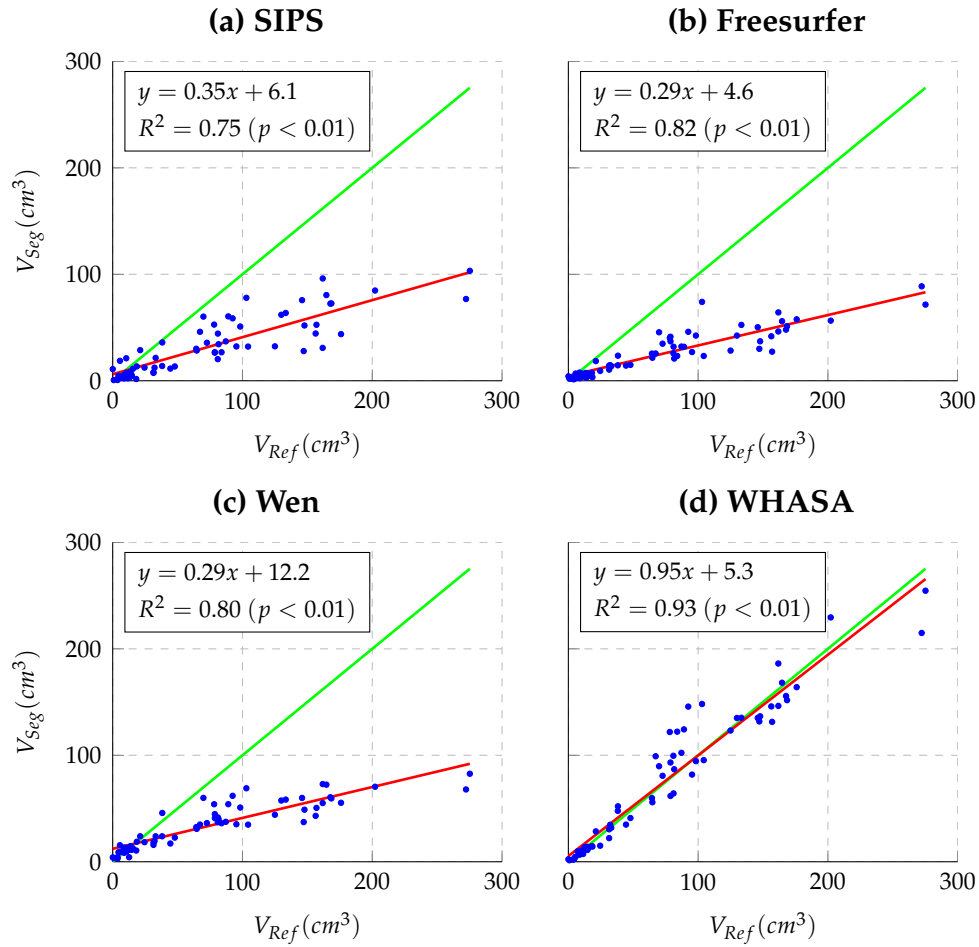


FIGURE 5.6 – Régression linéaire de V_{Seg} par rapport à V_{Ref} pour les méthodes non supervisées.

En rouge : droite de régression. En vert : droite identité.

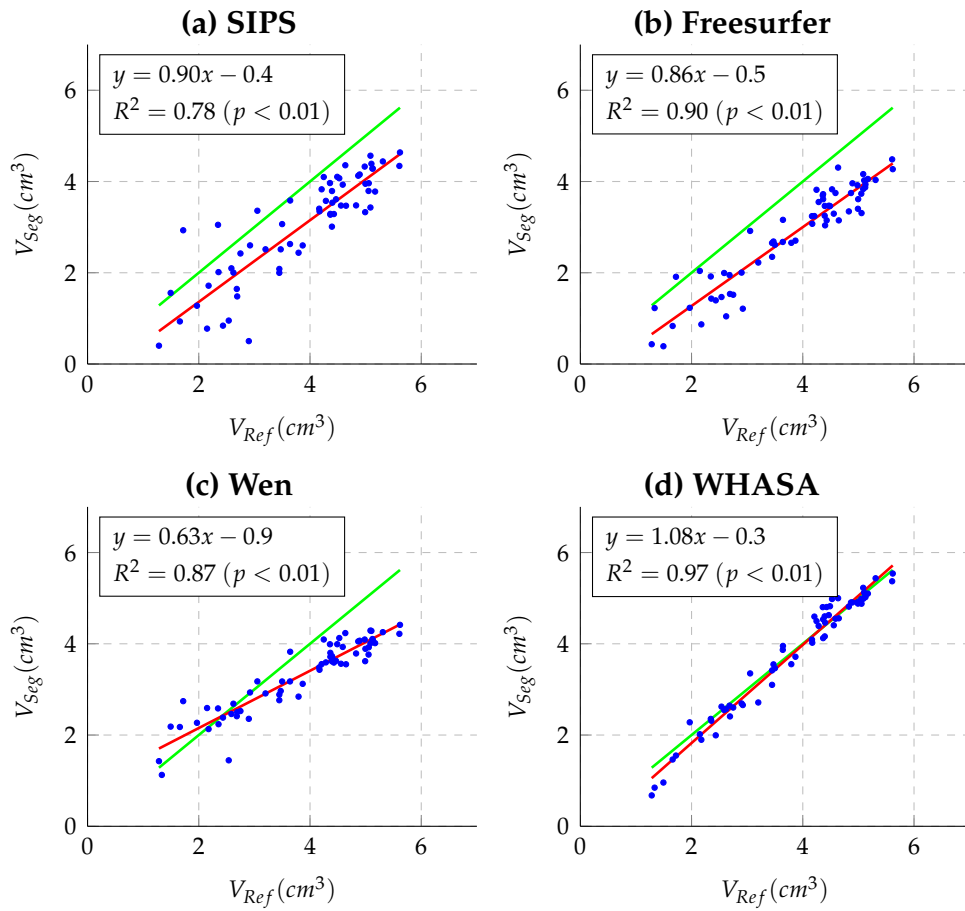


FIGURE 5.7 – Régression linéaire de $\log V_{Seg}$ par rapport à $\log V_{Ref}$ pour les méthodes non supervisées.

En rouge : droite de régression. En vert : droite identité.

5.4.1.2 Graphes de Bland-Altman

Les figures 5.8 et 5.9 présentent les graphes de Bland-Altman. La figure 5.8 analyse la différence entre les volumes en fonction de la moyenne arithmétique alors que la figure 5.9 analyse le ratio entre les volumes en fonction de la moyenne géométrique à partir de l'analyse des logarithmes des volumes (cf 4.3.3).

Pour les trois méthodes, on retrouve une forte sous-segmentation des HSB importantes. Le biais est ainsi fortement décroissant en fonction du volume (équation de droite pour l'estimation du biais : $y = -0,90x + 5,6$ pour SIPS, $y = -1,08x + 5,6$ pour Freesurfer et $y = -1,1x + 16,7$ pour Wen). Par comparaison, l'estimation du biais pour WHASA était indépendante du volume et proche de zéro. ($y = 1,2$ pour WHASA). L'analyse des *limits of agreement* est difficile étant donné les très grandes variations de volume. Globalement, il ne semble pas y avoir de très grandes différences entre les méthodes concernant ces *limits of agreements*.

L'analyse par logarithme et transformation réciproque permet de s'affranchir des effets de variation importante de volumes. Dans ce cas, l'estimation du biais devient relativement constante pour SIPS et Freesurfer. Le ratio moyen estimé est constant à 0,44 pour SIPS alors qu'il varie entre 0,41 et 0,32 pour Freesurfer. On retrouve à nouveau l'effet de sous-segmentation systématique pour ces deux méthodes. La variance est beaucoup plus importante pour les faibles volumes avec la méthode SIPS, avec des sous-estimations de l'ordre de 80% et des surestimations maximales de 100% voire 250% pour quelques sujets. Dans le cas de Freesurfer, en revanche, la variance est à peu près constante, avec des sous-estimations variant entre 80 et 50% environ et quasiment aucune surestimation. Pour Wen, le ratio moyen dépend plus fortement du volume, passant de 1,52 à 0,33, avec une variance importante pouvant aller jusqu'à une surestimation de plus de 250%. Pour la méthode WHASA, le ratio moyen demeure proche de 1, évoluant entre 0,75 et 1,13 selon le volume mesuré. De plus, de manière similaire à Freesurfer, les *limits of agreement* demeurent dans un intervalle plus restreint, entre des sous-estimations de l'ordre de 40% et des surestimations maximales de l'ordre de 60%.

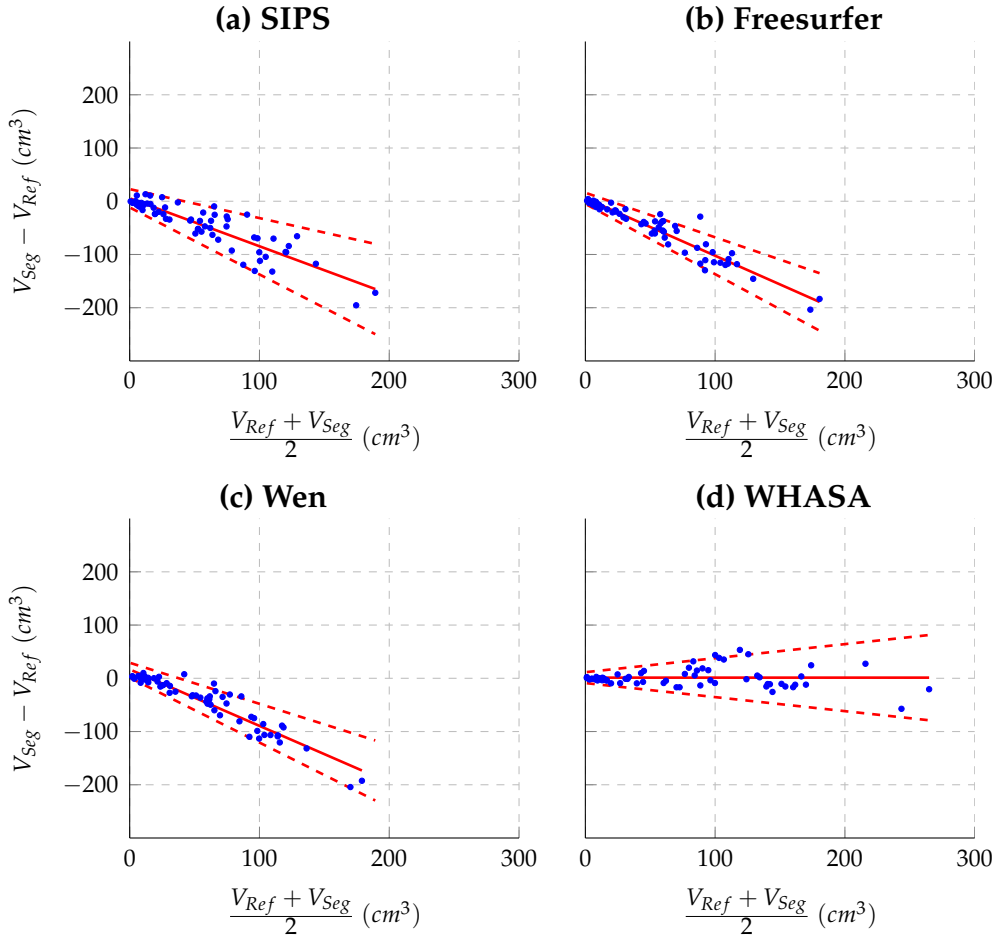


FIGURE 5.8 – Graphe de Bland-Altman entre V_{Seg} et V_{Ref} pour les méthodes non supervisées.

La ligne rouge indique l'estimation du biais, les lignes pointillées les *limits of agreement* comprenant 95% des points.

5.4.1.3 Validation spatiale

La figure 5.10 compare les distributions des SI pour les trois méthodes non supervisées et WHASA. La méthode SIPS présente les plus faibles performances avec de plus une variabilité élevée (valeur médiane du SI pour SIPS : 0,22 ; 1^{er} quartile (Q1) : 0,16 ; 3^{ème} quartile (Q3) : 0,38 ; écart interquartile (EIQ) : 0,22). Par ordre croissant de SI médian viennent ensuite les méthodes Freesurfer, Wen puis WHASA, avec un écart interquartile proche pour ces trois méthodes (Freesurfer : valeur médiane 0,40, [Q1 ; Q3] = [0,33 ; 0,47], EIQ = 0,15 ; Wen : valeur médiane 0,55, [Q1 ; Q3] = [0,47 ; 0,66], EIQ = 0,19 ; WHASA : valeur médiane 0,79, [Q1 ; Q3] = [0,70 ; 0,83], EIQ = 0,14). Les distributions des SI de SIPS,

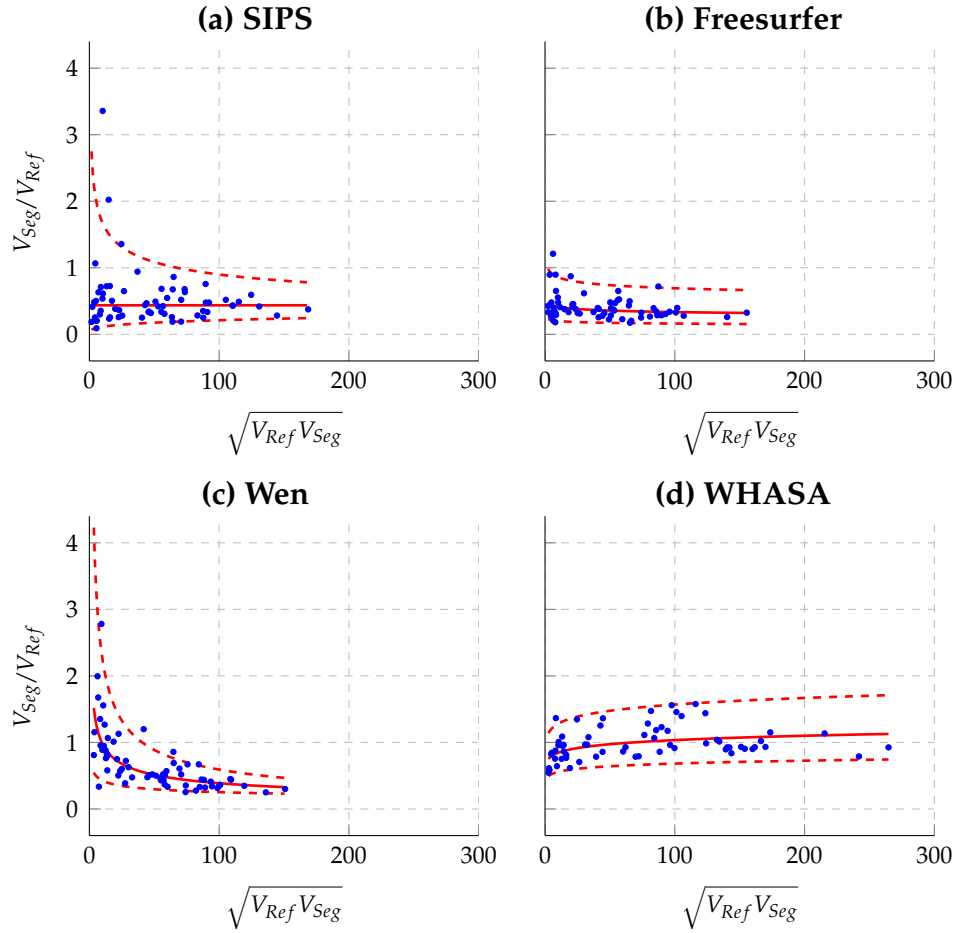


FIGURE 5.9 – Graphe de Bland-Altman obtenu après transformation réciproque de l’analyse des logarithmes pour les méthodes non supervisées. La ligne rouge continue indique l’estimation du biais, les lignes pointillées les *limits of agreement* comprenant 95% des points. Les deux sujets outliers sont exclus de l’analyse.

Freesurfer et Wen sont significativement différentes de celles de WHASA (test de Friedman, corrigé pour les comparaisons multiples par le HSD de Tukey, $p < 0,05$).

5.4.2 Comparaison avec les méthodes supervisées

Les méthodes d’apprentissage sont dépendantes de l’ensemble d’apprentissage fourni et du vecteur de caractéristiques utilisé. Afin de limiter le nombre de comparaisons, nous commençons par définir le vecteur de caractéristiques le plus performant pour le kNN et le SVM et comparons ensuite ses performances à celles de WHASA pour chacune des trois bases d’apprentissage.

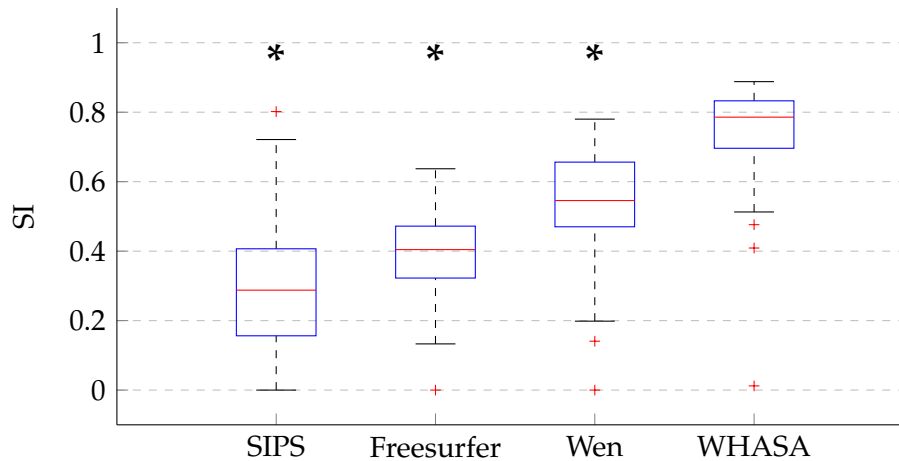


FIGURE 5.10 – Comparaison des distributions des SI des 67 sujets pour les différentes méthodes non supervisées.

* : différence statistiquement significative par rapport à la distribution des SI obtenue avec WHASA (test de Friedman, corrigé pour les comparaisons multiples par le HSD de Tukey, $p < 0,05$).

5.4.2.1 Effet de l'ensemble d'apprentissage et du vecteur de caractéristiques

Nous avons évalué les résultats obtenus par le kNN et le SVM en fonction du vecteur de caractéristiques utilisé (VC1, VC2, VC3 ou VC4) ainsi que de la base d'apprentissage/base de test (Base 1, Base 2 ou Base 3).

Un résumé des principaux indicateurs est présenté dans le tableau 5.1. Pour des raisons de clarté, nous n'y avons mentionné que le SI médian et l'écart interquartile, la valeur du coefficient de corrélation intraclasse, l'équation de la droite de régression de V_{Ref} par rapport à V_{Seg} et le coefficient de détermination associé, le biais et les *limits of agreement* de l'analyse de Bland-Altman. Pour ces deux derniers indicateurs, nous nous contentons d'estimer le biais et les *limits of agreement* en supposant que les différences sont normalement distribuées. Bien que cela ne soit pas le cas pour une grande partie des méthodes, le biais et les *limits of agreement* ainsi calculés sont d'interprétation beaucoup plus simple. Les *limits of agreement* sont surévaluées pour les faibles volumes et sous-évaluées pour les volumes importants mais permettent néanmoins de comparer les méthodes. Une analyse plus complète sera faite par la suite pour le vecteur de caractéristiques obtenant les meilleurs performances. La figure 5.11 résume graphiquement les distributions des SI pour les différents vec-

teurs de caractéristiques à travers les trois bases d'apprentissage pour la classification par kNN et par SVM.

En ce qui concerne le kNN, les vecteurs de caractéristiques VC1 et VC2, ne prenant pas l'information issue du voisinage du voxel, obtiennent des performances similaires pour tous les indicateurs. De même, les indicateurs sont similaires entre VC3 et VC4. Les vecteurs VC1 et VC2 obtiennent un meilleur accord spatial alors que les vecteurs VC3 et VC4 ont de meilleures performances au niveau volumique. La distribution de SI étant significativement meilleure pour VC1 et VC2 que pour VC3 et VC4 (test de Friedman, corrigé pour les comparaisons multiples par le HSD de Tukey, $p < 0,05$), nous avons choisi de retenir l'un des deux vecteurs VC1 ou VC2. Leur comportement étant très similaire, le choix entre ces deux vecteurs est assez arbitraire. VC1 obtient le meilleur SI médian et le plus faible écart interquartile sur les trois bases d'apprentissage, nous avons donc choisi de retenir VC1 pour la méthode de classification par kNN.

Pour le SVM, les trois vecteurs de caractéristiques VC1, VC2 et VC3 présentant des performances similaires, le vecteur VC4 affichant pour sa part des performances nettement inférieures. Le vecteur optimal dépend de la base d'apprentissage retenue bien que les différences soient assez faibles. Nous avons choisi de retenir VC3 pour la méthode de classification par SVM. Il présente en effet l'un des plus faibles écarts interquartile pour les SI et le meilleur coefficient de détermination R^2 sur les trois bases d'apprentissage, ce qui laisse supposer une performance plus homogène.

		SI médian (EIQ)	ICC	Regression			Bland-Altman		
				Pente	y-intercept	R ²	Biais	95% limits	
kNN	Test 1	VC1	0,77 (0,15)	0,90	0,79	29	0,83	11	[-44 66]
		VC2	0,77 (0,15)	0,90	0,77	27	0,83	8,3	[-47 64]
		VC3	0,70 (0,19)	0,92	0,85	22	0,86	8,9	[-41 59]
		VC4	0,70 (0,19)	0,92	0,84	21	0,86	7,3	[-43 58]
	Test 2	VC1	0,70 (0,34)	0,87	0,73	19	0,79	0,3	[-56 57]
		VC2	0,69 (0,36)	0,87	0,71	19	0,79	-0,2	[-58 58]
		VC3	0,64 (0,45)	0,91	0,93	5,6	0,82	1,3	[-53 55]
		VC4	0,64 (0,46)	0,91	0,92	5,8	0,82	0,7	[-54 55]
	Test 3	VC1	0,76 (0,18)	0,91	0,75	22	0,86	3,0	[-49 55]
		VC2	0,76 (0,20)	0,91	0,75	21	0,86	1,9	[-51 55]
		VC3	0,67 (0,25)	0,94	0,91	14	0,88	7,1	[-40 54]
		VC4	0,67 (0,25)	0,94	0,90	14	0,88	5,7	[-42 53]
SVM	Test 1	VC1	0,78 (0,20)	0,89	0,74	29	0,82	6,8	[-51 65]
		VC2	0,75 (0,19)	0,89	0,74	30	0,83	8,7	[-48 65]
		VC3	0,76 (0,14)	0,89	0,68	17	0,88	-10	[-65 44]
		VC4	0,74 (0,16)	0,88	0,67	21	0,87	-7,3	[-63 48]
	Test 2	VC1	0,73 (0,24)	0,88	0,77	24	0,80	9,0	[-47 65]
		VC2	0,72 (0,29)	0,88	0,75	21	0,81	4,3	[-51 59]
		VC3	0,72 (0,26)	0,92	0,78	9,4	0,87	-5,1	[-51 41]
		VC4	0,56 (0,46)	0,90	0,85	4,4	0,80	-5,7	[-61 49]
	Test 3	VC1	0,76 (0,19)	0,92	0,77	22	0,88	4,3	[-45 54]
		VC2	0,77 (0,18)	0,92	0,77	18	0,88	-0,2	[-49 49]
		VC3	0,76 (0,12)	0,94	0,80	15	0,90	-0,9	[-45 43]
		VC4	0,63 (0,27)	0,93	0,91	11	0,86	4,5	[-47 56]

TABLEAU 5.1 – Résultats de la classification par kNN et SVM pour les différents vecteurs de caractéristiques et les différentes bases d'apprentissage.

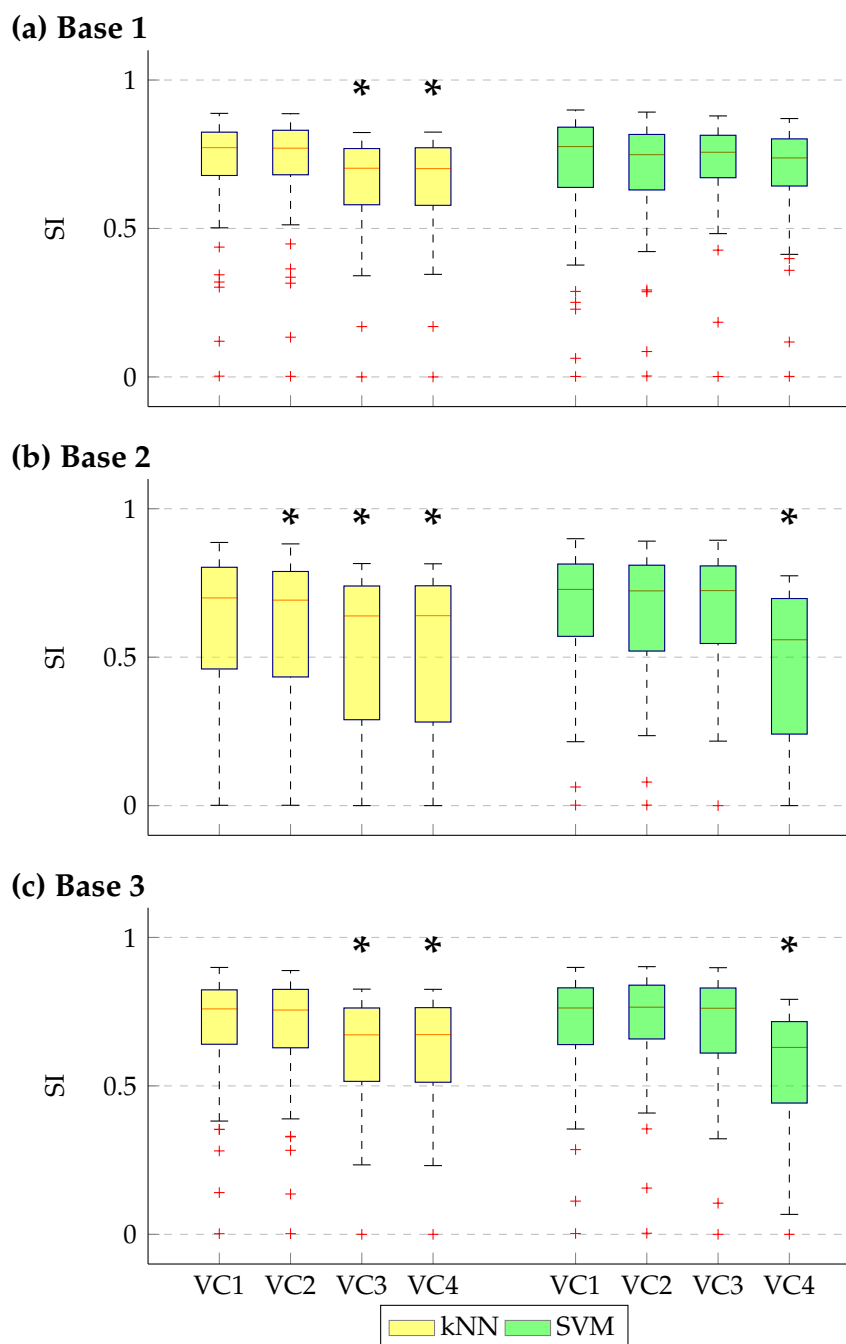


FIGURE 5.11 – Distribution des SI pour les différents vecteurs de caractéristiques (VC) et les différentes bases d'apprentissage.

* : différence statistiquement significative par rapport à la distribution des SI obtenue avec les autres VC pour la même méthode de classification (test de Friedman, corrigé pour les comparaisons multiples par le HSD de Tukey ($p < 0,05$)).

5.4.2.2 Comparaison

Nous avons donc retenu pour le kNN le vecteurs de caractéristiques VC1 et pour le SVM le vecteur de caractéristiques VC3. Pour chaque base d'apprentissage/base de test, nous évaluons les résultats obtenus en utilisant la fonction de classification définie à partir des 10 sujets de la base d'apprentissage et appliquée ensuite sur la base de test (57 patients restants). Pour permettre la comparaison, les résultats de WHASA sont alors restreints aux mêmes 57 sujets. De plus, les deux sujets outliers, lorsqu'ils appartenaient à la base de test, ont été exclus des analyses faisant intervenir les logarithmes.

Analyses par régression Les figures 5.12, 5.13 et 5.14 présentent les analyses par régression linéaire pour les bases d'apprentissage Base 1, Base 2 et Base 3.

Le comportement des trois méthodes est à peu près constant pour les trois bases d'apprentissage. Le kNN présente les moins bons accords volumiques : les faibles volumes sont la plupart du temps sursegmentés alors que les volumes importants sont sous-segmentés. La variance est augmentée pour les plus petits volumes et le kNN présente les coefficients de détermination R^2 les plus faibles (respectivement 0,83, 0,79 et 0,86 pour Base 1, Base 2 et Base 3 dans le cas des volumes ; 0,83, 0,72 et 0,86 pour Base 1, Base 2 et Base 3 dans le cas des logarithmes des volumes). Le SVM obtient de meilleurs résultats, la variance est à peu près constante en fonction du volume pour la Base 1, mais augmente pour les petits volumes dans le cas de la Base 2 et de la Base 3. Pour aucune de ces méthodes, on ne retrouve de sous-estimation ou surestimation systématique. Enfin, dans les trois bases d'apprentissage, l'équation de régression obtenue dans le cas de WHASA est la plus proche de la droite identité et le coefficient de détermination est toujours le plus élevé pour WHASA (respectivement 0,92, 0,93 et 0,93 pour Base 1, Base 2 et Base 3 dans le cas des volumes ; 0,97 pour les trois bases dans le cas des logarithmes).

5. COMPARAISON AUX AUTRES MÉTHODES DE LA LITTÉRATURE

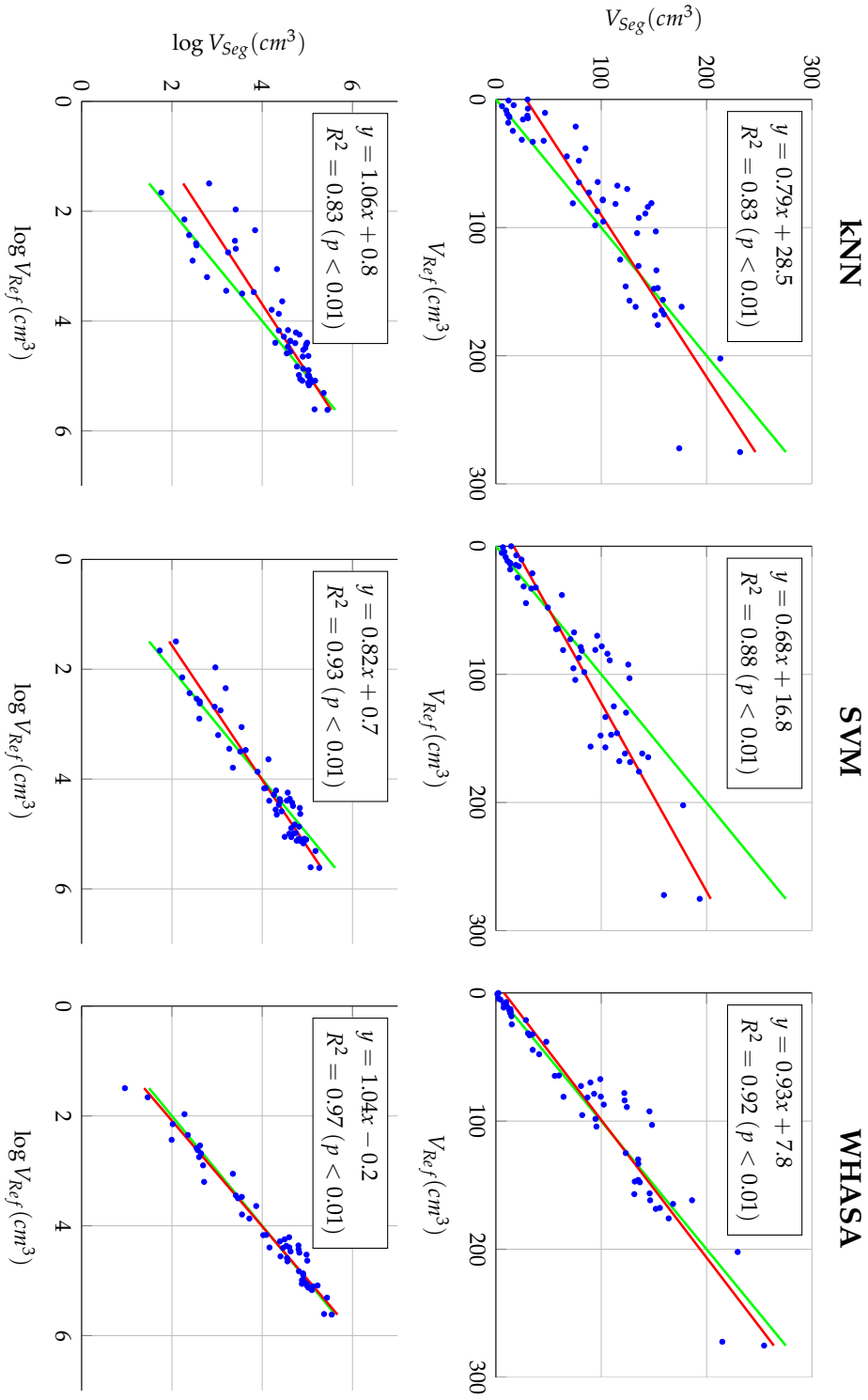


FIGURE 5.12 – Analyses par régression linéaire pour les résultats du kNN, SVM et WHASA sur la base de test Base 1.
Ligne du haut : Régression de V_{Seg} par rapport à V_{Ref} . **Ligne du bas :** Régression de $\log V_{Seg}$ par rapport à $\log V_{Ref}$. Pour tous les graphiques, la ligne rouge représente la droite de régression, la ligne verte la droite identité.

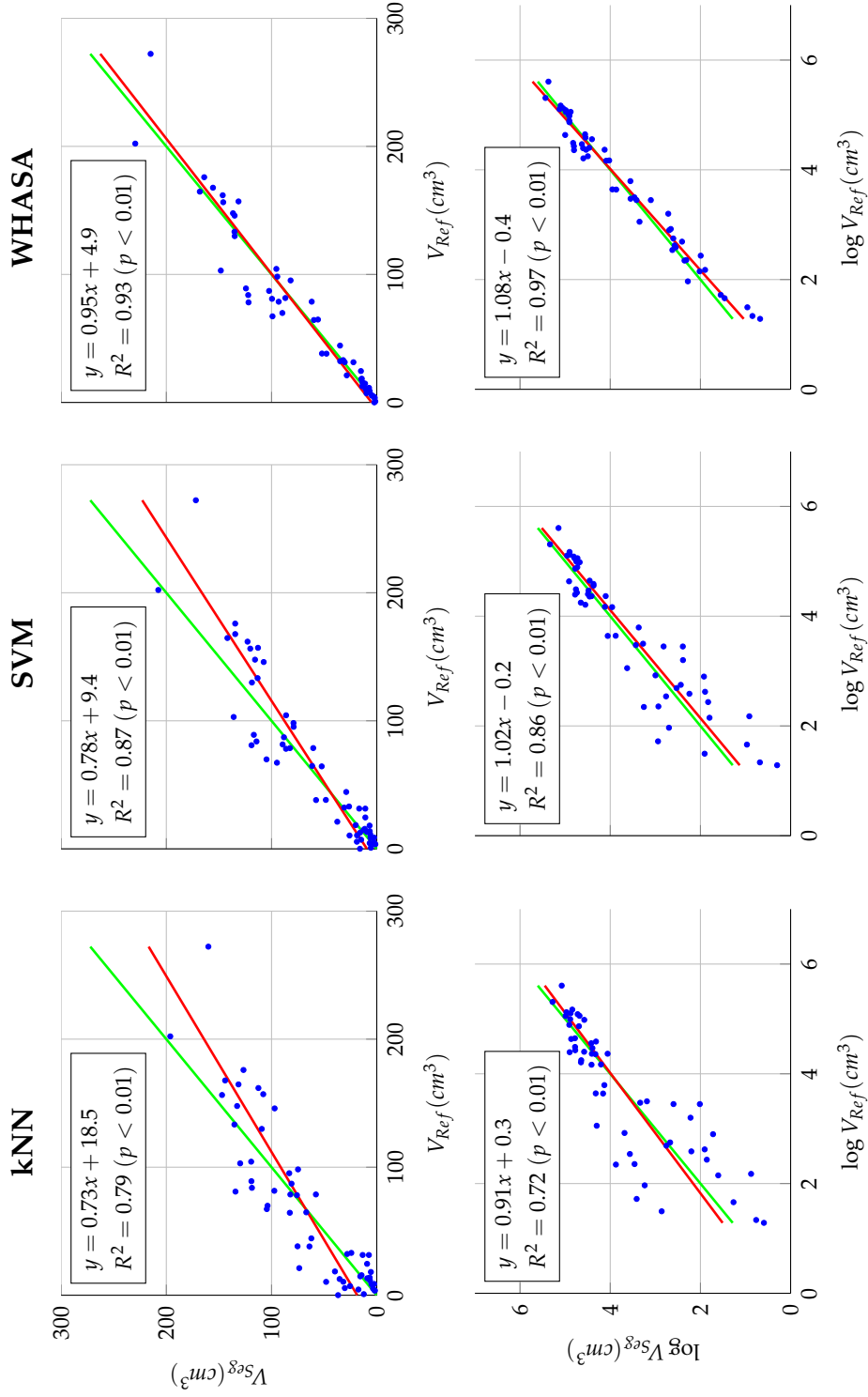


FIGURE 5.13 – Analyses par régression linéaire pour les résultats du kNN, SVM et WHASA sur la base de test Base 2.
Ligne du haut : Régression de V_{seg} par rapport à V_{ref} . **Ligne du bas :** Régression de $\log V_{seg}$ par rapport à $\log V_{ref}$. Pour tous les graphiques, la ligne rouge représente la droite de régression, la ligne verte la droite identité.

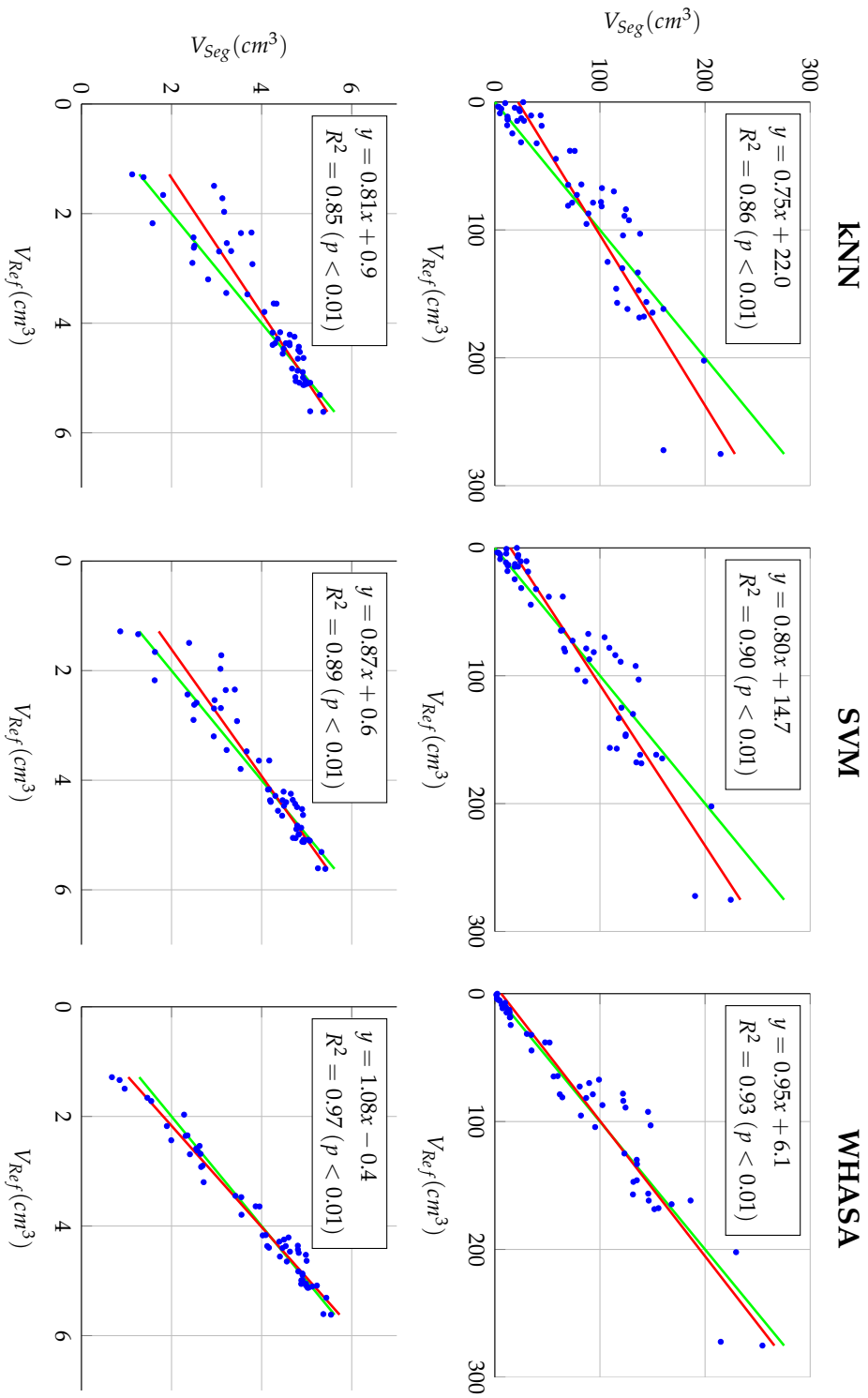


FIGURE 5.14 – Analyses par régression linéaire pour les résultats du kNN, SVM et WHASA sur la base de test Base 3.
Ligne du haut : Régression de V_{Seg} par rapport à V_{Ref} . **Ligne du bas** : Régression de $\log V_{Seg}$ par rapport à $\log V_{Ref}$. Pour tous les graphiques, la ligne rouge représente la droite de régression, la ligne verte la droite identité.

Graphes de Bland-Altman Les figures 5.15, 5.16 et 5.17 présentent les graphes de Bland-Altman pour les bases d'apprentissage Base 1, Base 2 et Base 3.

Pour le kNN et le SVM, le biais est décroissant en fonction du volume pour les trois bases (**kNN** : $y = -0,15x + 24,1$ pour Base 1, $y = -0,21x + 14,6$ pour Base 2 et $y = -0,22x + 20,2$ pour Base 3. **SVM** : $y = -0,33x + 15,7$ pour Base 1, $y = -0,18x + 6,5$ pour Base 2 et $y = -0,18x + 12,9$ pour Base 3). En revanche, pour WHASA le biais est constant à travers les trois bases ($y = 1,92$ pour Base 1, $y = 1,32$ pour Base 2 et $y = 1,82$ pour Base 3). Les *limits of agreement* sont également plus espacées pour le kNN, puis pour le SVM et sont le plus resserrées pour WHASA. Dans le cas de Base 1, en particulier, les *limits of agreement* pour le kNN sont particulièrement larges, y compris pour les faibles volumes. L'écart entre les *limits of agreement* est alors constant et vaut 105 cm^3 .

En analysant les ratios $\frac{V_{Seg}}{V_{Ref}}$, le ratio moyen est cette fois proche de 1 pour l'ensemble des méthodes et toutes les bases d'apprentissage (**kNN** : il varie entre 1,8 et 1 pour Base 1, est constant et égal à 1 pour Base 2, varie entre 1,7 et 1 pour Base 3. **SVM** : il varie entre 1,4 et 0,8 pour Base 1, est constant à 0,9 pour Base 2, est constant à 1,1 pour Base 3. **WHASA** : il varie entre 0,9 et 1,1 pour Base 1, varie entre 0,8 et 1,1 pour Base 2, varie entre 0,8 et 1,1 pour Base 3). En revanche, la variabilité des mesures en fonction du volume est importante pour le kNN pour les trois bases avec des surestimations assez fréquentes pour les faibles volumes. Ce comportement s'observe également de manière plus faible pour le SVM, alors que les performances obtenues par WHASA sont similaires pour les trois bases et comprises entre une sous-estimation de l'ordre de 40% et une surestimation maximale de l'ordre de 60%.

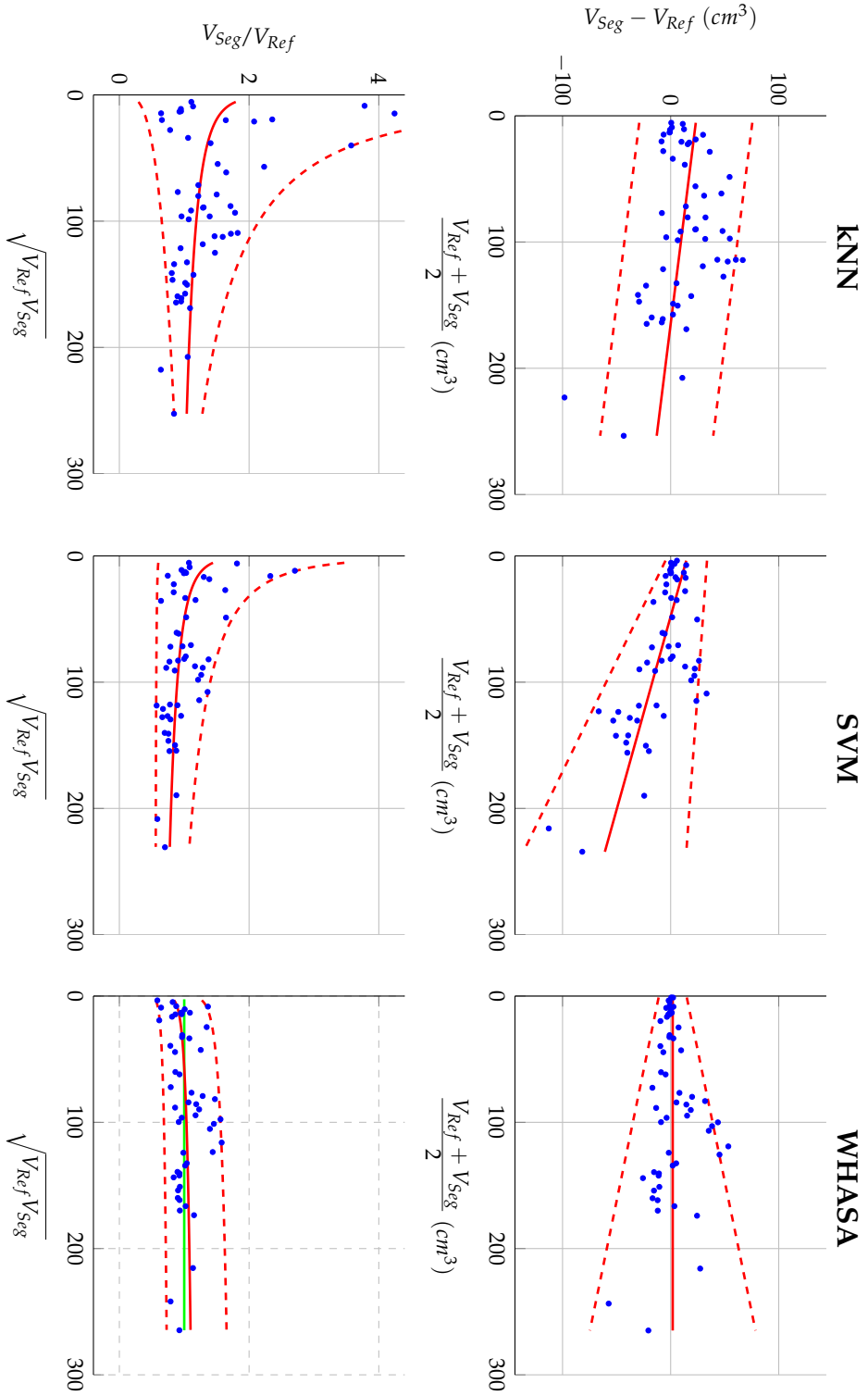


FIGURE 5.15 – Graphes de Bland-Altman pour les résultats du KNN, SVM et WHASA sur la base de test Base 1.
Ligne du haut : Graph de Bland-Altman entre V_{Seg} et V_{Ref} . **Ligne du bas :** graph de Bland-Altman obtenu après transformation réciproque de l'analyse des logarithmes. Pour tous les graphiques, la ligne rouge continue indique l'estimation du biais, les lignes pointillées les *limits of agreement* comprenant 95% des points.

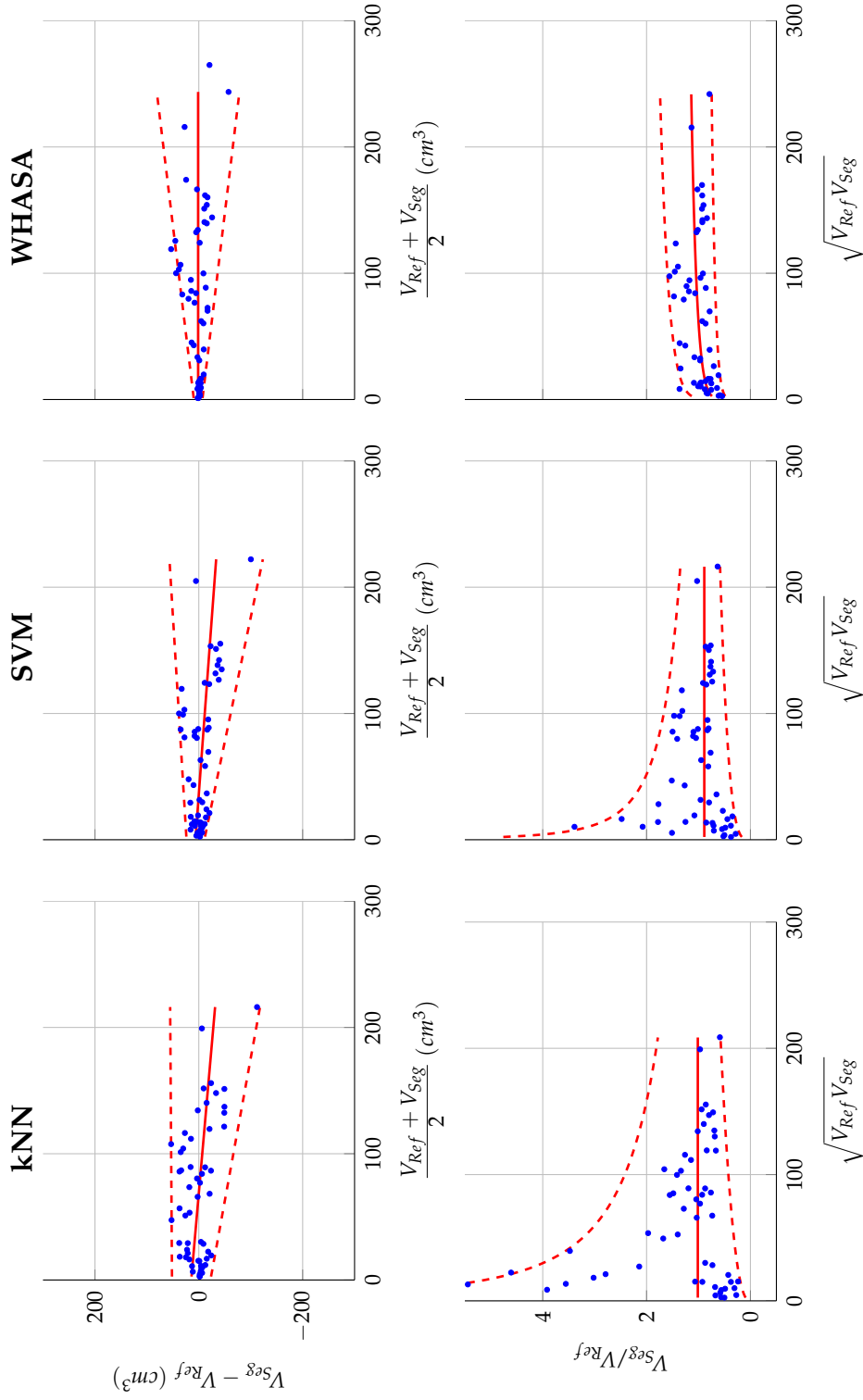


FIGURE 5.16 – Graphes de Bland-Altman pour les résultats du KNN, SVM et WHASA sur la base de test Base 2.
Ligne du haut : Graphe de Bland-Altman entre V_{Seg} et V_{Ref} . **Ligne du bas** : graphe de Bland-Altman obtenu après transformation réciproque de l'analyse des logarithmes. Pour tous les graphiques, la ligne rouge continue indique l'estimation du biais, les lignes pointillées les *limits of agreement* comprenant 95% des points.

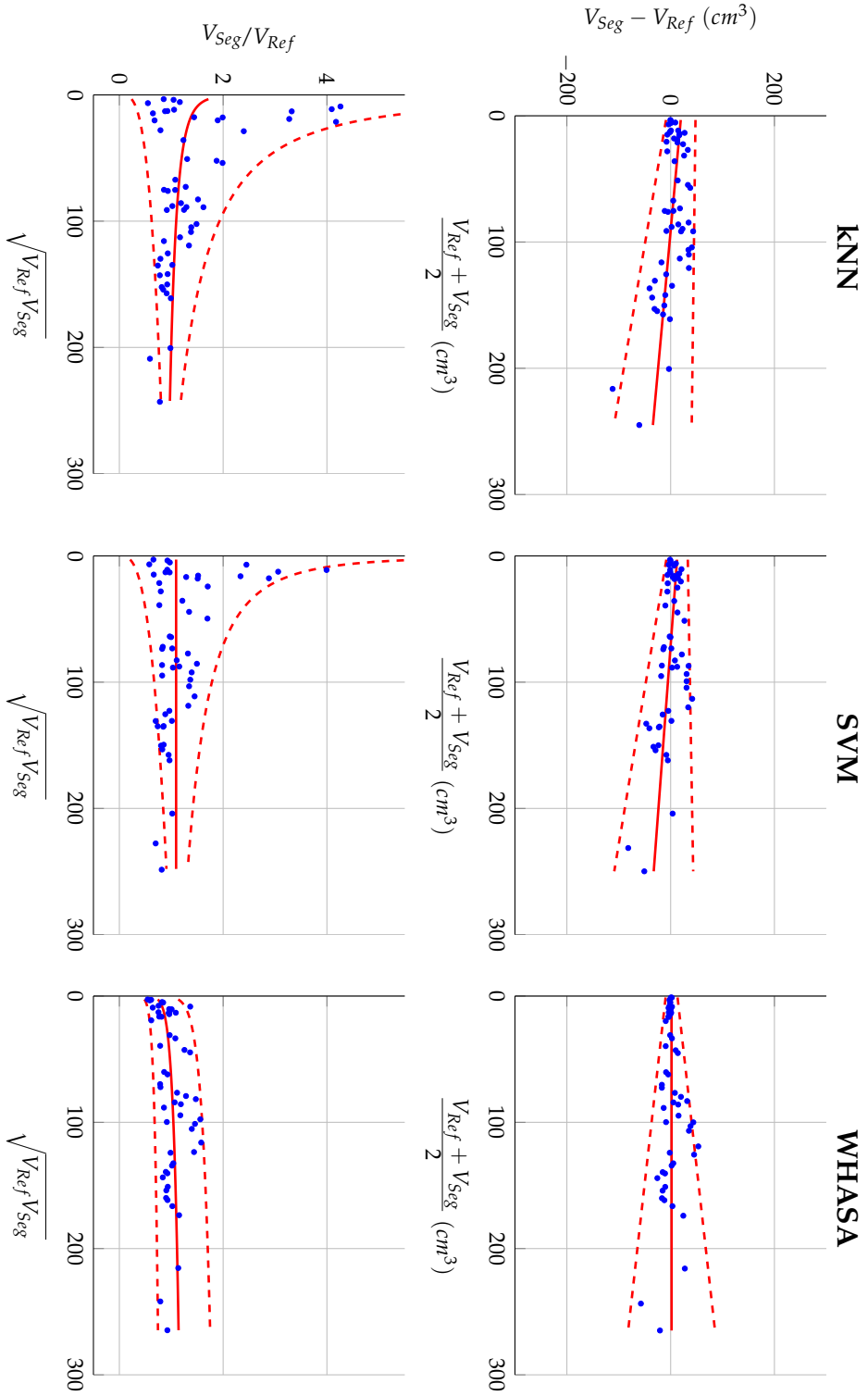


FIGURE 5.17 – Graphes de Bland-Altman pour les résultats du KNN, SVM et WHASA sur la base de test Base 3.
Ligne du haut : Graph de Bland-Altman entre V_{Seg} et V_{Ref} . **Ligne du bas :** graph de Bland-Altman obtenu après transformation réciproque de l'analyse des logarithmes. Pour tous les graphiques, la ligne rouge continue indique l'estimation du biais, les lignes pointillées les *limits of agreement* comprenant 95% des points.

Validation spatiale La figure 5.18 montre les distributions des SI pour le kNN, le SVM et WHASA en fonction des bases d'apprentissage Base 1, Base 2 et Base 3.

Pour la base 1, la valeur médiane est de 0,77 pour le kNN, 0,76 pour le SVM et de 0,78 pour WHASA. Les premiers et troisièmes quartiles sont eux égaux à $[0,68 \ 0,82]$ ($EIQ = 0,15$) pour le kNN, $[0,67 \ 0,81]$ ($EIQ = 0,14$) pour le SVM et $[0,72 \ 0,84]$ ($EIQ = 0,12$) pour WHASA. Pour la base 2, la valeur médiane est de 0,70 pour le kNN, 0,72 pour le SVM et de 0,78 pour WHASA. ($[Q1 \ Q3] = [0,46 \ 0,80]$ ($EIQ = 0,34$) pour le kNN, $[Q1 \ Q3] = [0,55 \ 0,81]$ ($EIQ = 0,26$) pour le SVM et $[0,70 \ 0,83]$ ($EIQ = 0,13$) pour WHASA. Pour la base 3, la valeur médiane est de 0,76 pour le kNN, 0,76 pour le SVM et de 0,78 pour WHASA. ($[Q1 \ Q3] = [0,64 \ 0,82]$ ($EIQ = 0,18$) pour le kNN, $[Q1 \ Q3] = [0,61 \ 0,83]$ ($EIQ = 0,22$) pour le SVM et $[Q1 \ Q3] = [0,70 \ 0,83]$ ($EIQ = 0,13$) pour WHASA. Dans les trois cas, la distribution des SI pour le kNN et le SVM est significativement différente de celle de WHASA (test de Friedman corrigé pour les comparaisons multiples par le HSD de Tukey, $p < 0,05$). Il est à noter que les performances du kNN et du SVM sont moins bonnes dans le cas de la Base 2, alors que les performances de WHASA sont similaires pour les trois bases d'apprentissage.

5.5 Discussion

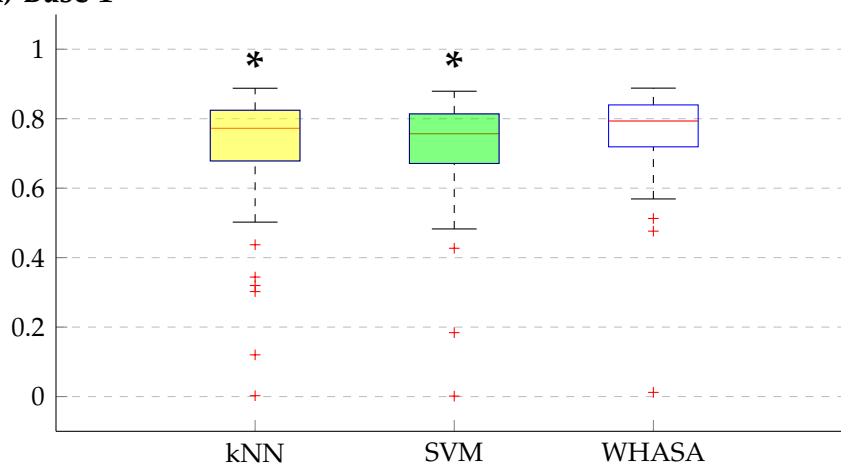
5.5.1 Comparaison avec les méthodes non supervisées

Les résultats présentés dans la section précédente montrent que les performances obtenues par WHASA sont supérieures à celles des trois méthodes non supervisées (SIPS, Freesurfer et Wen), tant pour l'accord volumique que pour l'accord spatial.

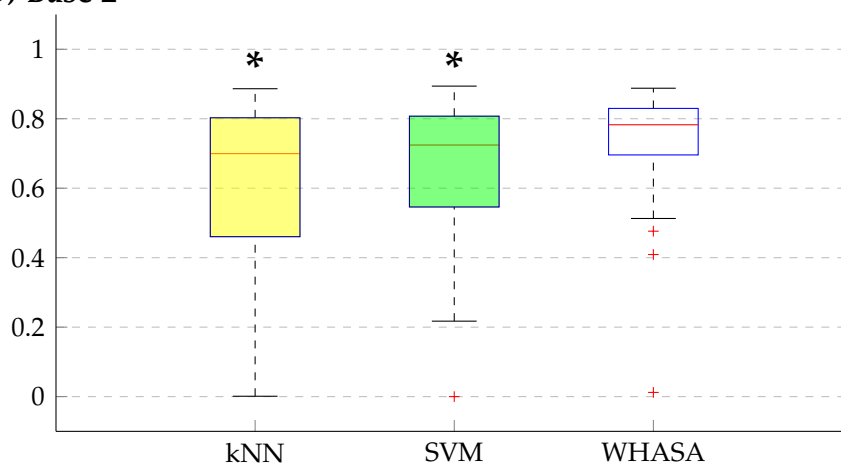
SIPS et Freesurfer sous-segmentent de manière quasi systématique les volumes de HSB et cette sous-estimation est de plus en plus importante lorsque le volume de HSB augmente. Pour SIPS, les résultats sont sensiblement moins bons que dans l'article de Gibson et al. [2010], qui indiquait un SI moyen de 0,65 pouvant aller jusqu'à 0,81 après post-traitement incluant des corrections manuelles. Dans notre cas, la valeur médiane du SI est de 0,22 (SI moyen de 0,28). Cette différence pourrait être expliquée par l'aspect multicentrique de notre étude, alors que Gibson et al. [2010] avaient

5. COMPARAISON AUX AUTRES MÉTHODES DE LA LITTÉRATURE

(a) Base 1



(b) Base 2



(c) Base 3

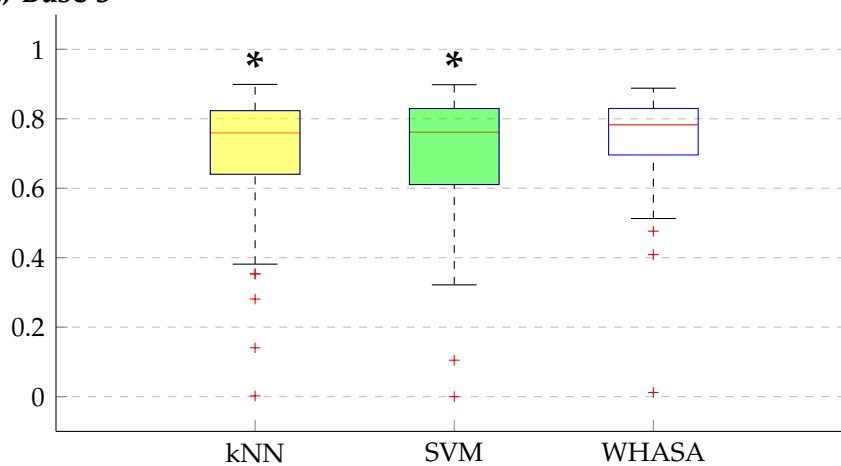


FIGURE 5.18 – Comparaison de la distribution des SI pour le KNN, le SVM et WHASA sur les différentes bases de test.

* : différence statistiquement significative par rapport à la distribution obtenue par WHASA (test de Friedman corrigé pour les comparaisons multiples par le HSD de Tukey, ($p < 0,05$)).

évalué leur méthode sur 18 individus, dont les acquisitions avaient toutes été faites dans le même centre. Concernant Freesurfer, la segmentation des HSB n'est pas l'objectif principal de ce logiciel, mais est prise en compte afin d'éviter de fausser les résultats de la segmentation des structures corticales et sous-corticales. La sous-segmentation systématique s'explique de plus facilement par le fait que Freesurfer ne s'appuie que sur les images T1, beaucoup moins sensibles pour la détection des HSB.

En ce qui concerne la méthode de seuillage proposée par Wen & Sachdev [2004], les résultats montrent une sursegmentation des faibles volumes et une sous-segmentation des volumes importants. Cet effet pourrait s'expliquer par la procédure de détermination du seuil. En effet, ce dernier dépend de l'écart-type de la distribution des intensités FLAIR du masque de substance blanche obtenu par les prétraitements. Dans le cas de faibles HSB, supposons que la distribution des intensités de la substance blanche suit une loi normale. Le fait de sélectionner tous les voxels situés à plus de 2,1 écart-types (valeur optimale pour le SI obtenue sur l'ensemble des 67 sujets) implique alors de sélectionner 2,5% des voxels de la distribution. Dans le cas d'importantes HSB, en revanche, il est probable qu'une partie importante d'entre eux seront segmentés dans la substance blanche lors des prétraitements. L'écart-type de la distribution sera donc augmenté et les HSB seront alors sous-évaluées. Il apparaît ainsi qu'un même seuil ne permettra jamais d'obtenir une segmentation correcte lorsque les charges lésionnelles varient beaucoup.

5.5.2 Comparaison avec les méthodes supervisées

Les performances des méthodes supervisées sont plus proches de celles de WHASA. Néanmoins, les performances de WHASA demeurent supérieures dans tous les cas évalués.

L'évaluation montre que les performances du kNN et du SVM sont meilleures que celles des méthodes non supervisées. De plus, le SVM semble légèrement plus performant que le kNN, obtenant des résultats plus homogènes à travers les trois bases d'apprentissage. Cette supériorité des méthodes supervisées et du SVM sur le kNN correspond aux résultats rapportés par Klöppel et al. [2011], seule étude ayant précédemment comparé plusieurs méthodes de segmentation des HSB sur un même jeu de données.

Nous avons également évalué l'impact du vecteur de caractéristiques et de la base d'apprentissage sur les résultats des méthodes supervisées. Pour le kNN, l'utilisation des informations issues du voisinage dans le vecteur de caractéristiques affaiblit les performances de manière significative. Les résultats sont ainsi moins bons avec VC3 et VC4, alors qu'ils sont très similaires entre VC1 et VC2. Ce n'est pas le cas pour le SVM, puisque les performances obtenues avec VC3 sont équivalentes et à celles obtenues avec VC1 et VC2. Cette différence peut probablement s'expliquer par la gestion des grandes dimensions. En effet, VC1 et VC2 sont de dimension 3 ou 6 alors que VC3 et VC4 sont de dimension 1401 ou 1404. Or, le SVM est adapté aux problèmes de grande dimension, pour lesquels l'existence d'un hyperplan séparant les données est plus probable. Au contraire, les performances du kNN diminuent avec la dimension de l'espace. En ce qui concerne les performances plus faibles du SVM entre VC3 et VC4, il est probable que le fait de simplement concaténer les informations spatiales avec les informations d'intensité ne soit pas optimal.

5.5.3 Temps de calcul

Le tableau 5.2 résume les temps de calcul nécessaires pour les différentes méthodes. Il ne s'agit ici que d'estimations, les calculs ayant été lancés sur différentes machines de manière simultanée qui ne présentaient pas toutes les mêmes caractéristiques techniques. De plus, des gains substantiels de temps auraient probablement pu être obtenus en optimisant la programmation de chaque méthode. Cela n'était néanmoins pas le but de cette thèse et ce tableau est principalement à titre indicatif.

		Temps d'appren- tissage	Temps de classi- fication par sujet
Non supervisées	SIPS	Aucun	5 minutes
	Freesurfer	Aucun	24 heures
	Wen	Aucun	< 1 minute
	WHASA	Aucun	30 minutes
Supervisées	kNN	VC1/VC2	Minutes
		VC3/VC4	Minutes
	SVM	VC1/VC2	Minutes
		VC3/VC4	15-20 heures

TABLEAU 5.2 – Estimation des temps de calcul pour les différentes méthodes.

5.6 Conclusion

Ce chapitre a présenté les résultats obtenus par différentes méthodes de segmentation des HSB proposées dans la littérature, qui ont été comparés aux résultats obtenus par la méthode que nous proposons, WHASA. Nous avons essayé d'évaluer un large spectre de méthodes de segmentation, comprenant aussi bien des méthodes non supervisées que des méthodes supervisées. Ces méthodes ont été comparées en utilisant des prétraitements identiques sur un même jeu de données multicentriques. Les segmentations obtenues ont été comparées avec une segmentation manuelle de référence. Cette comparaison a évalué l'accord volumique et l'accord spatial.

Les méthodes non supervisées évaluées (SIPS, Freesurfer et Wen) obtiennent des performances décevantes et ne permettent pas une segmentation automatique correcte des HSB sur une population présentant des HSB variées. Les résultats obtenus par WHASA sont largement supérieurs à ceux de ces méthodes tant au niveau volumique que spatial.

Les méthodes supervisées (kNN, SVM) présentent des performances correctes, mais néanmoins inférieures à celles de WHASA. Parmi ces méthodes, le SVM obtient des résultats légèrement meilleurs que le kNN. Le SVM semble en particulier plus robuste à la base d'apprentissage et au vecteur de caractéristique utilisé.

Chapitre 6

Application sur d'autres jeux de données

Le chapitre précédent a montré que WHASA obtenait des performances supérieures aux autres méthodes proposées dans la littérature. L'une des forces de WHASA est d'être robuste aux conditions d'acquisition et de ne pas dépendre d'un ensemble d'apprentissage. Il peut donc être utilisé directement sur de nouveaux jeux de données et ne nécessite pas l'optimisation de paramètres pour chaque jeu de données. Dans ce chapitre, nous présentons une validation de WHASA sur de nouvelles données. Dans une première étude, en collaboration avec le Centre Hospitalier Régional Universitaire (CHRU) de Lille, nous comparons les résultats automatiques avec l'échelle visuelle de Fazekas. Nous présentons ensuite l'utilisation de WHASA dans le cadre d'une initiative de mise à disposition à grande échelle d'outils d'analyses en neuroimagerie, le Centre pour l'Acquisition et le Traitement des Images (CATI). En particulier, la méthode WHASA a été appliquée sur les 87 sujets de la phase pilote et sur 47 sujets de la phase de déploiement d'une importante étude de cohorte multicentrique, dénommée MEMENTO, qui est gérée par le CATI.

6.1 Collaboration avec le CHRU de Lille

Les données présentées dans cette section ont été analysées dans le cadre d'une collaboration avec Christine Delmaire, Stéphanie Bombois et Xavier Delbheuck du CHRU de Lille.

6.1.1 Description des données

Les patients considérés ici font partie d'une cohorte de patients MCI (*Mild Cognitive Impairment*) suivis de manière prospective dans l'étude intitulée "MCI : facteurs diagnostiques et pronostiques" ayant débuté en 2005 au Centre Mémoire de Ressources et de Recherche (CMRR) de Lille.

Les patients devaient être âgés d'au moins 50 ans, parler le français et, de préférence, être accompagnés d'un informant pouvant renseigner de manière fiable sur l'existence de signes cognitifs ou comportementaux dont les patients pourraient ne pas avoir conscience. Leurs premiers symptômes devaient dater de plus de six mois et de moins de cinq ans. Enfin, ils devaient remplir les critères de MCI définis par la Mayo Clinic [Petersen, 2004], c'est-à-dire (i) une plainte cognitive, de préférence corroborée par un informant, (ii) un fonctionnement cognitif global normal, (iii) au moins une des fonctions cognitives déficitaires à plus d'1,5 écart-type en comparaison d'une population témoin de même âge et de même niveau culturel, (iv) une préservation de l'autonomie dans les actes de la vie quotidienne et (v) une absence de démence selon les critères du DSM-IV. Les patients ont tous signé au préalable un consentement éclairé. Les critères de non inclusion étaient l'existence ou la présence d'un antécédent de maladie psychiatrique, l'illettrisme, la non-francophonie, une contre-indication à l'IRM, une épilepsie ou un antécédent de traumatisme crânien sévère.

L'étude comprenait deux ans de suivi des patients pendant lesquels était effectué, de façon standardisée, le recueil :

- de données cliniques et neuropsychologiques tous les six mois,
- de données d'imagerie (morphologique et fonctionnelle) et biologiques (tests sanguins et analyse du liquide cébrospinal) à la date d'inclusion et après deux ans de suivi.

Les données que nous avons utilisées pour l'évaluation de WHASA sont les données d'imagerie acquises lors de la visite d'inclusion. L'examen IRM est réalisé sur une machine 1,5 Tesla, Philips Achieva. Le patient est placé en position standard d'exploration cérébrale après vérification de l'absence de contre-indication à l'examen. Les paramètres des séquences utilisées par WHASA sont les suivants :

- image T1 : TR = 8,02ms, TE = 3,8 ms, angle de bascule = 8°, résolution dans le plan = $1 \times 1 \text{ mm}^2$, épaisseur de coupe = 1,3 mm.
- image FLAIR : TR = 11 000 ms, TE = 140 ms, résolution dans le plan

= $0,938 \times 0,938$, épaisseur de coupe = 5 mm.

6.1.2 Segmentation des HSB par WHASA

Les images T1 et FLAIR étaient disponibles pour 59 patients. La segmentation des HSB de ces 59 patients a été obtenue en suivant la procédure WHASA décrite au chapitre 3. Aucun paramètre n'a été modifié pour l'application à ces nouvelles données.

6.1.3 Corrélation avec les échelles visuelles

La sévérité des HSB a également été cotée visuellement selon l'échelle de Fazekas & Schmidt (voir chapitre 2). La cotation s'effectue en deux parties :

- un score entre zéro et trois est donné pour les HSB périventriculaires, noté Score PVH ;
- un score entre zéro et trois est donné pour les HSB situées dans la substance blanche profonde, noté Score DWMH.

Score PVH La figure 6.1 montre la répartition des volumes obtenus par WHASA en fonction du score visuel PVH. La population n'étant pas répartie équitablement entre les groupes, nous avons montré le volume pour chaque sujet et un diagramme de type "boîte à moustaches".

Les volumes médians, les premiers et troisièmes quartiles des différentes catégories sont les suivants :

- **Score PVH de 0.** Volume médian de $1,2 \text{ cm}^3$, $EIQ = [0,6 ; 1,8]$ ($N = 16$).
- **Score PVH de 1.** Volume médian de $4,3 \text{ cm}^3$, $EIQ = [2,4 ; 6,7]$ ($N = 33$).
- **Score PVH de 2.** Volume médian de $10,9 \text{ cm}^3$, $EIQ = [9,5 ; 27,1]$ ($N = 7$).
- **Score PVH de 3.** Volume médian de $53,7 \text{ cm}^3$, $EIQ = [30,7 ; 75,1]$ ($N = 3$).

Une analyse de variance de Kruskal-Wallis confirme une différence statistiquement significative entre les groupes ($p < 0,01$). Les tests post-hoc font apparaître que les sujets ayant un Score PVH de 0 sont significativement différents des autres groupes ($p < 0,05$, corrigé pour les comparaisons multiples), de même pour les sujets ayant un Score PVH

de 1. La différence entre les sujets ayant un Score PVH de 2 et ceux ayant un Score PVH de 3 n'est par contre pas significative, très probablement à cause du faible nombre de sujets dans le dernier groupe.

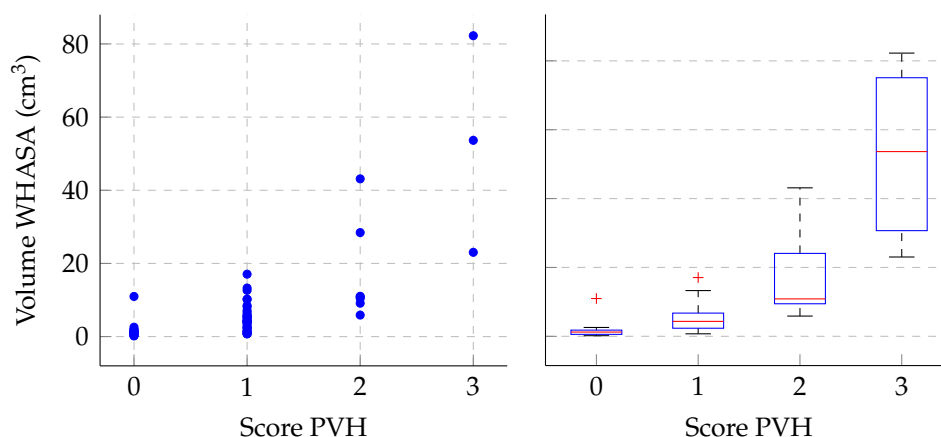


FIGURE 6.1 – Volume WHASA obtenu automatiquement en fonction du Score PVH attribué pour les 59 sujets de l'étude.

Gauche : Chaque sujet est représenté par un point. **Droite :** Diagramme en "boîte à moustaches".

Un patient se détache particulièrement avec un volume supérieur à 10 cm^3 pour un Score PVH de 0. Des coupes illustratives de ce patient sont présentées dans la figure 6.2, et on peut remarquer que ce patient présente un important infarctus cortical, ce qui explique la discordance entre la mesure automatique et l'échelle visuelle.

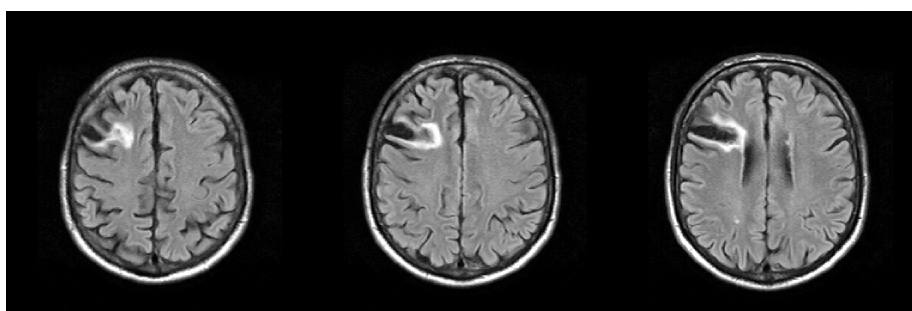


FIGURE 6.2 – Coupes illustratives du sujet pour lequel le volume WHASA et le Score PVH sont discordants.

Score DWMH La figure 6.3 présente la répartition des volumes obtenus par WHASA en fonction du score visuel DWMH. Comme précédemment, nous avons montré le volume pour chaque sujet et un diagramme de type "boîte à moustaches".

Les volumes médians, premiers et troisièmes quartiles des différentes catégories sont les suivants :

- **Score DWMH de 0.** Volume médian de $1,3 \text{ cm}^3$, $EIQ = [0,7 ; 1,5]$ ($N = 13$).
- **Score DWMH de 1.** Volume médian de $4,4 \text{ cm}^3$, $EIQ = [2,3 ; 7,9]$ ($N = 39$).
- **Score DWMH de 2.** Volume médian de $19,7 \text{ cm}^3$, $EIQ = [10,7 ; 35,8]$ ($N = 4$).
- **Score DWMH de 3.** Volume médian de $53,7 \text{ cm}^3$, $EIQ = [30,7 ; 75,1]$ ($N = 3$).

Une analyse de variance de Kruskal-Wallis confirme une différence statistiquement significative entre les groupes ($p < 0,01$). Les tests post-hoc font apparaître que les sujets ayant un Score DWMH de 0 sont significativement différents des autres groupes ($p < 0,05$, corrigé pour les comparaisons multiples). La différence entre les autres groupes n'est par contre pas significative pour un risque de 5% (Si l'on accepte un risque de 10%, le groupe Score DWMH=1 est alors significativement différent des autres groupes. La différence entre les groupes Score DWMH=2 et Score DWMH=3 demeure, elle, non significative, très probablement à cause du faible nombre de sujets dans le dernier groupe).

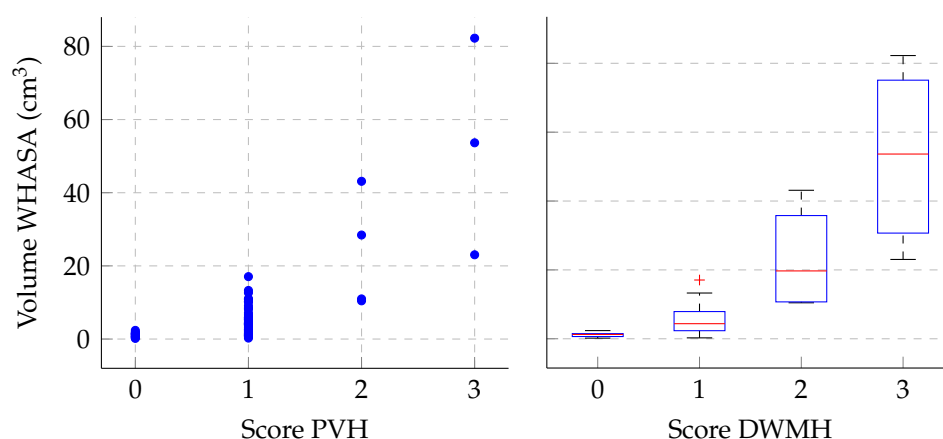


FIGURE 6.3 – Volume WHASA obtenu automatiquement en fonction du Score DWMH attribué pour les 59 sujets de l'étude.

Gauche : Chaque sujet est représenté par un point. **Droite** : Diagramme en "boîte à moustaches".

6.1.4 Relation avec l'âge

Une autre manière indirecte de vérifier le bon comportement de WHASA est de regarder la relation entre le volume obtenu par WHASA et des variables dont la relation avec les HSB est connue. En particulier, l'âge a été démontré comme l'un des principaux facteurs de risque des HSB.

La figure 6.4 montre la répartition des âges des sujets en fonction de leur score visuel PVH ou DWMH. L'utilisation de l'échelle visuelle ne fait pas apparaître de relation claire avec l'âge. Ainsi, pour les sujets ayant un Score PVH de 0, 1, 2 ou 3, l'âge médian est respectivement de 66, 69, 72 et 69 ans (les effectifs sont respectivement de 16, 33, 7 et 3 sujets). Pour les scores DWMH de 0, 1, 2 ou 3, les âges médians sont de 60 ans, 70 ans, 70 ans et demi et 69 ans (effectifs respectifs de 13, 39, 4 et 3 sujets). Une analyse de variance de Kruskal-Wallis ne montre aucune différence

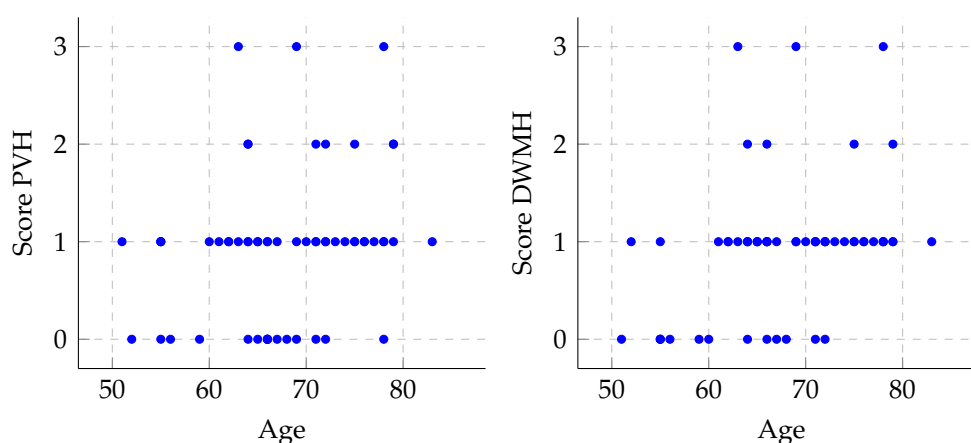


FIGURE 6.4 – Score visuel des sujets en fonction de leur âge.
Gauche : Score PVH. Droite : Score DWMH.

statistiquement significative d'âge médian entre les groupes définis par le Score PVH ($p = 0,20$). Une différence statistique significative entre tous les groupes est par contre trouvée en utilisant Score DWMH. Les tests post-hoc montrent cependant que seuls les âges médians des groupes Score DWMH=0 et Score DWMH=1 sont différents ($p < 0,05$, corrigé pour les comparaisons multiples), les différences entre tous les autres groupes n'étant pas significatives.

La figure 6.5 présente la répartition des volumes de HSB obtenus par WHASA en fonction de l'âge des sujets. Ce graphique montre une relation faible et non linéaire entre le volume des HSB et l'âge, avec un sujet très

différent des autres, présentant un très fort volume de HSB. Le coefficient de détermination R^2 entre le logarithme des volumes obtenus et l'âge est de 0,09, cette corrélation étant statistiquement significative ($p = 0,02$). Si l'on exclut le sujet outlier de l'analyse, le coefficient de détermination devient 0,12 ($p = 0,009$).

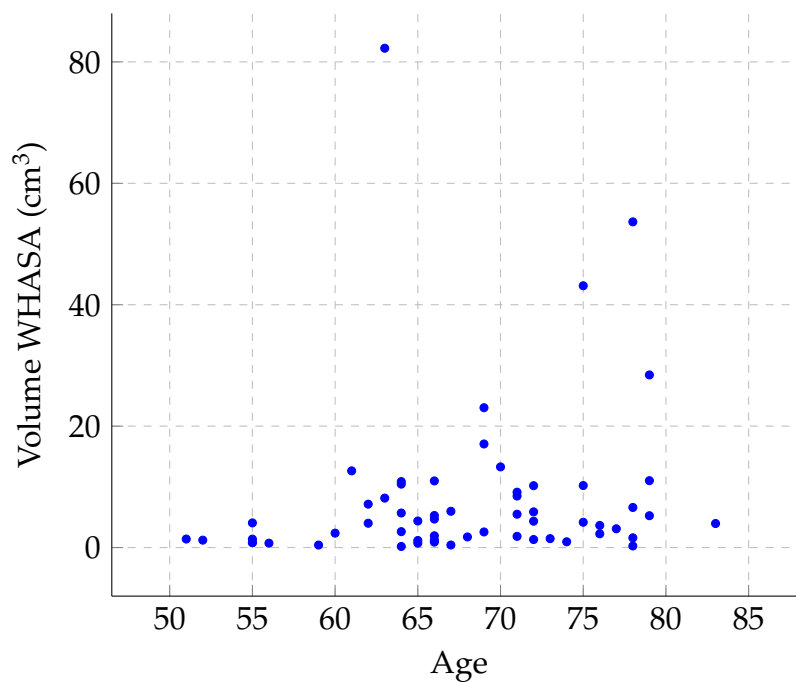


FIGURE 6.5 – Volume des HSB obtenu par WHASA en fonction de l'âge des sujets.

Le sujet outlier a un volume important de HSB pour son âge par rapport au reste de la population (volume de 82 cm³ à 63 ans). La figure 6.6 montre quelques coupes illustratives de ce sujet. La segmentation automatique est tout à fait correcte et ce patient présente bien un volume de HSB très important. L'analyse volumétrique permet d'obtenir une meilleure discrimination de la charge lésionnelle en fonction de l'âge, qui n'apparaît pas de façon évidente avec les échelles visuelles.

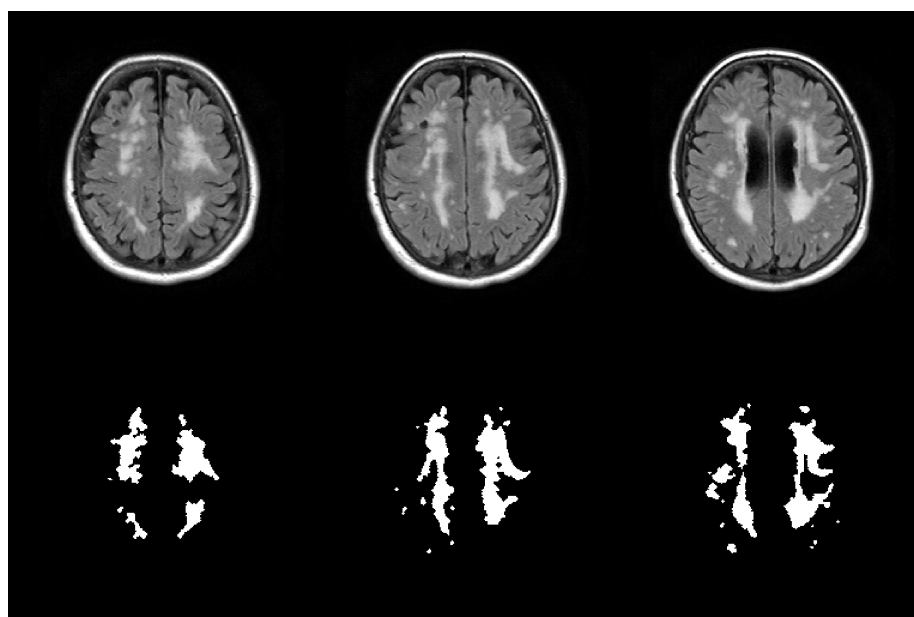


FIGURE 6.6 – Coupes illustratives du sujet présentant un volume anormalement élevé pour son âge.

Pour ce sujet, le volume des HSB mesuré automatiquement est de 82 cm^3 , le patient est âgé de 63 ans.

6.2 Mise en place au sein du Centre pour l'Acquisition et le Traitement des Images (CATI)

Les données présentées dans cette section sont issues de la cohorte MEMENTO, détaillée ci-dessous.

6.2.1 Cohorte MEMENTO

La cohorte MEMENTO (Déterminants et évolution de la Maladie d'Alzheimer débutante et des déficits cognitifs) est issue de la mesure 29 du plan Alzheimer, qui prévoit la constitution d'une cohorte permanente de patients français ayant des signes cliniques pouvant évoquer une forme débutante de la Maladie d'Alzheimer (MA) ou une maladie apparentée. Les participants sont recrutés au sein des Centres Mémoire de Ressources et de Recherche (CMRR) volontaires et disposant de l'infrastructure nécessaire en termes d'imagerie et de constitution de biothèque. Selon les termes de l'appel d'offre de la Fondation Plan Alzheimer, cette cohorte doit à la fois contribuer à la compréhension de la maladie, de son évolu-

6.2. Mise en place au sein du Centre pour l'Acquisition et le Traitement des Images (CATI)

tion, de ses conséquences sur la vie des patients, permettre d'accélérer le développement de nouveaux traitements et être une plateforme nationale pour le développement de projets de recherche sur un groupe de patients pour lesquels un recueil homogène et standardisé de données cliniques, sociodémographiques, biologiques, d'imagerie par résonance magnétique et de tomographie par émission de positons aura été réalisé.

Cette cohorte multicentrique nationale prospective inclura 2300 patients suivis pendant une durée minimale de 5 ans. Les critères d'inclusion sont les suivants :

- âge ≥ 18 ans ;
- avoir
 - des performances inférieures à la moyenne moins un écart-type (par comparaison à des normes attendues par âge et par sexe) dans au moins un domaine cognitif (évalué à partir de tests neuropsychologiques évaluant la mémoire, le langage, les praxies et les fonctions exécutives). Cette déviation à la norme doit être identifiée à partir de tests datant de moins de six mois ;
 - OU avoir des plaintes cognitives, quelle qu'en soit l'ancienneté, pour les plus de 60 ans avec des performances supérieures à la moyenne moins un écart-type dans tous les domaines cognitifs explorés par la batterie neuropsychologique (taille maximale de la strate "plainte isolée" = 300) ;
- Clinical Dementia Rating Scale $\leq 0,5$ (patient non dément) ;
- Consentement libre, éclairé et écrit, signé par le patient et l'investigateur (au plus tard le jour de la pré-inclusion et avant tout examen nécessité par la cohorte) ;
- personne affiliée à ou bénéficiaire d'un régime de sécurité sociale tel que défini dans la loi.

Les critères de non inclusion comprennent les formes génétiques de MA (mutation génétique connue) ; l'existence d'une maladie neurologique, telle que maladie de Parkinson traitée, épilepsie traitée, maladie de Huntington, tumeur cérébrale, hématome sous-dural, sclérose en plaques, paralysie supranucléaire progressive, antécédent de traumatisme crânien suivi de déficits neurologiques persistants, existence d'anomalies cérébrales structurelles ; l'existence d'un accident vasculaire cérébral ayant nécessité une hospitalisation survenue dans les trois mois précédents ; des antécédents de schizophrénie (critères DSM-IV).

6.2.2 Centre pour l'Acquisition et le Traitement des Images (CATI)

La partie sur l'imagerie des participants de cette cohorte est gérée par le Centre pour l'Acquisition et le Traitement des images (CATI). Il s'agit d'une plateforme nationale pour la neuroimagerie, soutenue par le plan Alzheimer. Le but du CATI est de promouvoir la recherche sur les troubles neurologiques en s'appuyant sur des techniques d'imagerie avancées. La plateforme est mise en place par un centre expert composé d'équipes de Neurospin (Jean-François Mangin, directeur du CATI) et de plusieurs équipes localisées sur le groupe Hospitalier Pitié-Salpêtrière (Stéphane Lehéricy, co-directeur du CATI), dont l'équipe Cogimage du CRICM (Olivier Colliot et Marie Chupin).

Le CATI a pour vocation de mutualiser les ressources requises pour réaliser des études multicentriques en neuroimagerie. Le centre expert est associé à un réseau de fournisseurs d'images (une trentaine de centres de neuroradiologie, dont un certain nombre de plateformes de recherche, et de médecine nucléaire) ainsi qu'un réseau d'équipes ayant une expertise particulière en analyse d'images cérébrales. Les services du CATI couvrent l'acquisition des données, leur suivi et leur contrôle qualité, l'organisation d'une base de données ainsi que le traitement et, si besoin, l'analyse d'image.

6.2.3 Protocole d'acquisition

La mission prioritaire du CATI étant sur les démences, le protocole d'imagerie CATIv1 est un protocole permettant d'obtenir à la fois des images de bonne qualité pour la recherche tout en gardant une partie pour le diagnostic différentiel des démences. Ce protocole a été développé de façon à être compatible avec le protocole ADNI et a été uniformisé autant que possible sur les centres du réseau, les machines identiques ayant le même protocole, et les différences entre constructeurs étant minimisées quand cela était possible. Le protocole se présentait donc ainsi :

Il est à noter que les séquences dites "cliniques" ou "vasculaires" ont été optimisées en deux temps : au cours de la phase pilote de la cohorte MEMENTO, qui ne rassemblait que cinq centres (dont quatre disposant d'une IRM 3T), les séquences étaient acquises en orientation axiale avec

6.2. Mise en place au sein du Centre pour l'Acquisition et le Traitement des Images (CATI)

N°	Noms CATI	Séquence	Orientation	Durée
1	Survey	Repérage 3 plans	-	~ 0 :10
2	3DT1	3D-T1	Sag stricte	~ 9 :00
3	T2FLAIR	2D-T2 FLAIR	Ax bicalleux	~ 4 :00
4	T2GRE + T2GREph	2D-T2* (GRE) + phase	Ax bicalleux	~ 5 :30
5	T2TSE	2D-T2 TSE/FSE 1 écho	Ax bicalleux	~ 1 :45
6	rsfMRI	IRMf (BOLD EPI) au repos	Ax oblique	~ 10 :00
	DTI1 + DTI2	Tenseur de diffusion	Ax stricte	~ 4 :30
7	DTI3 + DTI4	(DTI-DWI EPI)		× 2-4
	B0MAP + B0MAPph	+ carte de champ B0 + phase		~ 1 :45

TABLEAU 6.1 – Description des séquences du protocole CATIv1.

Les séquences 6 et 7 sont optionnelles. La durée totale du protocole obligatoire est d'environ 20 minutes. L'ajout des séquences optionnelles entraîne une durée totale comprise entre approximativement 40 et 50 minutes d'acquisition.

des coupes de 5mm. Pour la phase de déploiement, les séquences ont pu être modifiées sur les imageurs SIEMENS 3T et Philips 3T, et sont désormais acquises en coupes axiales de 2,5 mm d'épaisseur. Cela n'a pas été possible pour les imageurs GE, car l'option d'imagerie parallèle n'était pas disponible sur la séquence FLAIR 2D et la séquence d'écho de gradient T2* 2D. Dans tous les cas, les acquisitions de ces séquences "vasculaires" sont réalisées dans le plan bicalleux (figure 6.7). Les paramètres d'acquisition pour les séquences T1 et FLAIR sont, eux, résumés dans les tableaux 6.2 et 6.3.

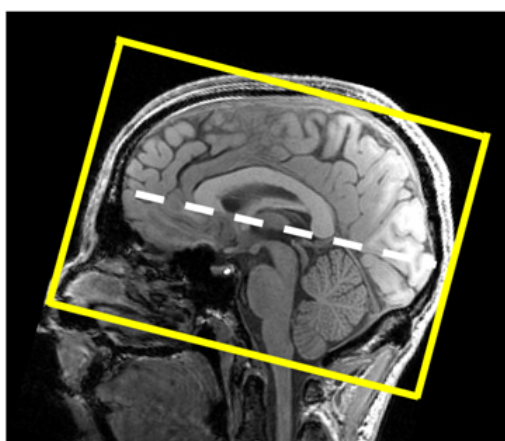


FIGURE 6.7 – Illustration d'une acquisition dans le plan bicalleux.

	1,5 T	3 T
Séquence	Inversion Récupération (IR-FSPGR, MPRAGE)	
Matrice	256 × 256	
FoV	256 × 256 mm ²	
Epaisseur de coupe	1,3 mm	1 mm
Voxel	1 mm isotrope, sans interpolation	
Nex	1	
Phase	Antéro-postérieure	
Facteur d'accélération	NON	
Nb coupes	160	176

TABLEAU 6.2 – Paramètres d'acquisition pour la séquence T1 de la cohorte MEMENTO

	1,5 T	3 T
Séquence	T2 FLAIR	
Matrice	240 × 240 ou 256 × 256	
FoV	240 × 240 mm ² ou 256 × 256 mm ²	
Epaisseur de coupe	5 mm entrelacées jointives	2,5 mm entrelacées jointives
Voxel	1 × 1 mm ² , sans interpolation	
Nex	1	
Phase	Droite - Gauche	
Facteur d'accélération	2, quand disponible	
Nb coupes	35	70

TABLEAU 6.3 – Paramètres d'acquisition pour la séquence FLAIR de la cohorte MEMENTO

6.2.4 Implémentation dans la suite logicielle Brainvisa

Le CATI s'appuie pour une partie des analyses liées à la cohorte MEMENTO sur la suite logicielle Brainvisa, disponible librement en ligne¹. La méthode WHASA proposée dans cette thèse a été implémentée en tant que module disponible à travers Brainvisa par Ludovic Fillon, ingénieur au sein du CATI (figure 6.8). Cette implémentation de WHASA n'est cependant pas disponible dans la version librement accessible sur Internet.

Il est ainsi possible de lancer directement WHASA à travers l'interface graphique de Brainvisa (figure 6.9), soit en utilisant les paramètres par défaut, soit en modifiant ceux-ci. L'utilisation peut également se faire de manière séquentielle ou parallèle sur des architectures disposant de plusieurs processeurs pour permettre le traitement de larges bases de données.

1. <http://brainvisa.info/doc/axon/en/help/aboutBrainVISA.html>

6.2. Mise en place au sein du Centre pour l'Acquisition et le Traitement des Images (CATI)

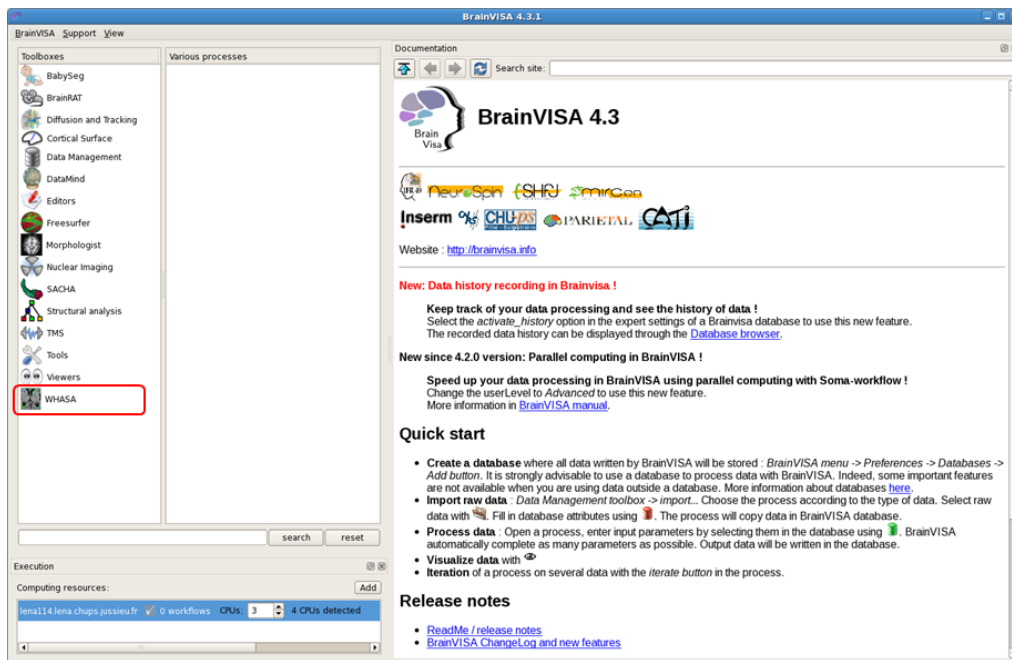


FIGURE 6.8 – Page d'accueil de la suite logicielle Brainvisa. La méthode WHASA est implémentée comme un module de Brainvisa, accessible depuis l'écran d'accueil.

Enfin, la version actuelle de l'implémentation permet d'afficher à la fin du traitement les images $mT1$ et $mFLAIR$, $LPE_{grossière}$ et le résultat de la segmentation HSB_{final} afin d'évaluer rapidement la qualité des résultats obtenus (figure 6.10).

6. APPLICATION SUR D'AUTRES JEUX DE DONNÉES

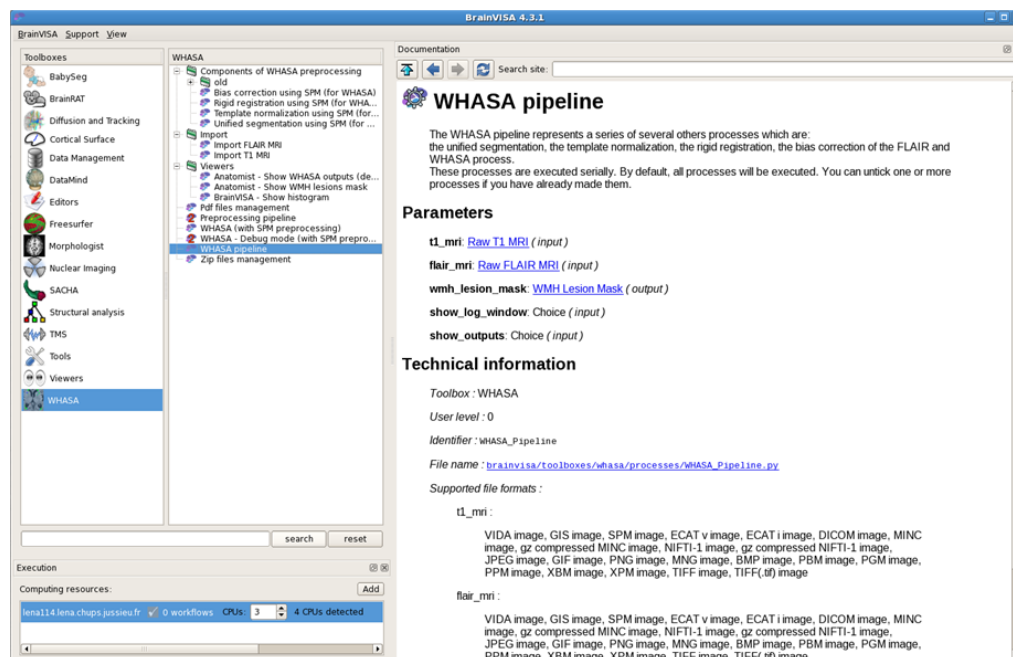


FIGURE 6.9 – Utilisation de WHASA dans Brainvisa.

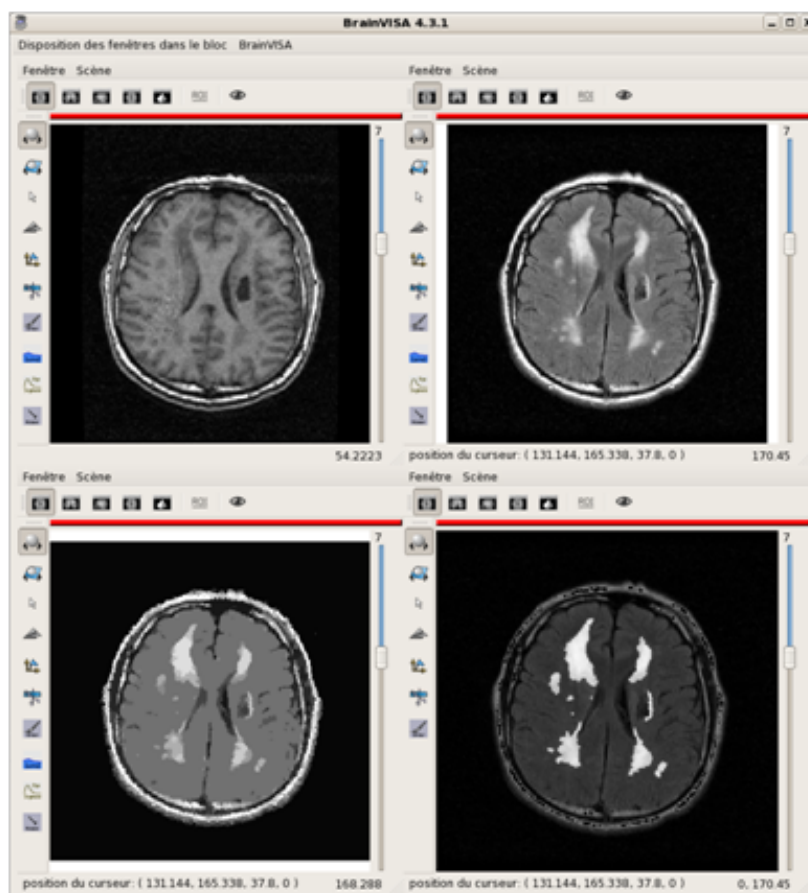


FIGURE 6.10 – Visualisation des résultats de WHASA dans Brainvisa.

6.2.5 Validation de la méthode WHASA sur la cohorte

MEMENTO

6.2.5.1 Estimation visuelle de la qualité de segmentation

WHASA a été appliqué sur les 87 sujets de la phase pilote (acquisitions réalisées dans cinq centres différents) et sur 47 sujets de la phase de déploiement (acquisitions réalisées dans deux des centres participant à la phase pilote, protocole modifié pour avoir une épaisseur de coupe de 2,5 mm sur les images FLAIR). Une estimation visuelle de la qualité de la segmentation a alors été réalisée par Ludovic Fillon, ingénieur du CATI. Une note entre 0 et 4 a ainsi été attribuée pour chacun des sujets.

Un score de 0 indique un échec de la segmentation, correspondant à une absence totale de segmentation alors que le sujet présente des HSB importantes ou, à l'inverse, à la présence dans la segmentation de coupes sur lesquelles aucune HSB n'est présente. Un score de 1 indique une mauvaise segmentation, lorsqu'une partie des HSB est correctement segmentée mais que des erreurs évidentes sont visibles. Un score de 2 est attribué pour une segmentation de qualité moyenne, lorsque les erreurs de segmentation sont situées dans des zones de fort volume partiel. Un score de 3 indique une bonne segmentation avec la majorité des HSB détectées correctement et un bon accord spatial visuel. Un certain nombre de faux positifs demeure cependant ou une sursegmentation des HSB confluentes. Enfin, un score de 4 indique une excellente segmentation, avec une bonne délimitation des HSB isolées, un bon accord spatial visuel avec l'ensemble des lésions et une présence très faible de faux positifs. Les images 6.11, 6.12, 6.13 et 6.14 montrent des exemples de patients correspondant aux différents stades de l'échelle. Aucun patient n'ayant été noté 1 sur les 134 sujets, il n'est pas possible de montrer un exemple pour ce stade.

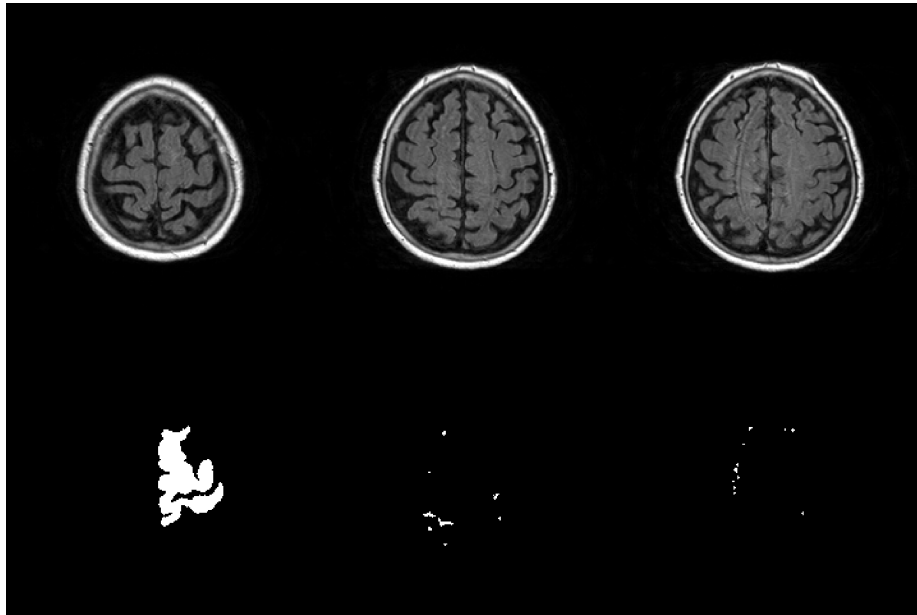


FIGURE 6.11 – Exemple de sujet noté 0 (Echec de la segmentation)

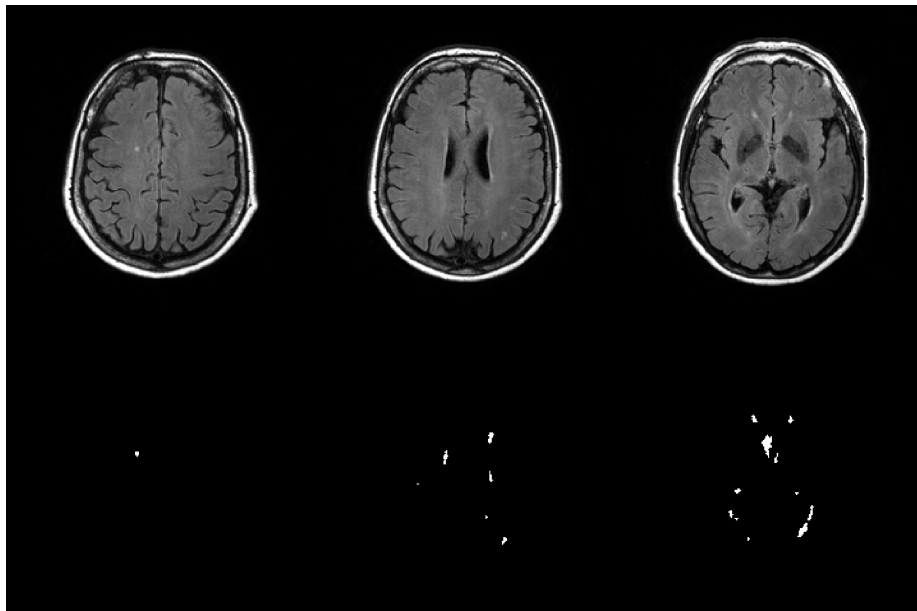


FIGURE 6.12 – Exemple de sujet noté 2 (Segmentation moyenne)

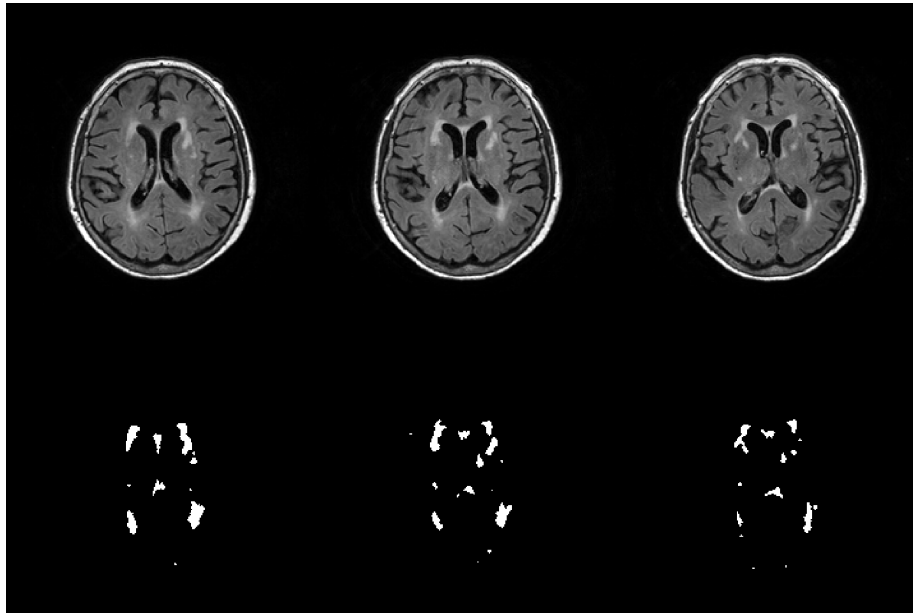


FIGURE 6.13 – Exemple de sujet noté 3 (Segmentation bonne)

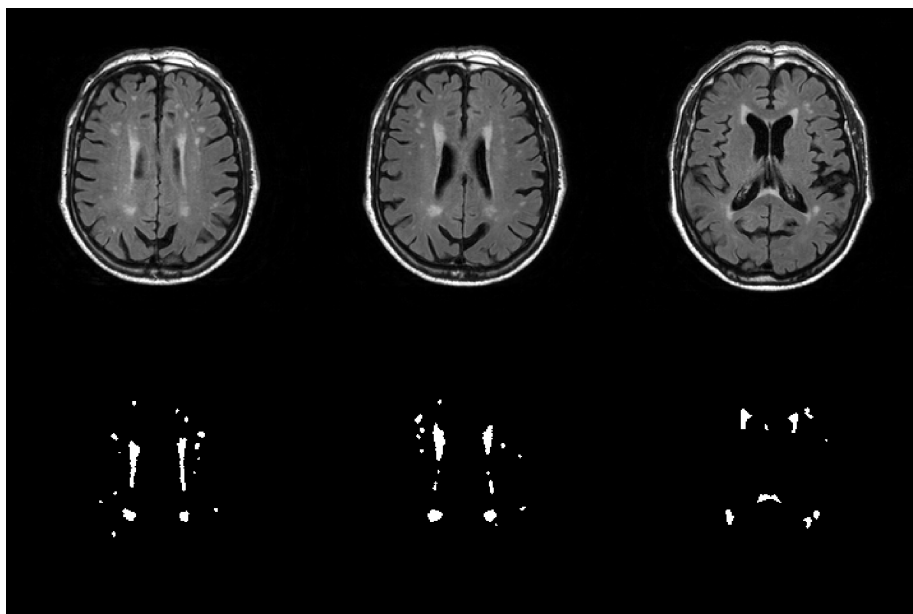


FIGURE 6.14 – Exemple de sujet noté 4 (Segmentation excellente)

6.2.5.2 Résultats

Le tableau 6.4 montre les résultats de la segmentation automatique sur les 87 sujets de la phase pilote et les 47 sujets de la phase déploiement. Seuls 3 échecs sont observés, uniquement sur la phase de déploiement. Après analyse, dans les trois cas, ces échecs étaient dus à une mauvaise conversion des fichiers qui modifiaient le codage des intensités. Le seuil sur les intensités $\tilde{i}$ était alors trop bas, entraînant la sursegmentation de certaines coupes. Hormis ces trois échecs, la qualité de la segmentation est satisfaisante : la segmentation est jugée bonne ou excellente dans 98 % des cas pour la phase pilote et dans 66 % des cas pour la phase déploiement, soit un taux de 87 % pour les deux phases combinées. La plus grande proportion de segmentations de qualité moyenne dans le cas de la phase de déploiement est le plus souvent lié à la présence de petits faux positifs, qui apparaissent plus souvent à cause de la plus faible épaisseur de coupe. Ces résultats demeurent préliminaires mais il est intéressant d'observer que les résultats obtenus sont globalement satisfaisants alors que la méthode a été appliquée sans aucune modification sur des données provenant de machines différentes, de champ différent. Le comportement de la méthode demeure également satisfaisant pour des épaisseurs de coupes plus fines.

Qualité de la segmentation	Phase Pilote	Phase déploiement
0 (Echec)	0 (0%)	3 (6 %)
1 (Faible)	0 (0%)	0 (0%)
2 (Moyenne)	2 (2 %)	13 (28 %)
3 (Bonne)	54 (62 %)	27 (57 %)
4 (Excellente)	31 (36 %)	4 (9 %)
Total	87 (100 %)	47 (100 %)

TABLEAU 6.4 – Estimation visuelle de la qualité de la segmentation sur la phase pilote et la phase de déploiement de la cohorte MEMENTO

6.3 Conclusion

Dans ce chapitre, nous avons montré que WHASA pouvait être appliqué directement sur de nouvelles données en conservant des performances satisfaisantes.

Dans une première étude de 59 sujets, nous montrons une forte concordance entre le volume des HSB segmentées automatiquement et les cotations visuelles. De plus, l'utilisation de WHASA met en évidence une corrélation faible mais significative entre les HSB et l'âge en accord avec la littérature. Cette corrélation n'était pas retrouvée avec l'utilisation des cotations visuelles ou aurait nécessité un plus grand nombre de sujets pour être significative.

Dans un deuxième temps, WHASA a également été appliqué sur la phase pilote et la phase de déploiement (87 sujets et 47 sujets respectivement) d'une importante étude multicentrique, la cohorte MEMENTO. Une estimation visuelle de la qualité de la segmentation automatique montre que la segmentation est jugée bonne ou excellente dans 87% des cas. Il est à souligner ici que les images de la cohorte MEMENTO provenaient de cinq centres différents, mélangeant des imageurs 1,5 T et 3T.

Au total, WHASA a donc été appliqué sur 193 sujets supplémentaires. Il n'a pas été possible d'effectuer une comparaison avec une référence manuelle sur un si grand nombre de sujets. La comparaison avec les échelles visuelles, la corrélation avec l'âge et les estimations visuelles de la qualité des résultats permettent néanmoins une validation indirecte des bonnes performances de WHASA sur six machines IRM différentes.

Conclusion

Les facteurs vasculaires jouent un rôle de plus en plus reconnu dans l'apparition des démences. Alors que de nombreuses méthodes d'analyses existent pour quantifier les processus neurodégénératifs, la prise en compte des facteurs vasculaires est plus éparse. L'étude de ces facteurs requiert le suivi d'un grand nombre de sujets à des stades différents de la pathologie. La constitution de larges études de cohorte rend ainsi indispensable le développement de méthodes d'analyse automatique permettant la caractérisation des anomalies vasculaires.

* *
*

Dans cette thèse, nous avons proposé une nouvelle méthode de segmentation automatique des hyperintensités de la substance blanche (HSB) à partir d'une image T1 et d'une image FLAIR, applicable dans un environnement multicentrique.

Après un rappel du contexte clinique général, nous avons présenté une revue de littérature sur les méthodes de segmentation automatique des HSB. Nous avons ensuite présenté la nouvelle méthode de segmentation des HSB, nommée WHASA. Afin d'être robuste aux conditions d'acquisition et à la variabilité de forme et de localisation des HSB, nous avons introduit une combinaison de diffusion non linéaire et de segmentation par ligne de partage des eaux pour obtenir deux représentations de l'image FLAIR à différents niveaux de détail. L'utilisation de trois critères simples (intensité, localisation dans la substance blanche, non connexité à l'interface substance grise/liquide cérébrospinal) sur le niveau de détail le plus grossier permet d'obtenir une initialisation des HSB. Cette initialisation est ensuite affinée par reconstruction géodésique en utilisant le niveau de détail plus fin pour obtenir la segmentation finale.

Afin de valider cette approche, nous avons comparé la segmentation obtenue par WHASA avec une segmentation manuelle sur un ensemble de 67 sujets, présentant une grande variabilité au niveau des conditions d'acquisition et de la charge lésionnelle. De plus, nous avons réimplémenté ou utilisé plusieurs méthodes proposées dans la littérature ou librement disponibles et comparé leurs performances sur ces mêmes données. La méthode WHASA obtient des résultats sensiblement meilleurs que toutes les autres méthodes évaluées.

Enfin, nous avons appliqué WHASA sur des données provenant d'autres études. Dans le cadre d'une collaboration avec le Centre Hospitalier Régional Universitaire de Lille, les volumes obtenus automatiquement sont cohérents avec les scores des échelles visuelles classiques. La relation entre le volume des HSB et l'âge est également retrouvée. Dans le cadre du développement du Centre pour l'Acquisition et le Traitement des Images (CATI), WHASA a également été intégré à l'environnement Brainvisa et utilisé sur 87 patients de la phase pilote et 47 patients de la phase de déploiement d'une large étude de cohorte multicentrique, la cohorte MEMENTO, avec des performances satisfaisantes.

Au total, WHASA a donc été appliqué sur 260 sujets, provenant d'une dizaine de centres, comprenant à la fois des imageurs 1,5 T et 3 T et avec des épaisseurs de coupes pour les images FLAIR variant entre 2,5 mm et 6,2 mm.

* *
*

Ce travail de thèse offre différentes perspectives.

Dans un premier temps, une extension naturelle de la méthode WHASA serait son adaptation au traitement de données longitudinales. La volumétrie apparaît en effet beaucoup plus sensible que les échelles visuelles pour distinguer une évolution des HSB [van den Heuvel et al., 2006]. L'analyse de l'évolution des HSB serait un indicateur important du mécanisme pathologique sous-jacent. Elle est également essentielle pour évaluer l'effet de potentiels traitements.

Dans un second temps, les volumes de HSB peuvent être intégrés dans les méthodes de classification afin d'améliorer le diagnostic différentiel des sous-types de démence. En particulier, la prise en compte des

anomalies vasculaires pourrait permettre d'expliquer en partie les insuffisances des méthodes actuelles pour la prédiction de la conversion du stade prédéméntiel au stade démentiel [Cuingnet et al., 2011].

Récemment, le développement de marqueurs du peptide amyloïde en tomographie par émission de positons (TEP) a également rendu possible la détection *in vivo* du peptide amyloïde. Les premières applications dans le cas de la MA mettent en évidence une proportion non négligeable de patients présentant d'importants dépôts amyloïdes mais peu de symptômes cliniques [Aizenstein et al., 2008]. Certaines études préliminaires, portant sur un faible nombre de patients, semblent indiquer que les HSB constitueraient alors un facteur aggravant permettant d'expliquer les différences cliniques [Provenzano et al., 2013]. La méthode WHASA permettrait alors de confirmer cette observation à plus grande échelle.

Par ailleurs, l'obtention d'une segmentation des HSB peut permettre des analyses plus fines de leur topographie et de leur localisation. Au niveau de la topographie, Yoshita et al. [2006] ont par exemple montré que les patients souffrant de la maladie d'Alzheimer présentaient une distribution des HSB plus postérieure que les sujets sains. Au niveau de la localisation, il a été suggéré que l'impact clinique des HSB serait notamment lié à une interruption des faisceaux cholinergiques [Bohnen et al., 2009 ; Bocti et al., 2005 ; Behl et al., 2007 ; Keverne et al., 2007]. La combinaison de la segmentation des HSB avec les méthodes de tractographie pourrait permettre de tester cette hypothèse.

Enfin, les HSB sont l'anomalie vasculaire la plus fréquente chez les personnes âgées. Néanmoins, la prise en compte d'autres anomalies vasculaires, telles que les lacunes ou les microsaignements, est nécessaire pour caractériser de manière plus complète l'intégrité vasculaire chez les patients. Couplée avec des méthodes de détection des microsaignements, des lacunes, et les méthodes déjà existantes de mesure d'atrophie, la méthode WHASA est une brique permettant une caractérisation plus complète des différents processus pouvant mener à la démence.

Valorisation de la thèse

Article de journal à comité de lecture

T. Samaille, L. Fillon, R. Cuingnet, E. Jouvent, H. Chabriat, D. Dormont, O. Colliot, M. Chupin, *Contrast-based fully automatic segmentation of white matter hyperintensities : method and validation*, PLoS ONE. 2012, 7(11) :e48953. DOI :10.1371/journal.pone.0048953.

Articles de conférence à comité de lecture

T. Samaille, O. Colliot, R. Cuingnet, E. Jouvent, H. Chabriat, D. Dormont, M. Chupin, *Automatic segmentation of white matter hyperintensities robust to multicentre acquisition and pathological variability*, Proceedings of the SPIE symposium on Medical Imaging (Image Processing) 2012, Volume 8314, pp. 831446-9 (2012). DOI : 10.1117/12.910268.

T. Samaille, O. Colliot, D. Dormont, M. Chupin, *Automatic segmentation of age-related white matter changes on flair images : Method and multicentre validation*. Proceedings of the 9th IEEE International Symposium on Biomedical Imaging : From Nano to Macro (ISBI' 11), Chicago, USA, March 30-April 2, 2011, pp. 2014-2017.

Actes de conférences

L. Fillon, T. Samaille, E. Jouvent, H. Chabriat, D. Dormont, O. Colliot, M. Chupin, *Segmentation of white matter hyperintensities : a comparison of automatic methods on multicentre data*, Proceedings of the 18th International Conference on Functional Mapping of the Human Brain, 2012, Beijing, China.

T. Samaille, O. Colliot, D. Dormont, M. Chupin, *Automated segmentation of white matter lesions using FLAIR images*, Proceedings of the 16th International Conference on Functional Mapping of the Human Brain, 2010, Barcelona, Spain.

Logiciel

La méthode WHASA de segmentation des hyperintensités de la substance blanche est déposé auprès de l'Agence de Protection des Programmes (Numéro temporaire DL 04554-01).

Annexe A

Critères diagnostiques des démences

A.1 Critères diagnostiques de la démence de type Alzheimer, DSM-IV-TR

A Apparition de déficits cognitifs multiples, comme en témoignent à la fois :

- a) une altération de la mémoire (altération de la capacité à apprendre des informations nouvelles ou à se rappeler les informations apprises antérieurement) ;
- b) une (ou plusieurs) des perturbations cognitives suivantes :
 - i. aphasie (perturbation du langage)
 - ii. apraxie (altération de la capacité à réaliser une activité motrice malgré des fonctions motrices intactes)
 - iii. agnosie (impossibilité de reconnaître ou d'identifier des objets malgré des fonctions sensorielles intactes)
 - iv. perturbation des fonctions exécutives (faire des projets, organiser, ordonner dans le temps, avoir une pensée abstraite)

B Les déficits cognitifs des critères A1 et A2 sont tous les deux à l'origine d'une altération significative du fonctionnement social ou professionnel et représentent un déclin significatif par rapport au niveau de fonctionnement antérieur.

C L'évolution est caractérisée par un début progressif et un déclin cognitif continu.

D Les déficits cognitifs des critères A1 et A2 ne sont pas dus :

- a) à d'autres affections du système nerveux central qui peuvent entraîner des déficits progressifs de la mémoire et du fonctionnement cognitif (par exemple : maladie cérébrovasculaire, maladie de Parkinson, maladie de Huntington, hématome sous-dural, hydrocéphalie à pression normale, tumeur cérébrale) ;
- b) à des affections générales pouvant entraîner une démence (par exemple : hypothyroïdie, carence en vitamine B12 ou en folates, hypercalcémie, neurosyphilis, infection par le VIH) ;
- c) à des affections induites par une substance.

E Les déficits ne surviennent pas de façon exclusive au cours de l'évolution d'un syndrome confusionnel

F La perturbation n'est pas mieux expliquée par un autre trouble de l'Axe I du DSM-IV-TR (par exemple : trouble dépressif majeur, schizophrénie)

Codification fondée sur la présence ou l'absence d'une perturbation cliniquement significative du comportement :

Sans perturbation du comportement : si les troubles cognitifs ne s'accompagnent d'aucune perturbation cliniquement significative du comportement.

Avec perturbation du comportement : si les troubles cognitifs s'accompagnent d'une perturbation cliniquement significative (par exemple : errance, agitation) du comportement.

Préciser le sous-type :

A début précoce : si le début se situe à 65 ans ou avant.

A début tardif : si le début se situe après 65 ans.

A.2 Critères diagnostiques de la maladie d'Alzheimer, NINCDS-ADRDA

1. Critères de maladie d'Alzheimer probable :
 - syndrome démentiel établi sur des bases cliniques et documenté par le *Mini Mental State Examination*, le *Blessed Dementia Scale* ou tout autre test équivalent et confirmé par des preuves neuropsychologiques
 - déficit d'au moins deux fonctions cognitives
 - altérations progressives de la mémoire et des autres fonctions cognitives
 - absence de trouble de conscience
 - survenue entre 40 et 90 ans, le plus souvent au-delà de 65 ans
 - en l'absence de désordres systémiques ou d'une autre maladie cérébrale pouvant rendre compte par eux-mêmes, des déficits mnésiques et cognitifs progressifs
2. Ce diagnostic de maladie d'Alzheimer probable est renforcé par :
 - la détérioration progressive des fonctions telles que le langage (aphasie), les habilités motrices (apraxie) et perceptives (agnosie)
 - la perturbation des activités de la vie quotidienne et la présence de troubles du comportement
 - une histoire familiale de troubles similaires surtout si confirmés histologiquement
 - le résultat aux examens standards suivants :
 - normalité du liquide céphalo-rachidien
 - EEG normal ou siège de perturbations non spécifiques comme la présence d'ondes lentes
 - présence d'atrophie cérébrale d'aggravation progressive
3. Autres caractéristiques cliniques compatibles avec le diagnostic de maladie d'Alzheimer probable après exclusion d'autres causes :
 - période de plateaux au cours de l'évolution
 - présence de symptômes tels que la dépression, insomnie, incontinence, idées délirantes, illusions, hallucinations, réactions de catastrophe; désordres sexuels et perte de poids. Des anomalies neurologiques sont possibles surtout aux stades évolués de la maladie, notamment des signes moteurs tels qu'une hypertonie, des

- myocloines ou des troubles de la marche.
 - crises comitiales aux stades tardifs
 - scanner cérébral normal pour l'âge
4. Signes rendant le diagnostic de la maladie d'Alzheimer probable incertain ou improbable :
- début brutal
 - déficit neurologique focal tel que hémiparésie, hypoesthésie, déficit du champ visuel, incoordination motrice à un stade précoce
 - crises convulsives ou troubles de la marche en tout début de maladie
5. Le diagnostic clinique de la maladie d'Alzheimer possible :
- peut être porté sur la base du syndrome démentiel, en l'absence d'autre désordre neurologique, psychiatrique ou systémique susceptible de causer une démence, en présence de variante dans la survenue, la présentation ou le cours de la maladie ;
 - peut être porté en présence d'une seconde maladie systémique ou cérébrale susceptible de produire un syndrome démentiel mais qui n'est pas considéré comme la cause de cette démence ;
 - et pourrait être utilisé en recherche clinique quand un déficit cognitif sévère progressif est identifié en l'absence d'autre cause identifiable.
6. Les critères de diagnostic de la maladie d'Alzheimer certaine sont :
- les critères cliniques de la maladie d'Alzheimer probable ;
 - et la preuve histologique apportée par la biopsie ou l'autopsie.

A.3 Critères de démence vasculaire, DSM-IV-TR

A Apparition de déficits cognitifs multiples, comme en témoignent à la fois :

- a) une altération de la mémoire (altération de la capacité à apprendre des informations nouvelles ou à se rappeler les informations apprises antérieurement) ;
- b) une (ou plusieurs) des perturbations cognitives suivantes :
 - i. aphasie (perturbation du langage)
 - ii. apraxie (altération de la capacité à réaliser une activité motrice malgré des fonctions motrices intactes)
 - iii. agnosie (impossibilité de reconnaître ou d'identifier des objets malgré des fonctions sensorielles intactes)
 - iv. perturbation des fonctions exécutives (faire des projets, organiser, ordonner dans le temps, avoir une pensée abstraite)

B Les déficits cognitifs des critères A1 et A2 sont tous les deux à l'origine d'une altération significative du fonctionnement social ou professionnel et représentent un déclin significatif par rapport au niveau de fonctionnement antérieur.

C Signes et symptômes neurologiques en foyer (p. ex., exagération des réflexes ostéo-tendineux, réflexe cutané plantéaire en extension, paralysie pseudo-bulbaire, troubles de la marche, faiblesse d'une extrémité) ou mise en évidence d'après les examens complémentaires d'une maladie cérébro-vasculaire (p. ex., infarctus multiples dans le cortex et la substance blanche sous-corticale) jugée liée étiologiquement à la perturbation.

D Les déficits ne surviennent pas exclusivement au cours de l'évolution d'un syndrome confusionnel.

Codification fondée sur les caractéristiques prédominantes :

Avec syndrome confusionnel : si un syndrome confusionnel est surajouté à la démence.

Avec idées délirantes : si les idées délirantes sont la caractéristique prédominante.

Avec humeur dépressive : si l'humeur dépressive (notamment des tableaux cliniques comportant les critères symptomatiques d'un épisode dépressif majeur) est la caractéristique prédominante, on ne fait pas un diagnostic séparé de troubles de l'humeur dus à une affection médicale générale.

Non compliquée : ce sous-type est utilisé si aucun des symptômes ou syndromes ci-dessus ne prédomine dans le tableau actuel.

Spécifier si : (peut s'appliquer à n'importe lequel des sous-types cités ci-dessus).

Avec perturbation du comportement : s'il existe une perturbation du comportement cliniquement significative (p. ex., errance)

A.4 Critères diagnostiques de la démence vasculaire probable ou possible, NINDS-AIREN

Ces critères associent une démence avec la définition suivante, une maladie cérébro-vasculaire, et une relation entre les deux, établie soit par la survenue de la démence dans les 3 mois qui suivent l'accident vasculaire cérébral, soit par une évolution fluctuante, en marche d'escalier.

Les critères suivants permettent de diagnostiquer une démence vasculaire :

- **Probable**

1. **Démence :**

- déclin cognitif par rapport au niveau antérieur
- déclin mnésique et d'au moins 2 autres domaines cognitifs
- interférant avec les activités de la vie quotidienne

2. **Maladie cérébro-vasculaire :**

- présence de signes focaux à l'examen neurologique
- preuve TDM ou IRM de lésion vasculaire cérébrale

3. **Relation entre les 2 affections ci-dessus, d'après :**

- début de la démence dans les 3 mois après un AVC
- détérioration brusque des fonctions cognitives, ou aggravation fluctuante ou par à coups des déficits cognitifs

4. **Critères en faveur du diagnostic (facultatifs) :**

- troubles de la marche, précoces
- antécédents d'instabilité, de chutes spontanées
- troubles du contrôle mictionnel
- paralysie pseudo-bulbaire, incontinence émotionnelle
- modification de la personnalité et de l'humeur

- **Possible :** Mêmes que probable, sauf :

- Absence de preuve TDM ou IRM de lésion vasculaire cérébrale
- Ou absence de relation temporelle claire entre démence et AVC
- Ou début insidieux à évolution variable (plateau, amélioration)

- **Certaine :** Mêmes que probable, plus :

- Signes histopathologiques de maladie cérébro-vasculaire
- Absence de dégénérescence neurofibrillaire et de plaques séniles

- **Absence d'autres affections cliniques ou neuropathologiques pouvant être la cause de démence**

Annexe B

Notions de base sur l'imagerie par résonance magnétique (IRM)

Nous nous contenterons ici d'exposer l'origine du signal et les principes physiques permettant d'obtenir différents types de contrastes, appelés pondérations, classiques en IRM (T1, T2, T2* et FLAIR). En particulier, les principes de codage spatial, les séquences d'imagerie rapide et les techniques plus avancées telles que l'imagerie de diffusion ou fonctionnelle ne seront pas évoquées. Pour plus de détails, le lecteur intéressé peut se référer à McRobbie et al. [2007] ou, concernant les aspects mathématiques, à Liang & Lauterbur [2000].

B.1 Origine du signal

L'IRM repose sur le moment magnétique, noté $\vec{\mu}$, du noyau d'hydrogène, présent notamment dans les molécules d'eau du corps humain. Bien que le moment magnétique soit une propriété quantique, une analogie classique, bien que limitée, est de le considérer comme l'aiguille d'une boussole.

La propriété de résonance magnétique nucléaire désigne le fait que, lorsqu'ils sont placés dans un champ magnétique $\vec{B}_0$, ces moments magnétiques vont se mettre à tourner autour de l'axe de $\vec{B}_0$ avec une vitesse angulaire $\omega_0 = \gamma B_0$, où γ est une constante dépendant du noyau considéré, appelée rapport gyromagnétique. On préfère généralement utiliser la notation en fréquence et la fréquence de rotation des moments ma-

B. NOTIONS DE BASE SUR L'IMAGERIE PAR RÉSONANCE MAGNÉTIQUE (IRM)

gnétiques $f_0 = \frac{\gamma}{2\pi} B_0$ est appelée *fréquence de Larmor*¹. Ce mouvement de rotation du moment magnétique est appelé *précession* (figure B.1).

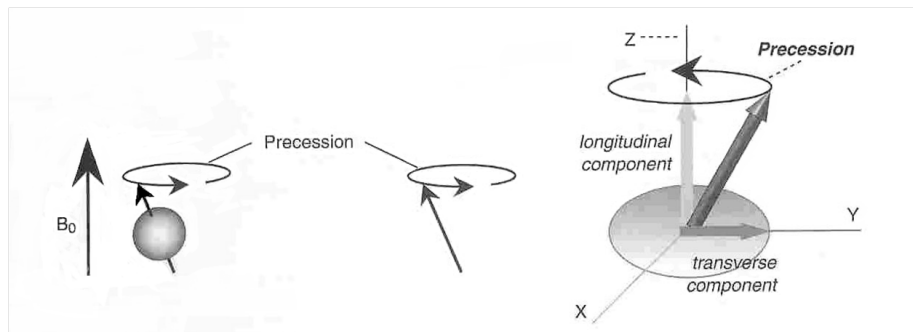


FIGURE B.1 – Mouvement de précession du spin autour de l'axe du champ magnétique (Source : http://www.sas.upenn.edu/~wwalsh/MRI_BASICS.html).

Lorsque de nombreux protons en interaction avec leur environnement sont placés dans un champ magnétique, ils vont donc se mettre en mouvement de précession autour de l'axe de ce champ magnétique et vont progressivement s'orienter de manière à être colinéaires avec le champ $\vec{B}_0$. Deux orientations sont alors possibles : parallèle et antiparallèle. Les principes de physique statistique impliquent qu'un plus grand nombre de moments magnétiques vont se placer dans la configuration parallèle, plus stable énergétiquement. Ainsi, bien que les moments parallèles et antiparallèles s'annulent deux à deux, la somme de tous les moments magnétiques, ou résultante, est non nulle. On note généralement cette résultante $\vec{M}$. Pour simplifier les notations, on considère un référentiel tel que l'axe $\vec{z}$ est parallèle au champ $\vec{B}_0$ et on a donc $\vec{M} = M \vec{z}$ (figure B.2).

Le signal de résonance magnétique nucléaire est alors obtenu par l'application transitoire dans le plan perpendiculaire à $\vec{B}_0$ d'un second champ magnétique, noté $\vec{B}_1$, oscillant à la fréquence de Larmor du noyau étudié. Ce champ $\vec{B}_1$ est également appelé onde radiofréquence (ou onde RF). Les moments magnétiques se mettent alors également en précession autour de $\vec{B}_1$. Il apparaît donc une composante transversale M_{xy} à la résultante $\vec{M}$, qui peut se décomposer de la manière suivante : $\vec{M} = M_z \vec{z} + M_{xy} \vec{x}\vec{y}$. C'est l'amplitude de ce champ magnétique transversal qui est mesurée à l'aide d'une bobine par le principe d'induction. En fonction de la durée de l'impulsion, il est possible de faire basculer de

1. du nom de Sir Joseph Larmor, physicien irlandais (1^{er} juillet 1857 - 19 mai 1942)

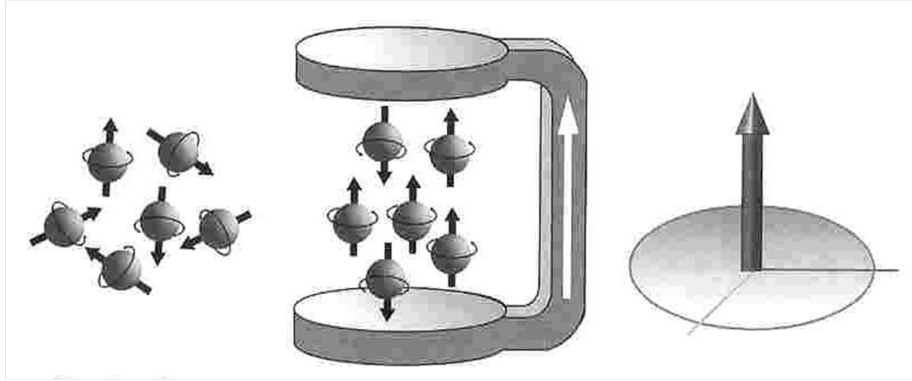


FIGURE B.2 – Origine du vecteur de magnétisation résultant (Source : http://www.sas.upenn.edu/~wwalsh/MRI_BASICS.html).

Gauche : Hors d'un champ magnétique, la distribution de l'orientation des spins est aléatoire. **Milieu** : Placés dans un champ magnétique, les spins s'orientent colinéairement à l'axe du champ magnétique. Statistiquement, le nombre de spins dans l'état parallèle (ici, 4) est légèrement supérieur à celui des spins dans l'état antiparallèle (ici, 2). **Droite** : La somme des spins est donc non nulle et peut être modélisée par un unique vecteur résultant.

manière plus ou moins importante le vecteur résultant par rapport à l'axe du champ $\vec{B}_0$. La composante transversale maximale que l'on peut obtenir correspond au cas où la totalité de l'aimantation longitudinale initiale est basculée dans le plan perpendiculaire, c'est-à-dire un angle de bascule de 90° .

A la fin de l'impulsion B_1 , les moments magnétiques reprennent leur position initiale, c'est le phénomène de relaxation. Cette relaxation ne s'effectue pas à la même vitesse pour la composante longitudinale et pour la composante transversale, mais fait intervenir deux constantes de temps T_1 et T_2 :

$$\begin{aligned} M_z(t) &= M_z^0 \left(1 - e^{-\frac{t}{T_1}} \right) \\ M_{xy}(t) &= M_{xy}^0 e^{-\frac{t}{T_2}} \end{aligned}$$

Les constantes de temps T_1 et T_2 permettent de caractériser la vitesse de ces phénomènes de relaxation. Le temps de relaxation T_1 indique la durée nécessaire, après une impulsion de 90° , pour que l'aimantation longitudinale récupère approximativement 63% de sa valeur initiale avant impulsion. Le temps T_2 représente la durée nécessaire pour que la composante transversale perde environ 67% de sa valeur après une impulsion RF de bascule de 90° . Ainsi, le signal mesuré par une bobine placée dans le plan perpendiculaire est une sinusoïde amortie par une exponentielle

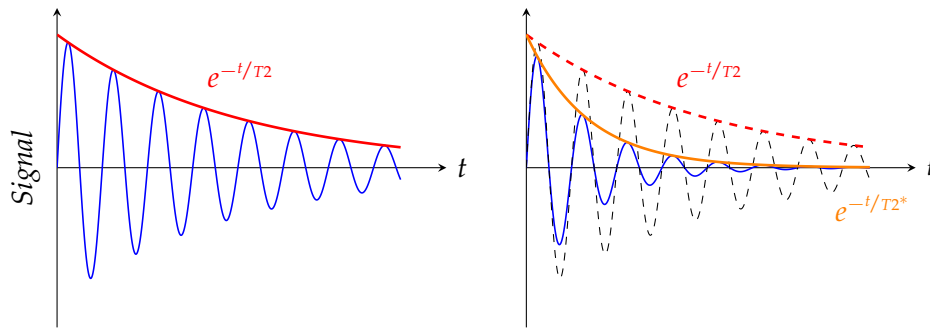


FIGURE B.3 – Signal mesuré après une onde de radiofréquence de 90° . **Gauche** : Signal théorique dans un champ parfaitement homogène avec une décroissance caractérisée par T_2 (*Free Induction Decay*). **Droite** : En réalité, les inhomogénéités locales de champ entraînent une décroissance plus rapide du signal, caractérisée par le temps T_2^* .

de constante de temps T_2 . Ce signal est appelé *Free Induction Decay* (FID). En réalité cependant, le champ magnétique $\vec{B}_0$ n'est pas parfaitement homogène, ce qui entraîne un déphasage des spins, et donc une décroissance du signal plus rapide que prévu. On parle alors de temps T_2^* pour désigner la constante de temps de cette exponentielle décroissant plus rapidement (figure B.3). Il est possible de corriger ces inhomogénéités par l'envoi d'une impulsion de 180° , que nous ne détaillerons pas ici.

B.2 Pondération des images

Le principe des séquences IRM est de manipuler les champs magnétiques et les processus de phasage/déphasage des spins pour obtenir un signal mettant en évidence les différences de temps T_1 et T_2 existant entre les tissus. Ces constantes de temps sont en effet dépendantes de l'agitation moléculaire et donc de la structure d'un matériau ainsi que des interactions avec les molécules environnantes. Le tableau B.1 montre par exemple les constantes de temps T_1 et T_2 pour différents types de tissus cérébraux.

Les séquences IRM s'appuient essentiellement sur deux paramètres, appelés *Temps de répétition* (TR) et *Temps d'écho* (TE). TR désigne le temps séparant l'envoi de deux ondes RF, TE celui séparant l'envoi de l'onde de radiofréquence de la lecture du signal.

Le mécanisme pour mesurer la différence entre des tissus ayant des T_2 différents est le plus facile à comprendre. Considérons deux tissus A

Tissue	Muscle	SG	SB	LCS
T1 (ms)	870	920	790	2400
T2 (ms)	47	100	92	160

TABLEAU B.1 – Estimation des constantes de temps T1 et T2 pour différents types de tissus à 1.5T (SB = Substance grise ; SG= Substance blanche ; LCS = Liquide cérébrospinal). (Source : http://en.wikibooks.org/wiki/Basic_Physics_of_Nuclear_Medicine/MRI_&_Nuclear_Medicine).

et B tels que $T2_A \leq T2_B$, auxquels on applique une impulsion de bascule de 90° . Toute la composante longitudinale est donc basculée dans le plan transversal. Si l'on attend très peu de temps avant de lire le signal (TE court), le signal provenant des deux tissus n'aura pas eu le temps de décroître ni pour A, ni pour B, et il y aura peu de contraste entre les deux tissus. Si, au contraire, on attend plus longtemps avant de lire le signal (TE long), le tissu A, qui a un T2 plus court, aura perdu plus rapidement son aimantation transversale et émettra donc un signal plus faible que celui du tissu B. Il y a donc création d'un contraste provenant majoritairement de la différence de T2, on parle de **pondération T2** (figure B.4). Si l'on veut mesurer le signal en un autre point du volume, il faut recommencer ce processus en attendant un temps suffisamment long pour que l'aimantation longitudinale soit revenue à sa valeur initiale. Il faut donc également un TR long.

Pour observer un contraste basé sur le T1, il faut cette fois parvenir à ce que l'aimantation transversale soit dépendante du T1. L'une des possibilités pour obtenir cet effet est d'utiliser un TR court. En effet, supposons désormais $T1_A \leq T1_B$. Après une impulsion RF de 90° , l'aimantation longitudinale du tissu A va repousser plus lentement que celle du tissu B. Si l'on envoie une nouvelle onde de radiofréquence au moment où le tissu B a récupéré la quasi totalité de son aimantation, l'aimantation transversale obtenue sera plus grande pour le tissu B que pour le tissu A. Il est alors nécessaire de lire rapidement le signal (TE court) pour observer cette différence avant qu'elle ne disparaisse à cause des processus de relaxation. Dans ce cas, la différence de signal observée est obtenue à partir de la différence de T1 et on parle de **pondération T1** (figure B.5).

Enfin, certaines séquences, dites d'**inversion-récupération** (figure B.6) font également intervenir un troisième paramètre, le *temps d'inversion* (TI). Dans ce cas, on envoie une impulsion RF de 180° avant l'impulsion RF de

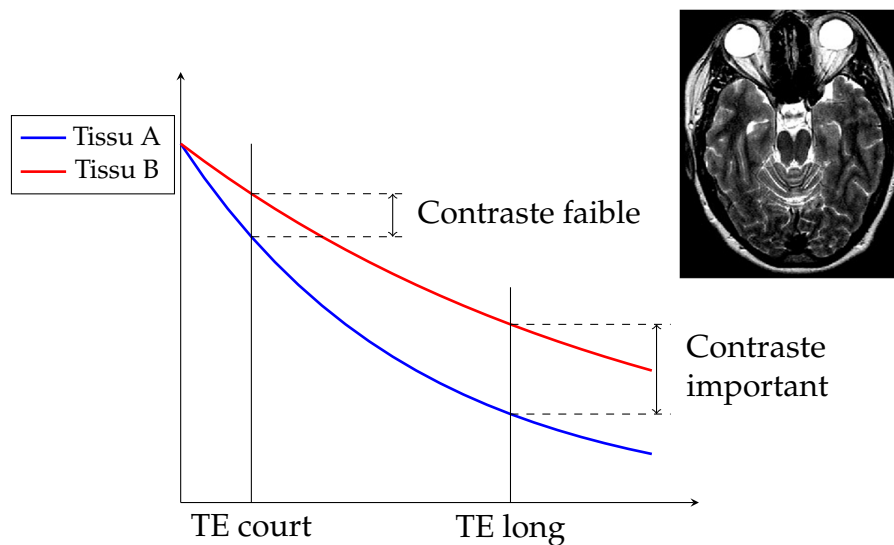


FIGURE B.4 – Principe de la pondération T2.

Les courbes représentent le signal mesuré dans le plan transversal pour un tissu A et un tissu B tel que $T2_A \leq T2_B$. Le contraste entre les deux tissus sera dépendant du temps d'écho, c'est-à-dire de la durée entre l'onde RF et la lecture du signal. Le schéma montre de plus que le signal du tissu ayant un temps T2 court sera plus important que celui du tissu ayant un temps T2 long. Le LCS ayant le T2 le plus long, son signal dans une séquence pondérée en T2 sera plus important que celui de la substance grise et de la substance blanche comme montré sur l'image illustrative.

basculer de 90° et TI désigne le délai entre l'onde d'inversion (180°) et celle d'excitation (90°). L'onde d'inversion a pour effet d'inverser le vecteur résultant, qui devient donc antiparallèle au champ $\vec{B}_0$ ². Le phénomène de relaxation va ensuite entraîner la repousse de l'aimantation longitudinale, repousse dont la vitesse sera dépendante de T1. En choisissant à quel moment de la repousse est envoyé l'onde de 90° , il est possible de choisir la valeur de l'aimantation qui sera basculée dans le plan transversal.

En particulier, puisque l'aimantation longitudinale passe d'une valeur négative à une valeur positive, elle va s'annuler à un moment donné. Si l'on choisit pour valeur de TI le temps mis par un tissu pour que son aimantation longitudinale s'annule, son aimantation transversale après la bascule de 90° sera alors nulle elle aussi. Autrement dit, en fonction du TI choisi, il est possible d'annuler le signal provenant d'un tissu particulier tout en gardant une pondération choisie pour les autres tissus.

2. En mécanique quantique, cela revient à considérer que tous les spins sont passés dans l'état de plus haute énergie.

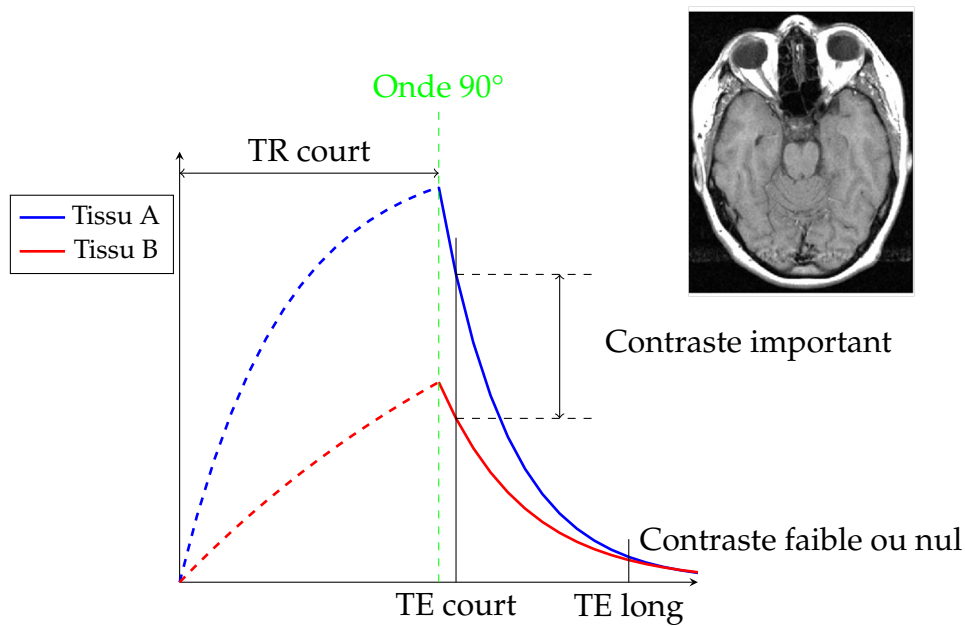


FIGURE B.5 – Principe de la pondération T1.

Les courbes représentent le signal mesuré dans le plan transversal pour un tissu A et un tissu B tel que $T1_A \leq T1_B$. Le contraste entre les deux tissus sera dépendant du temps de répétition TR. De plus, la lecture du signal doit être faire rapidement, avant que le contraste ne disparaisse à cause de la relaxation T2. Le schéma montre que le signal du tissu ayant un temps T1 court sera plus important que celui du tissu ayant un temps T1 long. Le LCS ayant le T1 le plus long, son signal dans une séquence pondérée en T1 sera plus faible que celui de la substance grise et de la substance blanche comme montré sur l'image illustrative.

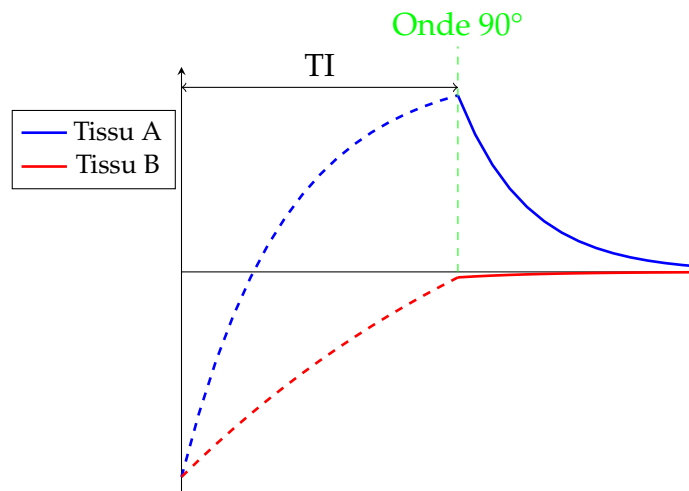


FIGURE B.6 – Inversion-récupération.

La composante longitudinale a été inversée à l'aide d'une onde de 180°. Les traits pointillés indiquent la repousse de la composante longitudinale. La courbe pleine indique la décroissance du signal T2 observée dans le plan perpendiculaire. Ici, la figure illustre comment le TI peut être adapté pour que l'onde de 90° soit envoyée au moment où le signal du tissu B est nul.

Table des figures

1.1	Estimation du nombre de personnes dans le monde atteintes de démence entre 2010 et 2050	5
1.2	Répartition des diagnostics de sous-types de démence en fonction de l'âge	8
1.3	Plaque amyloïde.	11
1.4	Dégénérescence neurofibrillaire dans un neurone du cortex.	12
1.5	Illustration du rôle physiologique de la protéine tau.	12
1.6	Evolution de la pathologie tau en 10 étapes selon les régions cérébrales touchées.	13
1.7	Préparation anatomique mettant en évidence les grandes, moyennes et petites artères cérébrales	14
1.8	Artériolosclérose, caractérisée par un épaississement concentrique de la paroi du vaisseau.	14
1.9	Angiopathie amyloïde affectant la paroi d'un vaisseau.	15
1.10	Représentation schématique des trois principaux sous types de démence vasculaire.	16
1.11	Exemples de différents contrastes IRM sur le même plan de coupe pour un même sujet.	19
1.12	Illustrations des trois types d'anomalies vasculaires en IRM.	21
1.13	Analyse postmortem des HSB.	23
2.1	Echelle modifiée de Fazekas	28
2.2	Echelle de Fazekas & Schmidt.	29
2.3	Exemples de coupes des IRM de référence utilisée par les études CHS et ARIC pour la cotation des HSB	30
2.4	Seuillage de l'histogramme FLAIR [Jack et al., 2001]	31
2.5	Principe de la segmentation des HSB par Admiraal-Behloul et al. [2005]	33

3.1	Schéma général de la chaîne de traitements WHASA.	43
3.2	Schéma de principe de la segmentation unifiée	44
3.3	Recalage rigide entre l'espace T1 et l'espace FLAIR	46
3.4	Définition et graphes des fonctions de diffusivité g_1 , g_2 et g_3 . . .	49
3.5	Représentation graphique des fonctions de flux Φ_1 , Φ_2 et Φ_3 . .	50
3.6	Comparaison du processus de diffusion non linéaire pour les trois fonctions de diffusivité g_1 , g_2 et g_3 pour différents nombre d'itérations.	53
3.7	Détermination de l'interface substance blanche/substance grise (SB SG) dans l'espace FLAIR.	54
3.8	Effet de la diffusion non linéaire sur une image $mFLAIR$ pour différentes valeurs du paramètre de contraste ($\Delta t = 0, 1, 10000$ itérations).	55
3.9	Principe de la ligne de partage des eaux	56
3.10	Illustration de l'effet de la LPE sur une image $mFLAIR$ non diffusée	57
3.11	Illustration de l'effet de la LPE sur une image diffusée	57
3.12	Résultat de la LPE sur une image diffusée avec différents nombres d'itérations ($\Delta t = 0, 1$).	58
3.13	Critère d'arrêt pour le processus de segmentation de l'image $mFLAIR$	58
3.14	Comparaison de $LPE_{grossière}$ et de $LPE_{détaillée}$	60
3.15	Histogramme des intensités d'images $mFLAIR$ avec différentes charges lésionnelles.	61
3.16	Effet de la carte de probabilité utilisée pour déterminer la pro- babilité moyenne d'une région	63
3.17	Procédure d'élimination des régions corticales.	64
3.18	Détermination de l'interface $(SG LCS)_{cortex}$	65
3.19	Echec du processus de diffusion sur certaines coupes	66
3.20	Mauvaise détection des HSB temporales	67
3.21	Principe de la dilatation géodésique.	68
3.22	Apport de $LPE_{détaillée}$	69
3.23	Illustration de l'étape de croissance dans le cas de l'échec du processus de diffusion sur certaines coupes.	71
3.24	Illustration de l'étape de croissance dans le cas d'une mauvaise détection des HSB temporales.	72

4.1	Répartition des volumes manuels pour les 24 sujets des données <i>Hippocampe</i>	76
4.2	Répartition des volumes manuels pour les 43 sujets des données <i>Cadasil</i>	76
4.3	Schéma présentant les différentes catégories de voxels lors d'une comparaison entre <i>Ref</i> et <i>Seg</i>	81
4.4	Régression linéaire de V_{Seg} par rapport à V_{Ref}	82
4.5	Régression linéaire de $\log V_{Seg}$ par rapport à $\log V_{Ref}$	83
4.6	Coupes illustratives du premier sujet <i>outlier</i>	84
4.7	Coupes illustratives du second sujet <i>outlier</i>	84
4.8	Régression linéaire de $\log V_{Seg}$ par rapport à $\log V_{Ref}$ en excluant les sujets <i>outliers</i>	85
4.9	Graphes de Bland-Altman entre V_{Seg} et V_{Ref}	86
4.10	Graphes de Bland-Altman entre $\log V_{Seg}$ et $\log V_{Ref}$	87
4.11	Graphes de Bland-Altman obtenu après transformation réciproque de l'analyse des logarithmes.	87
4.12	Répartition des SI en fonction du volume.	88
4.13	Coupes illustratives d'un sujet ayant un mauvais accord volumique.	90
4.14	Coupes illustratives d'un sujet ayant un mauvais accord volumique.	90
4.15	Analyses par régression de la segmentation avant et après l'étape de croissance.	92
4.16	Graphes de Bland-Altman avant et après l'étape de croissance.	93
4.17	Comparaison des distributions des SI pour les 67 sujets avant et après l'étape de croissance.	94
4.18	Distribution des SI sur les 67 sujets en fonction des variations du paramètre de contraste	96
4.19	Distribution des SI sur les 67 sujets en fonction des variations du seuil sur les intensités	96
5.1	Motif de voisinage obtenu par la discrétisation d'une sphère de rayon 8 mm sur une grille de taille $0,9375 \times 0,9375 \times 5,5 \text{ mm}^3$	106
5.2	Construction des quatre vecteurs de caractéristiques (VC) évalués.	107
5.3	Principe de la classification par k plus proches voisins (kNN).	108

5.4	Schéma de principe de la classification par machines à vecteurs de support (SVM).	110
5.5	Procédure de recherche des hyperparamètres du SVM.	112
5.6	Régression linéaire de V_{Seg} par rapport à V_{Ref} pour les méthodes non supervisées.	114
5.7	Régression linéaire de $\log V_{Seg}$ par rapport à $\log V_{Ref}$ pour les méthodes non supervisées.	115
5.8	Graphe de Bland-Altman entre V_{Seg} et V_{Ref} pour les méthodes non supervisées.	117
5.9	Graphes de Bland-Altman obtenus après transformation réciproque de l'analyse des logarithmes pour les méthodes non supervisées.	118
5.10	Comparaison des distributions des SI des 67 sujets pour les différentes méthodes non supervisées.	119
5.11	Distribution des SI pour les différents vecteurs de caractéristiques (VC) et les différentes bases d'apprentissage.	122
5.12	Analyses par régression linéaire pour les résultats du kNN, SVM et WHASA sur la base de test Base 1.	124
5.13	Analyses par régression linéaire pour les résultats du kNN, SVM et WHASA sur la base de test Base 2.	125
5.14	Analyses par régression linéaire pour les résultats du kNN, SVM et WHASA sur la base de test Base 3.	126
5.15	Graphes de Bland-Altman pour les résultats du KNN, SVM et WHASA sur la base de test Base 1.	128
5.16	Graphes de Bland-Altman pour les résultats du KNN, SVM et WHASA sur la base de test Base 2.	129
5.17	Graphes de Bland-Altman pour les résultats du KNN, SVM et WHASA sur la base de test Base 3.	130
5.18	Comparaison de la distribution des SI pour le KNN, le SVM et WHASA sur les différentes bases de test.	132
6.1	Volume WHASA obtenu automatiquement en fonction du Score PVH attribué pour les 59 sujets de l'étude.	140
6.2	Coupes illustratives du sujet pour lequel le volume WHASA et le Score PVH sont discordants.	140
6.3	Volume WHASA obtenu automatiquement en fonction du Score DWMH attribué pour les 59 sujets de l'étude.	141

6.4	Score visuel des sujets en fonction de leur âge.	142
6.5	Volume des HSB obtenu par WHASA en fonction de l'âge des sujets.	143
6.6	Coupes illustratives du sujet présentant un volume anormale- ment élevé pour son âge.	144
6.7	Illustration d'une acquisition dans le plan bicalleux.	147
6.8	Page d'accueil de la suite logicielle Brainvisa.	149
6.9	Utilisation de WHASA dans Brainvisa.	150
6.10	Visualisation des résultats de WHASA dans Brainvisa.	150
6.11	Exemple de sujet noté 0 (Echec de la segmentation)	152
6.12	Exemple de sujet noté 2 (Segmentation moyenne)	152
6.13	Exemple de sujet noté 3 (Segmentation bonne)	153
6.14	Exemple de sujet noté 4 (Segmentation excellente)	153
B.1	Mouvement de précession du spin autour de l'axe du champ magnétique.	172
B.2	Origine du vecteur de magnétisation résultant.	173
B.3	Signal mesuré après une onde de radiofréquence de 90°.	174
B.4	Principe de la pondération T2.	176
B.5	Principe de la pondération T1.	177
B.6	Inversion-récupération	177

Liste des tableaux

1.1	Synthèse des critères diagnostiques de maladie d'Alzheimer. .	9
1.2	Synthèse des critères diagnostiques de démence vasculaire. . .	10
2.1	Indice de similarité moyen entre la segmentation obtenue par kNN et une segmentation de référence en fonction des contrastes IRM utilisés	37
2.2	Synthèse des méthodes de quantification utilisées pour les principales cohortes.	39
4.1	Principaux paramètres d'acquisition pour les différents centres des jeux de données <i>Hippocampe</i> et <i>Cadasil</i>	74
4.2	Répartition des volumes manuels de référence dans les deux jeux de données.	77
5.1	Résultats de la classification par kNN et SVM pour les différents vecteurs de caractéristiques et les différentes bases d'apprentissage.	121
5.2	Estimation des temps de calcul pour les différentes méthodes.	134
6.1	Description des séquences du protocole CATiv1.	147
6.2	Paramètres d'acquisition pour la séquence T1 de la cohorte MEMENTO	148
6.3	Paramètres d'acquisition pour la séquence FLAIR de la cohorte MEMENTO	148
6.4	Estimation visuelle de la qualité de la segmentation sur la phase pilote et la phase de déploiement de la cohorte MEMENTO	154
B.1	Estimation des constantes de temps T1 et T2 pour certains types de tissus cérébraux à 1.5T.	175

Bibliographie

3C Study Group (2003). Vascular factors and risk of dementia : design of the three-city study and baseline characteristics of the study population. *Neuroepidemiology*, 22(6), 316–325. PMID : 14598854.

Admiraal-Behloul, F., van den Heuvel, D. M. J., Olofsen, H., van Osch, M. J. P., van der Grond, J., van Buchem, M. A., & Reiber, J. H. C. (2005). Fully automatic segmentation of white matter hyperintensities in MR images of the elderly. *NeuroImage*, 28(3), 607–617. PMID : 16129626.

Aizenstein, H. J., Nebes, R. D., Saxton, J. A., Price, J. C., Mathis, C. A., Tsopelas, N. D., Ziolkowski, S. K., James, J. A., Snitz, B. E., Houck, P. R., Bi, W., Cohen, A. D., Lopresti, B. J., DeKosky, S. T., Halligan, E. M., & Klunk, W. E. (2008). Frequent amyloid deposition without significant cognitive impairment among the elderly. *Archives of Neurology*, 65(11), 1509–1517. PMID : 19001171.

Aizerman, A., Braverman, E., & Rozoner, L. (1964). Theoretical foundations of the potential function method in pattern recognition learning. *Automation and Remote Control*, 25, 821–837.

Alvarez, L., Guichard, F., Lions, P.-L., & Morel, J.-M. (1993). Axioms and fundamental equations of image processing. *Archive for Rational Mechanics and Analysis*, 123(3), 199–257.

Anbeek, P., Vincken, K. L., van Bochove, G. S., van Osch, M. J. P., & van der Grond, J. (2005). Probabilistic segmentation of brain tissue in MR imaging. *NeuroImage*, 27(4), 795–804. PMID : 16019235.

Anbeek, P., Vincken, K. L., van Osch, M. J. P., Bisschops, R. H. C., & van der Grond, J. (2004). Probabilistic segmentation of white matter lesions in MR imaging. *NeuroImage*, 21(3), 1037–1044. PMID : 15006671.

- Aronszajn, N. (1950). Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68(3), 337–404.
- Ashburner, J. & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, 26(3), 839–851. PMID : 15955494.
- Awad, I. A., Johnson, P. C., Spetzler, R. F., & Hodak, J. A. (1986). Incidental subcortical lesions identified on magnetic resonance imaging in the elderly. II. postmortem pathological correlations. *Stroke; a journal of cerebral circulation*, 17(6), 1090–1097. PMID : 3810706.
- Barnes, S. R. S., Haacke, E. M., Ayaz, M., Boikov, A. S., Kirsch, W., & Kido, D. (2011). Semiautomated detection of cerebral microbleeds in magnetic resonance images. *Magnetic resonance imaging*, 29(6), 844–852. PMID : 21571479.
- Behl, P., Bocti, C., Swartz, R. H., Gao, F., Sahlas, D. J., Lanctot, K. L., Streiner, D. L., & Black, S. E. (2007). Strategic subcortical hyperintensities in cholinergic pathways and executive function decline in treated alzheimer patients. *Archives of Neurology*, 64(2), 266–272. PMID : 17296844.
- Beucher, S. & Lantuejoul, C. (1979). Use of watersheds in contour detection. In *International Workshop on Image Processing : Real-time Edge and Motion Detection/Estimation* Rennes, France.
- Black, M. J., Sapiro, G., Marimont, D. H., & Heeger, D. (1998). Robust anisotropic diffusion. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 7(3), 421–432. PMID : 18276262.
- Bland, J. M. & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1(8476), 307–310. PMID : 2868172.
- Bland, J. M. & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical methods in medical research*, 8(2), 135–160. PMID : 10501650.
- Bocti, C., Swartz, R. H., Gao, F.-Q., Sahlas, D. J., Behl, P., & Black, S. E. (2005). A new visual rating scale to assess strategic white matter hyperintensities within cholinergic pathways in dementia. *Stroke*, 36(10), 2126–2131. PMID : 16179569.

- Bohnen, N. I., Müller, M. L., Kuwabara, H., Constantine, G. M., & Studdenski, S. A. (2009). Age-associated leukoariosis and cortical cholinergic deafferentation. *Neurology*, 72(16), 1411–1416. PMID : 19380700 PMCID : 2677506.
- Bousser, M. G. & Tournier-Lasserre, E. (1994). Summary of the proceedings of the first international workshop on CADASIL. Paris, may 19-21, 1993. *Stroke*, 25(3), 704–707. PMID : 7907448.
- Braak, H. & Braak, E. (1995). Staging of alzheimer's disease-related neurofibrillary changes. *Neurobiology of aging*, 16(3), 271–278 ; discussion 278–284. PMID : 7566337.
- Brayne, Matthews, Xuereb, Broome, McKenzie, Rossi, Ince, McKeith, Lowe, Esiri, & Morris (2001). Pathological correlates of late-onset dementia in a multicentre, community-based population in england and wales. neuropathology group of the medical research council cognitive function and ageing study (MRC CFAS). *Lancet*, 357(9251), 169–175. PMID : 11213093.
- Brion, J.-P., Passareiro, H., Nunez, J., & Flament-Durand, J. (1985). Mise en évidence immunologique de la protéine tau au niveau des lésions de dégénérescence neurofibrillaire de la maladie d'Alzheimer. *Archives de biologie*, 96(2), 229–235.
- Bronge, L., Bogdanovic, N., & Wahlund, L.-O. (2002). Postmortem MRI and histopathology of white matter changes in alzheimer brains. a quantitative, comparative study. *Dementia and geriatric cognitive disorders*, 13(4), 205–212. PMID : 12006730.
- Burzynska, A., Preuschhof, C., Bäckman, L., Nyberg, L., Li, S.-C., Lindenberg, U., & Heekeren, H. (2010). Age-related differences in white matter microstructure : Region-specific patterns of diffusivity. *NeuroImage*, 49(3), 2104–2112.
- Chabriat, H., Joutel, A., Dichgans, M., Tournier-Lasserre, E., & Bousser, M.-G. (2009). Cadasil. *Lancet Neurology*, 8(7), 643–653. PMID : 19539236.
- Chang, C.-C. & Lin, C.-J. (2011). LIBSVM : a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 27 :1–27 :27.

- Chimowitz, M. I., Estes, M. L., Furlan, A. J., & Awad, I. A. (1992). Further observations on the pathology of subcortical lesions identified on magnetic resonance imaging. *Archives of neurology*, 49(7), 747–752. PMID : 1497503.
- Chupin, M., Gérardin, E., Cuingnet, R., Boutet, C., Lemieux, L., Lehericy, S., Benali, H., Garnero, L., & Colliot, O. (2009). Fully automatic hippocampus segmentation and classification in alzheimer's disease and mild cognitive impairment applied on data from ADNI. *Hippocampus*, 19(6), 579–587. PMID : 19437497.
- Colliot, O., Chételat, G., Chupin, M., Desgranges, B., Magnin, B., Benali, H., Dubois, B., Garnero, L., Eustache, F., & Lehericy, S. (2008). Discrimination between alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus. *Radiology*, 248(1), 194–201. PMID : 18458242.
- Cordonnier, C., Al-Shahi Salman, R., & Wardlaw, J. (2007). Spontaneous brain microbleeds : systematic review, subgroup analyses and standards for study design and reporting. *Brain*, 130(Pt 8), 1988–2003. PMID : 17322562.
- Cordonnier, C., Potter, G. M., Jackson, C. A., Doubal, F., Keir, S., Sudlow, C. L. M., Wardlaw, J. M., & Al-Shahi Salman, R. (2009). Improving interrater agreement about brain microbleeds : development of the brain observer MicroBleed scale (BOMBS). *Stroke*, 40(1), 94–99. PMID : 19008468.
- Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
- Cuingnet, R. (2011). *Contributions à l'apprentissage automatique pour l'analyse d'images cérébrales anatomiques*. Thèse de doctorat, Université Paris Sud - Paris XI.
- Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehericy, S., Habert, M.-O., Chupin, M., Benali, H., & Colliot, O. (2011). Automatic classification of patients with alzheimer's disease from structural MRI : a comparison of ten methods using the ADNI database. *NeuroImage*, 56(2), 766–781. PMID : 20542124.

- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage*, 9(2), 179–194. PMID : 9931268.
- Davatzikos, C., Fan, Y., Wu, X., Shen, D., & Resnick, S. M. (2008). Detection of Prodromal Alzheimer's Disease via Pattern Classification of MRI. *Neurobiology of Aging*, 29(4), 514–523. PMID : 17174012 PMCID : PMC2323584.
- de Boer, R., Vrooman, H. A., van der Lijn, F., Vernooij, M. W., Ikram, M. A., van der Lugt, A., Breteler, M. M. B., & Niessen, W. J. (2009). White matter lesion extension to automatic brain tissue segmentation on MRI. *NeuroImage*, 45(4), 1151–1161. PMID : 19344687.
- DeCarli, C., Maisog, J., Murphy, D. G., Teichberg, D., Rapoport, S. I., & Horwitz, B. (1992). Method for quantification of brain, ventricular, and subarachnoid CSF volumes from MR images. *Journal of Computer Assisted Tomography*, 16(2), 274–284. PMID : 1545026.
- DeCarli, C., Miller, B. L., Swan, G. E., Reed, T., Wolf, P. A., Garner, J., Jack, L., & Carmelli, D. (1999). Predictors of brain morphology for the men of the NHLBI twin study. *Stroke*, 30(3), 529–536.
- Delacourte, A. (2003). Le retour de la protéine tau. *La Recherche*, Hors Série n° 10, 44–48.
- Delacourte, A., David, J. P., Sergeant, N., Buée, L., Wattez, A., Vermersch, P., Ghozali, F., Fallet-Bianco, C., Pasquier, F., Lebert, F., Petit, H., & Di Menza, C. (1999). The biochemical pathway of neurofibrillary degeneration in aging and alzheimer's disease. *Neurology*, 52(6), 1158–1165. PMID : 10214737.
- Dessi, F., Colle, M.-A., Grignon, Y., Hauw, J.-J., & Duyckaerts, C. (1997). Maladie d'Alzheimer : Neuropathologie. *Médecine thérapeutique*, 3(5), 353–363.
- Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3), 297–302.
- Dickerson, B. C., Bakkour, A., Salat, D. H., Feczko, E., Pacheco, J., Greve, D. N., Grodstein, F., Wright, C. I., Blacker, D., Rosas, H. D., Sperling, R. A.,

- Atri, A., Growdon, J. H., Hyman, B. T., Morris, J. C., Fischl, B., & Buckner, R. L. (2009). The cortical signature of alzheimer's disease : regionally specific cortical thinning relates to symptom severity in very mild to mild AD dementia and is detectable in asymptomatic amyloid-positive individuals. *Cerebral cortex*, 19(3), 497–510. PMID : 18632739.
- Divry, P. (1927). Etude histochemique des plaques seniles. *Journal Belge de Neurologie et de Psychiatrie*, 27, 643–657.
- Duarte-Carvajalino, J. M., Castillo, P. E., & Velez-Reyes, M. (2007). Comparative study of semi-implicit schemes for nonlinear diffusion in hyper-spectral imagery. *IEEE Transactions on Image Processing*, 16(5), 1303–1314. PMID : 17491461.
- Dubois, B., Feldman, H. H., Jacova, C., Cummings, J. L., Dekosky, S. T., Barberger-Gateau, P., Delacourte, A., Frisoni, G., Fox, N. C., Galasko, D., Gauthier, S., Hampel, H., Jicha, G. A., Meguro, K., O'Brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Sarazin, M., de Souza, L. C., Stern, Y., Visser, P. J., & Scheltens, P. (2010). Revising the definition of Alzheimer's disease : a new lexicon. *Lancet Neurology*, 9(11), 1118–1127. PMID : 20934914.
- Dubois, B., Feldman, H. H., Jacova, C., Dekosky, S. T., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G., Meguro, K., O'Brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Stern, Y., Visser, P. J., & Scheltens, P. (2007). Research criteria for the diagnosis of Alzheimer's disease : revising the NINCDS-ADRDA criteria. *Lancet Neurology*, 6(8), 734–746. PMID : 17616482.
- Dubois, B., Sarazin, M., Lehericy, S., Chupin, M., Garnero, L., Tonelli, I., Dormot, D., & Group, D. H. S. (2009). Effects of Donepezil on structural MRI and clinical markers in patients with amnesic mild cognitive impairment : a randomized, placebo controlled trial. In *2nd Clinical Trials Conference on Alzheimer's Disease*, Las Vegas.
- Dufouil, C., de Kersaint-Gilly, A., Besançon, V., Levy, C., Auffray, E., Brunereau, L., Alperovitch, A., & Tzourio, C. (2001). Longitudinal study of blood pressure and white matter hyperintensities : the EVA MRI cohort. *Neurology*, 56(7), 921–926. PMID : 11294930.

- Durgerian, S., Lancaster, M., Seidenberg, M., Nielson, K., Woodard, J., Carson Smith, J., Matthews, M., Butts, A., Hantke, N., & Rao, S. (2011). White matter diffusivity is higher among healthy elders at increased risk for alzheimer's disease. In *Proceedings of the 17th International Conference on Functional Mapping of the Human Brain*, Québec City.
- Dyrby, T. B., Rostrup, E., Baaré, W. F. C., van Straaten, E. C. W., Barkhof, F., Vrenken, H., Ropele, S., Schmidt, R., Erkinjuntti, T., Wahlund, L.-O., Pantoni, L., Inzitari, D., Paulson, O. B., Hansen, L. K., & Waldemar, G. (2008). Segmentation of age-related white matter changes in a clinical multi-center study. *NeuroImage*, 41(2), 335–345. PMID : 18394928.
- Erkinjuntti, T., Gao, F., Lee, D. H., Eliasziw, M., Merskey, H., & Hachinski, V. C. (1994). Lack of difference in brain hyperintensities between patients with early Alzheimer's disease and control subjects. *Archives of Neurology*, 51(3), 260–268. PMID : 8129637.
- Esiri, M. M., Nagy, Z., Smith, M. Z., Barnettson, L., & Smith, A. D. (1999). Cerebrovascular disease and threshold for dementia in the early stages of Alzheimer's disease. *Lancet*, 354(9182), 919–920. PMID : 10489957.
- Fazekas, F., Chawluk, J. B., Alavi, A., Hurtig, H. I., & Zimmerman, R. A. (1987). MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. *AJR. American journal of roentgenology*, 149(2), 351–356. PMID : 3496763.
- Fazekas, F., Kleinert, R., Offenbacher, H., Payer, F., Schmidt, R., Kleinert, G., Radner, H., & Lechner, H. (1991). The morphologic correlate of incidental punctate white matter hyperintensities on MR images. *AJNR. American journal of neuroradiology*, 12(5), 915–921. PMID : 1950921.
- Fazekas, F., Kleinert, R., Offenbacher, H., Schmidt, R., Kleinert, G., Payer, F., Radner, H., & Lechner, H. (1993). Pathologic correlates of incidental MRI white matter signal hyperintensities. *Neurology*, 43(9), 1683–1689. PMID : 8414012.
- Fazekas, F., Ropele, S., Enzinger, C., Gorani, F., Seewann, A., Petrovic, K., & Schmidt, R. (2005). MTI of white matter hyperintensities. *Brain : A Journal of Neurology*, 128(Pt 12), 2926–2932. PMID : 15958507.

- Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., & Dale, A. M. (2002). Whole brain segmentation : automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3), 341–355. PMID : 11832223.
- Fischl, B., Salat, D. H., van der Kouwe, A. J. W., Makris, N., Ségonne, F., Quinn, B. T., & Dale, A. M. (2004). Sequence-independent segmentation of magnetic resonance images. *NeuroImage*, 23 Suppl 1, S69–84. PMID : 15501102.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II : inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2), 195–207. PMID : 9931269.
- Fornage, M., Debette, S., Bis, J. C., Schmidt, H., Ikram, M. A., Dufouil, C., Sigurdsson, S., Lumley, T., DeStefano, A. L., Fazekas, F., Vrooman, H. A., Shibata, D. K., Maillard, P., Zijdenbos, A., Smith, A. V., Gudnason, H., de Boer, R., Cushman, M., Mazoyer, B., Heiss, G., Vernooij, M. W., Enzinger, C., Glazer, N. L., Beiser, A., Knopman, D. S., Cavalieri, M., Niessen, W. J., Harris, T. B., Petrovic, K., Lopez, O. L., Au, R., Lambert, J.-C., Hofman, A., Gottesman, R. F., Garcia, M., Heckbert, S. R., Atwood, L. D., Catellier, D. J., Uitterlinden, A. G., Yang, Q., Smith, N. L., Aspelund, T., Romero, J. R., Rice, K., Taylor, K. D., Nalls, M. A., Rotter, J. I., Sharrett, R., van Duijn, C. M., Amouyel, P., Wolf, P. A., Gudnason, V., van der Lugt, A., Boerwinkle, E., Psaty, B. M., Seshadri, S., Tzourio, C., Breteler, M. M. B., Mosley, T. H., Schmidt, R., Longstreth, W. T., DeCarli, C., & Launer, L. J. (2011). Genome-wide association studies of cerebral white matter lesion burden : the CHARGE consortium. *Annals of neurology*, 69(6), 928–939. PMID : 21681796.
- Fried, L. P., Borhani, N. O., Enright, P., Furberg, C. D., Gardin, J. M., Kronmal, R. A., Kuller, L. H., Manolio, T. A., Mittelmark, M. B., & Newman, A. (1991). The cardiovascular health study : design and rationale. *Annals of epidemiology*, 1(3), 263–276. PMID : 1669507.
- Gerard, G. & Weisberg, L. A. (1986). MRI periventricular lesions in adults. *Neurology*, 36(7), 998–1001. PMID : 3714061.

- Gerardin, E., Chételat, G., Chupin, M., Cuingnet, R., Desgranges, B., Kim, H.-S., Niethammer, M., Dubois, B., Lehéricy, S., Garnero, L., Eustache, F., & Colliot, O. (2009). Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. *NeuroImage*, 47(4), 1476–1486. PMID : 19463957.
- Gibson, E., Gao, F., Black, S. E., & Lobaugh, N. J. (2010). Automatic segmentation of white matter hyperintensities in the elderly using FLAIR images at 3T. *Journal of magnetic resonance imaging : JMRI*, 31(6), 1311–1322. PMID : 20512882.
- Gicquel, S. (1998). *Détection, quantification et localisation automatique des hypersignaux de la substance blanche en imagerie par résonance magnétique*. Thèse de doctorat, Université de Caen, Caen, France.
- Glenner, G. G. & Wong, C. W. (1984). Alzheimer's disease : initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochemical and biophysical research communications*, 120(3), 885–890. PMID : 6375662.
- Gouw, A. A., Seewann, A., van der Flier, W. M., Barkhof, F., Rozemuller, A. M., Scheltens, P., & Geurts, J. J. G. (2011). Heterogeneity of small vessel disease : a systematic review of MRI and histopathology correlations. *Journal of Neurology, Neurosurgery, and Psychiatry*, 82(2), 126–135. PMID : 20935330.
- Gouw, A. A., Seewann, A., Vrenken, H., van der Flier, W. M., Rozemuller, J. M., Barkhof, F., Scheltens, P., & Geurts, J. J. G. (2008). Heterogeneity of white matter hyperintensities in alzheimer's disease : post-mortem quantitative MRI and neuropathology. *Brain*, 131(12), 3286–3298.
- Greenberg, S. M., Vernooij, M. W., Cordonnier, C., Viswanathan, A., Al-Shahi Salman, R., Warach, S., Launer, L. J., Van Buchem, M. A., & Breteler, M. M. (2009). Cerebral microbleeds : a guide to detection and interpretation. *Lancet Neurology*, 8(2), 165–174. PMID : 19161908.
- Gregoire, S. M., Chaudhary, U. J., Brown, M. M., Yousry, T. A., Kallis, C., Jäger, H. R., & Werring, D. J. (2009). The Microbleed Anatomical Rating Scale (MARS) Reliability of a tool to map brain microbleeds. *Neurology*, 73(21), 1759–1766.

- Grinberg, L. T. & Thal, D. R. (2010). Vascular pathology in the aged human brain. *Acta neuropathologica*, 119(3), 277–290. PMID : 20155424.
- Hachinski, V. C., Potter, P., & Merskey, H. (1986). Leuko-araiosis : an ancient term for a new problem. *The Canadian journal of neurological sciences. Le journal canadien des sciences neurologiques*, 13(4 Suppl), 533–534. PMID : 3791068.
- Harris, T. B., Launer, L. J., Eiriksdottir, G., Kjartansson, O., Jonsson, P. V., Sigurdsson, G., Thorgeirsson, G., Aspelund, T., Garcia, M. E., Cotch, M. F., Hoffman, H. J., & Gudnason, V. (2007). Age, Gene/Environment susceptibility-reykjavik study : multidisciplinary applied phenomics. *American journal of epidemiology*, 165(9), 1076–1087. PMID : 17351290.
- Hofman, A., Breteler, M. M. B., van Duijn, C. M., Janssen, H. L. A., Krestin, G. P., Kuipers, E. J., Stricker, B. H. C., Tiemeier, H., Uitterlinden, A. G., Vingerling, J. R., & Witteman, J. C. M. (2009). The rotterdam study : 2010 objectives and design update. *European journal of epidemiology*, 24(9), 553–572. PMID : 19728115.
- investigators, A. (1989). The atherosclerosis risk in communities (ARIC) study : design and objectives. the ARIC investigators. *American journal of epidemiology*, 129(4), 687–702. PMID : 2646917.
- Jack, C R, J., O'Brien, P. C., Rettman, D. W., Shiung, M. M., Xu, Y., Muthupillai, R., Manduca, A., Avula, R., & Erickson, B. J. (2001). FLAIR histogram segmentation for measurement of leukoaraiosis volume. *Journal of Magnetic Resonance Imaging : JMRI*, 14(6), 668–676. PMID : 11747022.
- Jack, C R, J., Petersen, R. C., Xu, Y., O'Brien, P. C., Smith, G. E., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1998). Rate of medial temporal lobe atrophy in typical aging and Alzheimer's disease. *Neurology*, 51(4), 993–999. PMID : 9781519.
- Jack, Clifford R, J., Albert, M. S., Knopman, D. S., McKhann, G. M., Sperling, R. A., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). Introduction to the recommendations from the National Institute on Aging - Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia*, 7(3), 257–262. PMID : 21514247.

- Jeerakathil, T., Wolf, P. A., Beiser, A., Massaro, J., Seshadri, S., D'Agostino, R. B., & DeCarli, C. (2004). Stroke risk profile predicts white matter hyperintensity volume : the Framingham study. *Stroke*, 35(8), 1857–1861. PMID : 15218158.
- Kannel, W. B., Feinleib, M., McNamara, P. M., Garrison, R. J., & Castelli, W. P. (1979). An investigation of coronary heart disease in families the framingham offspring study. *American Journal of Epidemiology*, 110(3), 281–290. PMID : 474565.
- Keverne, J. S., Low, W. C. R., Ziabreva, I., Court, J. A., Oakley, A. E., & Kalaria, R. N. (2007). Cholinergic neuronal deficits in CADASIL. *Stroke*, 38(1), 188–191. PMID : 17122431.
- Khayati, R., Vafadust, M., Towhidkhah, F., & Nabavi, M. (2008). Fully automatic segmentation of multiple sclerosis lesions in brain MR FLAIR images using adaptive mixtures method and markov random field model. *Computers in biology and medicine*, 38(3), 379–390. PMID : 18262511.
- Kidd, M. (1963). Paired helical filaments in electron microscopy of Alzheimer's disease. *Nature*, 197, 192–193. PMID : 14032480.
- Klöppel, S., Abdulkadir, A., Hadjidemetriou, S., Issleib, S., Frings, L., Thanh, T. N., Mader, I., Teipel, S. J., Hüll, M., & Ronneberger, O. (2011). A comparison of different automated methods for the detection of white matter lesions in MRI data. *NeuroImage*, 57(2), 416–422. PMID : 21569857.
- Klöppel, S., Stonnington, C. M., Chu, C., Draganski, B., Scahill, R. I., Rohrer, J. D., Fox, N. C., Jack, Clifford R, J., Ashburner, J., & Frackowiak, R. S. J. (2008). Automatic classification of MR scans in Alzheimer's disease. *Brain : a journal of neurology*, 131(Pt 3), 681–689. PMID : 18202106.
- Lammie, G. A. (2002). Hypertensive cerebral small vessel disease and stroke. *Brain pathology*, 12(3), 358–370. PMID : 12146804.
- Lao, Z., Shen, D., Liu, D., Jawad, A. F., Melhem, E. R., Launer, L. J., Bryan, R. N., & Davatzikos, C. (2008). Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine. *Academic Radiology*, 15(3), 300–313. PMID : 18280928.

- Levy-Cooperman, N., Ramirez, J., Lobaugh, N. J., & Black, S. E. (2008). Misclassified tissue volumes in alzheimer disease patients with white matter hyperintensities : importance of lesion segmentation procedures for volumetric analysis. *Stroke*, 39(4), 1134–1141. PMID : 18323507.
- Liang, Z.-P. & Lauterbur, P. C. (2000). *Principles of Magnetic Resonance Imaging : A Signal Processing Perspective*. IEEE Press series in biomedical engineering. New York : IEEE Press.
- Liu, T., Surapaneni, K., Lou, M., Cheng, L., Spincemaille, P., & Wang, Y. (2012). Cerebral microbleeds : burden assessment by using quantitative susceptibility mapping. *Radiology*, 262(1), 269–278. PMID : 22056688.
- Longstreth, W. T., Manolio, T. A., Arnold, A., Burke, G. L., Bryan, N., Jungreis, C. A., Enright, P. L., O’Leary, D., & Fried, L. (1996). Clinical correlates of white matter findings on cranial magnetic resonance imaging of 3301 elderly people. the Cardiovascular Health Study. *Stroke*, 27(8), 1274–1282. PMID : 8711786.
- Madhuranthakam, A. J., Sarkar, S. N., Busse, R. F., Bakshi, R., & Alsop, D. C. (2011). Optimized double inversion recovery for reduction of t(1) weighting in fluid-attenuated inversion recovery. *Magnetic Resonance in Medicine : Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*. PMID : 21590725.
- Maillard, P., Delcroix, N., Crivello, F., Dufouil, C., Gicquel, S., Joliot, M., Tzourio-Mazoyer, N., Alperovitch, A., Tzourio, C., & Mazoyer, B. (2008). An automated procedure for the assessment of white matter hyperintensities by multispectral (T1, T2, PD) MRI and an evaluation of its between-centre reproducibility based on two large community databases. *Neuroradiology*, 50(1), 31–42. PMID : 17938898.
- Manolio, T. A., Kronmal, R. A., Burke, G. L., Poirier, V., O’Leary, D. H., Gardin, J. M., Fried, L. P., Steinberg, E. P., & Bryan, R. N. (1994). Magnetic resonance abnormalities and cardiovascular disease in older adults. the Cardiovascular Health Study. *Stroke*, 25(2), 318–327. PMID : 8303738.
- Marc, G. & Dubas, F. (2005). Microangiopathies cérébrales. *Psychologie & neuropsychiatrie du vieillissement*, 3(2), 83–88. PMID : 15929921.

- Marshall, V. G., Bradley, W. G. J., Marshall, C. E., Bhoopat, T., & Rhodes, R. H. (1988). Deep white matter infarction : correlation of MR imaging and histopathologic findings. *Radiology*, 167(2), 517–522. PMID : 3357964.
- McDonald, C. R., McEvoy, L. K., Gharapetian, L., Fennema-Notestine, C., Hagler, D. J., Holland, D., Koyama, A., Brewer, J. B., & Dale, A. M. (2009). Regional rates of neocortical atrophy from normal aging to early Alzheimer disease. *Neurology*, 73(6), 457–465. PMID : 19667321 PMCID : PMC2727145.
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, Clifford R. J., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer’s disease : recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease. *Alzheimer’s & dementia*, 7(3), 263–269. PMID : 21514250.
- McRobbie, D. W., Moore, E. A., Graves, M. J., & Prince, M. R. (2007). *MRI from Picture to Proton*. Cambridge University Press, 2nd edition.
- Mikula, K. & Ramarosy, N. (2001). Semi-implicit finite volume scheme for solving nonlinear diffusion equations in image processing. *Numerische Mathematik*, 89(3), 561–590.
- Mirsan, T. R., Lee, D. H., Wong, C. J., Diaz, J. F., Fox, A. J., Hachinski, V. C., & Merskey, H. (1991). Clinical correlates of white-matter changes on magnetic resonance imaging scans of the brain. *Archives of Neurology*, 48(10), 1015–1021. PMID : 1929891.
- Mäntylä, R., Erkinjuntti, T., Salonen, O., Aronen, H. J., Peltonen, T., Pohjasvaara, T., & Standertskjöld-Nordenstam, C. G. (1997). Variable agreement between visual rating scales for white matter hyperintensities on MRI. comparison of 13 rating scales in a poststroke cohort. *Stroke*, 28(8), 1614–1623. PMID : 9259759.
- Moody, D. M., Thore, C. R., Anstrom, J. A., Challa, V. R., Langefeld, C. D., & Brown, W. R. (2004). Quantification of afferent vessels shows reduced

- brain vascular density in subjects with leukoaraiosis¹. *Radiology*, 233(3), 883–890.
- Pantoni, L., Basile, A. M., Pracucci, G., Asplund, K., Bogousslavsky, J., Chabriat, H., Erkinjuntti, T., Fazekas, F., Ferro, J. M., Hennerici, M., O'Brien, J., Scheltens, P., Visser, M. C., Wahlund, L.-O., Waldemar, G., Wallin, A., & Inzitari, D. (2005). Impact of age-related cerebral white matter changes on the transition to disability – the LADIS study : rationale, design and methodology. *Neuroepidemiology*, 24(1-2), 51–62. PMID : 15459510.
- Pantoni, L. & Garcia, J. H. (1997). Pathogenesis of leukoaraiosis : a review. *Stroke*, 28(3), 652–659. PMID : 9056627.
- Perona, P. & Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7), 629–639.
- Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, 256(3), 183–194. PMID : 15324362.
- Petrovitch, H., Ross, G. W., Steinhorn, S. C., Abbott, R. D., Markesbery, W., Davis, D., Nelson, J., Hardman, J., Masaki, K., Vogt, M. R., Launer, L., & White, L. R. (2005). AD lesions and infarcts in demented and non-demented japanese-american men. *Annals of Neurology*, 57(1), 98–103. PMID : 15562458.
- Poels, M. M. F., Vernooij, M. W., Ikram, M. A., Hofman, A., Krestin, G. P., van der Lugt, A., & Breteler, M. M. B. (2010). Prevalence and risk factors of cerebral microbleeds : an update of the Rotterdam scan study. *Stroke*, 41(10 Suppl), S103–106. PMID : 20876479.
- Poggesi, A., Gouw, A., van der Flier, W., Pracucci, G., Chabriat, H., Erkinjuntti, T., Fazekas, F., Ferro, J. M., Hennerici, M., Langhorne, P., O'Brien, J. T., Visser, M. C., Wahlund, L.-O., Waldemar, G., Wallin, A., Scheltens, P., Inzitari, D., & Pantoni, L. (2012). Cerebral white matter changes are associated with abnormalities on neurological examination in non-disabled elderly : the LADIS study. *Journal of Neurology*. PMID : 23180181.
- Poggesi, A., Pantoni, L., Inzitari, D., Fazekas, F., Ferro, J., O'Brien, J., Hennerici, M., Scheltens, P., Erkinjuntti, T., Visser, M., Langhorne, P., Chabriat,

- H., Waldemar, G., Wallin, A., & Wahlund, A. (2011). 2001-2011 : A decade of the LADIS (leukoaraiosis and DISability) study : What have we learned about white matter changes and small-vessel disease? *Cerebrovascular Diseases*, 32(6), 577–588. PMID : 22277351.
- Provenzano, F. A., Muraskin, J., Tosto, G., Narkhede, A., Wasserman, B. T., Griffith, E. Y., Guzman, V. A., Meier, I. B., Zimmerman, M. E., & Brickman, A. M. (2013). White Matter Hyperintensities and Cerebral Amyloidosis : Necessary and Sufficient for Clinical Expression of Alzheimer Disease? *JAMA Neurology*, (pp. 1–7). PMID : 23420027.
- Psaty, B. M., O'Donnell, C. J., Gudnason, V., Lunetta, K. L., Folsom, A. R., Rotter, J. I., Uitterlinden, A. G., Harris, T. B., Witteman, J. C. M., Boerwinkle, E., & CHARGE Consortium (2009). Cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium : Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circulation. Cardiovascular genetics*, 2(1), 73–80. PMID : 20031568.
- Román, G. C., Tatemichi, T. K., Erkinjuntti, T., Cummings, J. L., Masdeu, J. C., Garcia, J. H., Amaducci, L., Orgogozo, J. M., Brun, A., & Hofman, A. (1993). Vascular dementia : diagnostic criteria for research studies. report of the NINDS-AIREN international workshop. *Neurology*, 43(2), 250–260. PMID : 8094895.
- Samaille, T., Fillon, L., Cuingnet, R., Jouvent, E., Chabriat, H., Dormont, D., Colliot, O., & Chupin, M. (2012). Contrast-based fully automatic segmentation of white matter hyperintensities : method and validation. *PloS one*, 7(11), e48953. PMID : 23152828.
- Scheltens, P., Barkhof, F., Leys, D., Pruvo, J. P., Nauta, J. J., Vermersch, P., Steinling, M., & Valk, J. (1993). A semiquantative rating scale for the assessment of signal hyperintensities on magnetic resonance imaging. *Journal of the neurological sciences*, 114(1), 7–12. PMID : 8433101.
- Schmidt, R., Lechner, H., Fazekas, F., Niederkorn, K., Reinhart, B., Grieshofer, P., Horner, S., Offenbacher, H., Koch, M., & Eber, B. (1994). Assessment of cerebrovascular risk profiles in healthy persons : definition of research goals and the austrian stroke prevention study (ASPS). *Neuroepidemiology*, 13(6), 308–313. PMID : 7800110.

- Ségonne, F., Pacheco, J., & Fischl, B. (2007). Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *IEEE Transactions on Medical Imaging*, 26(4), 518–529. PMID : 17427739.
- Shimada, K., Kawamoto, A., Matsubayashi, K., & Ozawa, T. (1990). Silent cerebrovascular disease in the elderly. correlation with ambulatory pressure. *Hypertension*, 16(6), 692–699. PMID : 2246035.
- Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations : uses in assessing rater reliability. *Psychological bulletin*, 86(2), 420–428. PMID : 18839484.
- Sigurdsson, S., Aspelund, T., Forsberg, L., Fredriksson, J., Kjartansson, O., Oskarsdottir, B., Jonsson, P. V., Eiriksdottir, G., Harris, T. B., Zijdenbos, A., van Buchem, M. A., Launer, L. J., & Gudnason, V. (2011). Brain tissue volumes in the general population of the elderly the AGES-Reykjavik study. *NeuroImage*. PMID : 22119006.
- Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE transactions on medical imaging*, 17(1), 87–97. PMID : 9617910.
- Smart, S. D., Firbank, M. J., & O'Brien, J. T. (2011). Validation of automated white matter hyperintensity segmentation. *Journal of Aging Research*, 2011. PMID : 21904678 PMCID : PMC3167190.
- Smith, C. D., Snowdon, D., & Markesbery, W. R. (2000). Periventricular white matter hyperintensities on MRI : correlation with neuropathologic findings. *Journal of neuroimaging : official journal of the American Society of Neuroimaging*, 10(1), 13–16. PMID : 10666976.
- Smith, E. E., Salat, D. H., Jeng, J., McCreary, C. R., Fischl, B., Schmahmann, J. D., Dickerson, B. C., Viswanathan, A., Albert, M. S., Blacker, D., & Greenberg, S. M. (2011). Correlations between MRI white matter lesion location and executive function and episodic memory. *Neurology*, 76(17), 1492–1499. PMID : 21518999.
- Snowdon, D. A., Greiner, L. H., Mortimer, J. A., Riley, K. P., Greiner, P. A., & Markesbery, W. R. (1997). Brain infarction and the clinical expression of Alzheimer disease. the Nun study. *JAMA : the journal of the American Medical Association*, 277(10), 813–817. PMID : 9052711.

- Thal, D. R., Grinberg, L. T., & Attems, J. (2012). Vascular dementia : different forms of vessel disorders contribute to the development of dementia in the elderly brain. *Experimental gerontology*, 47(11), 816–824. PMID : 22705146.
- Thompson, P. M., Hayashi, K. M., de Zubicaray, G., Janke, A. L., Rose, S. E., Semple, J., Herman, D., Hong, M. S., Dittmer, S. S., Doddrell, D. M., & Toga, A. W. (2003). Dynamics of gray matter loss in Alzheimer's disease. *The Journal of neuroscience*, 23(3), 994–1005. PMID : 12574429.
- Tournier-Lasserre, E., Joutel, A., Melki, J., Weissenbach, J., Lathrop, G. M., Chabriat, H., Mas, J. L., Cabanis, E. A., Baudrimont, M., & Maciazek, J. (1993). Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy maps to chromosome 19q12. *Nature Genetics*, 3(3), 256–259. PMID : 8485581.
- Turetschek, K., Wunderbaldinger, P., Bankier, A. A., Zontsich, T., Graf, O., Mallek, R., & Hittmair, K. (1998). Double inversion recovery imaging of the brain : initial experience and comparison with fluid attenuated inversion recovery imaging. *Magnetic Resonance Imaging*, 16(2), 127–135. PMID : 9508269.
- van den Heuvel, D. M. J., ten Dam, V. H., de Craen, A. J. M., Admiraal-Behloul, F., van Es, A. C. G. M., Palm, W. M., Spilt, A., Bollen, E. L. E. M., Blauw, G. J., Launer, L., Westendorp, R. G. J., & van Buchem, M. A. (2006). Measuring longitudinal white matter changes : comparison of a visual rating scale with a volumetric measurement. *AJNR. American journal of neuroradiology*, 27(4), 875–878. PMID : 16611781.
- Van Leemput, K., Maes, F., Vandermeulen, D., Colchester, A., & Suetens, P. (2001). Automated segmentation of multiple sclerosis lesions by model outlier detection. *IEEE transactions on medical imaging*, 20(8), 677–688. PMID : 11513020.
- van Straaten, E. C. W., Fazekas, F., Rostrup, E., Scheltens, P., Schmidt, R., Pantoni, L., Inzitari, D., Waldemar, G., Erkinjuntti, T., Mäntylä, R., Wahlund, L.-O., & Barkhof, F. (2006). Impact of white matter hyperintensities scoring method on correlations with clinical data : the LADIS study. *Stroke ; a Journal of Cerebral Circulation*, 37(3), 836–840. PMID : 16439704.

- van Swieten, J. C., Hijdra, A., Koudstaal, P. J., & van Gijn, J. (1990). Grading white matter lesions on CT and MRI : a simple scale. *Journal of Neurology, Neurosurgery, and Psychiatry*, 53(12), 1080–1083. PMID : 2292703 PMCID : PMC488320.
- van Swieten, J. C., van den Hout, J. H., van Ketel, B. A., Hijdra, A., Wokke, J. H., & van Gijn, J. (1991). Periventricular lesions in the white matter on magnetic resonance imaging in the elderly. a morphometric correlation with arteriolosclerosis and dilated perivascular spaces. *Brain : a journal of neurology*, 114 (Pt 2), 761–774. PMID : 2043948.
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer.
- Vemuri, P., Gunter, J. L., Senjem, M. L., Whitwell, J. L., Kantarci, K., Knopman, D. S., Boeve, B. F., Petersen, R. C., & Jack, Clifford R, J. (2008). Alzheimer's disease diagnosis in individual subjects using structural MR images : validation studies. *NeuroImage*, 39(3), 1186–1197. PMID : 18054253.
- Vernooij, M., de Groot, M., van der Lugt, A., Ikram, M., Krestin, G., Hofman, A., Niessen, W., & Breteler, M. (2008). White matter atrophy and lesion formation explain the loss of structural integrity of white matter in aging. *NeuroImage*, 43(3), 470–477.
- Vernooij, M. W., Ikram, M. A., Tanghe, H. L., Vincent, A. J. P. E., Hofman, A., Krestin, G. P., Niessen, W. J., Breteler, M. M. B., & van der Lugt, A. (2007). Incidental findings on brain MRI in the general population. *The New England Journal of Medicine*, 357(18), 1821–1828. PMID : 17978290.
- Viswanathan, A., Guichard, J.-P., Gschwendtner, A., Buffon, F., Cumurcuic, R., Boutron, C., Vicaud, E., Holtmannspötter, M., Pachai, C., Bousser, M.-G., Dichgans, M., & Chabriat, H. (2006). Blood pressure and haemoglobin a1c are associated with microhaemorrhage in CADASIL : a two-centre cohort study. *Brain*, 129(Pt 9), 2375–2383. PMID : 16844717.
- Wahlund, L. O., Agartz, I., Almqvist, O., Basun, H., Forssell, L., Sääf, J., & Wetterberg, L. (1990). The brain in healthy aged individuals : MR imaging. *Radiology*, 174(3 Pt 1), 675–679. PMID : 2305048.
- Wahlund, L. O., Barkhof, F., Fazekas, F., Bronge, L., Augustin, M., Sjögren, M., Wallin, A., Ader, H., Leys, D., Pantoni, L., Pasquier, F., Erkinjuntti, T.,

- & Scheltens, P. (2001). A new rating scale for age-related white matter changes applicable to MRI and CT. *Stroke*, 32(6), 1318–1322.
- Weickert, J. (1998). *Anisotropic diffusion in image processing*. Stuttgart : B.G. Teubner.
- Weickert, J., Romeny, B. H., & Viergever, M. A. (1998). Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 7(3), 398–410. PMID : 18276260.
- Wen, W. & Sachdev, P. (2004). The topography of white matter hyperintensities on brain MRI in healthy 60- to 64-year-old individuals. *NeuroImage*, 22(1), 144–154. PMID : 15110004.
- World Health Organization & Alzheimer's Disease International (2012). *Dementia a public health priority*. Geneva ; London : World Health Organization ; Alzheimer's Disease International.
- Wu, M., Rosano, C., Butters, M., Whyte, E., Nable, M., Crooks, R., Meltzer, C. C., Reynolds, C. F., & Aizenstein, H. J. (2006). A fully automated method for quantifying and localizing white matter hyperintensities on MR images. *Psychiatry research*, 148(2-3), 133–142. PMID : 17097277 PMCID : 1761950.
- Yoshita, M., Fletcher, E., Harvey, D., Ortega, M., Martinez, O., Mungas, D. M., Reed, B. R., & DeCarli, C. S. (2006). Extent and distribution of white matter hyperintensities in normal aging, MCI, and AD. *Neurology*, 67(12), 2192–2198. PMID : 17190943.
- Young, V. G., Halliday, G. M., & Kril, J. J. (2008). Neuropathologic correlates of white matter hyperintensities. *Neurology*, 71(11), 804–811. PMID : 18685136.
- Zijdenbos, A. P., Forghani, R., & Evans, A. C. (2002). Automatic "pipeline" analysis of 3-D MRI data for clinical trials : application to multiple sclerosis. *IEEE Transactions on Medical Imaging*, 21(10), 1280–1291. PMID : 12585710.

Résumé

Les hyperintensités de la substance blanche (HSB) sont visibles sur les séquences d'imagerie par résonance magnétique (IRM) pondérées en T2. Couramment observées chez les personnes âgées, il est cependant démontré que leur présence en quantité importante est un facteur de risque pour les accidents vasculaires et la démence. Leur impact sur la cognition ou l'effet de HSB plus modérées demeurent néanmoins controversés. A ce jour, la quantification des HSB est principalement réalisée à l'aide d'échelles visuelles alors qu'une segmentation serait plus sensible, permettant une mesure volumétrique, une meilleure caractérisation spatiale et un suivi des lésions dans le temps.

Cette thèse présente une nouvelle chaîne de traitements, nommée WHASA (*White matter Hyperintensities Automatic Segmentation Algorithm*), permettant la segmentation automatique des HSB à partir d'une image T1 et d'une image FLAIR dans un environnement multicentrique. Elle s'appuie sur l'utilisation couplée d'un filtrage par diffusion non linéaire et une segmentation par ligne de partage des eaux. Les régions correspondant aux HSB sont sélectionnées à partir de caractéristiques sur les intensités et la localisation. Les résultats obtenus par WHASA sur 67 sujets provenant de six centres différents sont évalués par rapport à une segmentation manuelle de référence. Les performances sont ensuite comparées avec celles de plusieurs méthodes issues de la littérature. Enfin, WHASA est appliquée sur des données issues d'autres études pour évaluer sa robustesse. Au total, WHASA a été utilisée sur 260 sujets provenant de onze machines IRM, comprenant des imageurs 1,5 T et 3 T, avec des épaisseurs de coupe variant entre 2,5 et 6,2 mm.

Abstract

White Matter Hyperintensities (WMH) may be observed on T2-weighted magnetic resonance imaging (MRI). Although they are commonly seen in elderly people, it has been demonstrated that a high level of WMH is a risk factor for stroke and dementia. The impact of WMH on cognition or the effect of a low level of WMH remains however controversial. Today, WMH quantification is still mainly performed by visual rating, although a segmentation would be more sensitive and would enable volume measurement, spatial distribution analysis and lesion follow-up.

This thesis introduces a new pipeline, named WHASA (*White matter Hyperintensities Automatic Segmentation Algorithm*), to automatically segment WMH from a T1-weighted and a FLAIR image in multicentre studies. WHASA relies on the coupling of non linear diffusion and watershed segmentation. Regions corresponding to WMH are selected based on intensity and location characteristics. Results obtained with WHASA were evaluated with respect to a manual reference on 67 subjects, for which images have been acquired in six different centres. Performances were also compared to state-of-the-art methods on the same dataset. Finally, WHASA was applied on subjects from new studies to assess its robustness. A total of 260 subjects, for which images were acquired from eleven MRI scanners, including 1,5 T and 3 T machines, and with slice thicknesses varying between 2.5 and 6.2 mm, has been analyzed with WHASA.