



HAL
open science

Modeling and measuring the brain morphological evolution using structural MRI in the context of neurodegenerative diseases

Raphaël Sivera

► **To cite this version:**

Raphaël Sivera. Modeling and measuring the brain morphological evolution using structural MRI in the context of neurodegenerative diseases. Medical Imaging. COMUE Université Côte d'Azur (2015 - 2019), 2019. English. NNT : 2019AZUR4082 . tel-02389924v2

HAL Id: tel-02389924

<https://theses.hal.science/tel-02389924v2>

Submitted on 12 Jun 2020

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

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

THÈSE DE DOCTORAT

Modélisation et mesure de l'évolution morphologique du cerveau à partir d'IRM structurelles pour l'étude des maladies neurodégénératives

Raphaël SIVERA

dirigée par Nicholas AYACHE

co-encadrée par Hervé DELINGETTE, Xavier PENNEC et Marco LORENZI

préparée à l'Inria Sophia Antipolis, équipe EPIONE

soutenue le 29 novembre 2019

Présentée en vue de l'obtention du grade de Docteur en Automatique, Traitement du Signal et des Images de l'UNIVERSITÉ CÔTE D'AZUR devant le jury composé de :

Jean-François MANGIN	- CEA (Neurospin)	Président
Pierrick COUPÉ	- CNRS (LaBRI - UMR 5800)	Rapporteur
Stanley DURRLEMAN	- Inria (Aramis Lab)	Rapporteur
Nicholas AYACHE,	- Inria (équipe Epione)	Examineur
Hervé DELINGETTE	- Inria (équipe Epione)	Examineur
Xavier PENNEC	- Inria (équipe Epione)	Examineur
Marco LORENZI	- Inria (équipe Epione)	Invité
Philippe ROBERT	- Université Côte d'Azur (EA Cobtek)	Invité

Membres du Jury

Titre français

Modélisation et mesure de l'évolution morphologique du cerveau à partir d'IRM structurelles pour l'étude des maladies neurodégénératives

Titre anglais

Modeling and measuring the brain morphological evolution using structural MRI in the context of neurodegenerative diseases

Cette thèse a été présentée en vue de l'obtention du grade de DOCTEUR EN SCIENCES de l'UNIVERSITÉ CÔTE D'AZUR le 29 novembre 2019 devant le jury composé de :

Président	Jean-François MANGIN	-	CEA (Neurospin)
Rapporteurs	Pierrick COUPÉ	-	CNRS (LaBRI - UMR 5800)
	Stanley DURRLEMAN	-	Inria (Aramis Lab)
Examineurs	Nicholas AYACHE,	-	Inria (Epione team)
	Hervé DELINGETTE	-	Inria (Epione team)
	Xavier PENNEC	-	Inria (Epione team)
Invités	Marco LORENZI	-	Inria (Epione team)
	Philippe ROBERT	-	Université Côte d'Azur (EA Cobtek)

Résumé

En imagerie médicale computationnelle, l'analyse statistique des déformations permet de caractériser les effets des maladies neurodégénératives sur la morphologie cérébrale. Elle soulève cependant des problèmes méthodologiques spécifiques et l'interprétation des résultats reste complexe. L'objectif de cette thèse est de mettre en place un cadre méthodologique et applicatif permettant une meilleure interprétation clinique des résultats.

Dans la première partie, nous proposons un modèle joint des effets du vieillissement et de la maladie d'Alzheimer sur le cerveau. Le modèle propose une description simple des deux processus et peut être utilisé pour générer des évolutions réalistes et personnalisées pour divers diagnostics.

Dans la seconde partie, une analyse statistique morphométrique est réalisée dans le cadre de l'essai clinique MAPT. Un effet de l'intervention multidomaine sur la morphologie cérébrale est mis en évidence en multivariée alors qu'il n'est pas directement observable sur les variables cliniques ou les mesures volumétriques. Nous montrons que ces différences associées au traitement sont corrélées à de meilleures performances cognitives.

La troisième partie détaille plus en profondeur la méthodologie statistique utilisée. Il s'agit d'un cadre général de test d'hypothèse à partir d'images multivariées qui généralise les outils non paramétriques existants et requiert peu d'hypothèses pour être appliqué.

Enfin, la dernière partie explore les relations existantes entre la morphologie cérébrale et la cognition chez les personnes âgées. Les corrélations spatiales et les schémas d'évolution observés suggèrent l'existence de plusieurs dynamiques d'évolutions.

Mots-clés : morphométrie, vieillissement, maladie d'Alzheimer, IRM, modèles longitudinaux, statistiques multivariées

Abstract

In medical imaging, the statistical analysis of deformations enables the characterization of the effects of neurodegenerative diseases on the brain morphology. Deformations are able to capture precise changes but their analysis raises specific methodological challenges and the results may be difficult to interpret. The objective of this thesis is to present deformation-based methods and to show applications that contribute towards a better clinical interpretation of morphological changes.

In the first part, we introduce a joint model of the effects of normal aging and Alzheimer's disease on the brain morphology. The model proposes a simple description of both processes and is used to generate realistic and personalized evolutions under several diagnosis conditions.

In the second part, a morphometric study is conducted on the MAPT cohort. We bring out an effect of the multidomain intervention on the longitudinal deformation of the brain using multivariate statistics. This effect is not observable using clinical assessments or volumetric measures, but we show that the differences associated with the treatment are correlated with better cognitive performance.

The third part extends the statistical methodology used in the second part. A complete hypothesis testing framework for multivariate images is presented. It generalizes existing non-parametric frameworks and requires few hypothesis on the data to be applied.

Finally the last part builds on the methodology of the previous sections to explore the relation between morphology and cognition in elderly subjects. The spatial correlations and the patterns of evolution described in this section suggest the existence of several dynamics of evolutions that are associated with specific cognitive changes.

Keywords: morphometry, aging, Alzheimer's disease, MRI, longitudinal models, multivariate statistical analysis

Remerciements

Cette thèse n'aurait évidemment pas été possible sans le soutien de mes encadrants Nicolas, Hervé, Xavier et Marco. Je tiens à vous remercier chaleureusement, ce travail est autant le votre que le mien. Merci pour tout ce que vous m'avez appris, je pense que vous êtes aussi arrivé à me transmettre un peu de la rigueur et du recul nécessaire au travail de recherche.

Je voudrais aussi remercier Philippe Robert, Valeria Manera et Nicolas Capet avec qui j'ai eu la chance de collaborer pour l'analyse des données MAPT. Travailler avec vous a été très stimulant et très agréable.

Enfin je remercie les membres du jury et en particulier Pierrick Coupé et Stanley Durrleman pour leur lecture attentive et leurs commentaires sur le travail présenté dans ce manuscrit et merci à Jean-François Mangin d'avoir accepté de présider mon jury de thèse.

À Rocío, Chloé, Loïc L., Nicolas C., Jan, Mathieu, Nina, Bishesh, Mehdi, Marc-Mich, Thomas, Sophie, Roch, Pawel, Shuman, Julian, Qiao, Wen, Luigi, Nicoco, Tania, Fanny, Yann, Benoit, Clément, Jaume, Sara, Buntheng, Nicolas G., Gaëtan et tous les autres membres d'Asclepios, d'Epione et d'ailleurs avec qui j'ai passé de nombreuses pauses café, soirées beachvolleys, journées à Auron... durant ces quatre années riches en souvenirs.

À ma famille.

Contents

1	Context and objectives	1
1.1	Age related neurological disorders	1
1.1.1	Alzheimer’s disease	1
1.1.2	Toward a better understanding of disease progression	2
1.2	Brain morphometry	3
1.2.1	Segmentation-based approaches	3
1.2.2	Voxel-wise morphometry and deformation modeling	4
1.2.3	The stationary velocity field framework	4
1.3	General objective	5
1.4	Manuscript organization and contributions	6
2	Modeling the effects of aging and Alzheimer’s disease on the brain morphology	11
2.1	Introduction	12
2.1.1	Modeling brain morphological aging	12
2.1.2	Disease progression modeling	13
2.1.3	Toward a joint model of brain aging and disease progression	14
2.2	Definition of the generative model	15
2.2.1	A space of reference morphologies	15
2.2.2	Modeling the morphological evolution of the template	16
2.2.3	Individual morphological variability and generative model	17
2.3	Estimation of the model parameters	18
2.3.1	Estimation of the template trajectories v_A and v_D in a given template space	19
2.3.2	Estimation of the template morphology T_0	20
2.3.3	Estimation of the subject’s parameters	21
2.4	Results	21
2.4.1	Experiments with synthetic data	21
2.4.2	Experiments on ADNI data	24
2.4.3	Generating diagnosis driven morphological evolution	33
2.5	Discussion and perspectives	36
2.5.1	Approximations in the model estimation	37
2.5.2	Limitations of the description	38
2.5.3	Perspectives	38
2.6	Appendix	39
2.6.1	Validation on a synthetic dataset	39
2.6.2	Bias on the estimated template trajectories	40

3	Exploratory morphometry for the MAPT clinical trial	43
3.1	Introduction	44
3.1.1	The MAPT study	44
3.1.2	Measuring the evolution of the brain morphology	45
3.2	Method	46
3.2.1	Participants	46
3.2.2	Regional atrophy and cortical thickness assessments	46
3.2.3	Statistical analysis of the deformations	46
3.3	Results	49
3.3.1	Cognitive, behavioral and functional assessment	49
3.3.2	Segmentation-based statistical analysis of the interventions' effect	49
3.3.3	Deformation-based statistical analysis of the interventions' effect	51
3.3.4	Description of the MDI effect on the brain morphology	51
3.3.5	Relation to cognitive and behavioral scores	55
3.4	Discussion	56
3.4.1	Interpretation of the morphological changes	57
3.4.2	Limitations	57
3.4.3	Perspectives	58
3.5	Conclusion	58
4	Statistical framework for hypothesis testing on deformations	61
4.1	Context and objectives	62
4.1.1	Deformation-based morphometry	62
4.1.2	Clinical motivation: the MAPT case study	63
4.1.3	Chapter overview	63
4.2	Method	64
4.2.1	Multivariate General Linear Model	64
4.2.2	Likelihood ratio test	65
4.2.3	Design for non-parametric statistical testing on images	65
4.3	Experimental results in the MAPT study	67
4.3.1	Results of the multivariate analysis	67
4.3.2	Comparison to the atrophy-based assessments	68
4.4	Alternative methodological choices	70
4.4.1	Voxel statistic	70
4.4.2	Summary statistic	72
4.5	Discussion	76
4.5.1	Statistical analysis parameters	76
4.5.2	Interpretation of morphological differences	77
4.6	Conclusion	78
4.7	Appendix: the Likelihood ratio test and Aikake's weights	78

5	Relations between cognitive assessments and brain morphological evolution in the MAPT cohort	79
5.1	Introduction	80
5.2	Data and Method	81
5.2.1	Participants: the MAPT cohort	81
5.2.2	Modeling the brain evolution	81
5.2.3	Statistical analysis of the correlation	81
5.2.4	Cross decomposition of the cognitive and morphological evolution	82
5.3	Correlation between morphological changes and cognitive and behavioral scores	82
5.4	PLS between cognitive variables and morphological evolution	84
5.4.1	Dimension reduction with PCA	86
5.4.2	Cross decomposition with PLS	86
5.5	Discussion	88
5.5.1	Correlations with cognitive scores	88
5.5.2	Cross-decomposition of morphological and cognitive evolutions	88
5.5.3	Limitations and perspectives	88
6	Conclusion and Perspectives	93
6.1	Summary of the main contributions	93
6.2	Publications	95
6.3	Perspectives	95
6.3.1	Describing the brain deformations	95
6.3.2	Multimodal approach	96
6.3.3	Spatio-temporal patterns and disease evolution subtypes	97
A	Appendix A: Algorithmic contribution to the measure of the brain evolution	99
A.1	Robust-FOV	99
A.1.1	Context	99
A.1.2	Model	100
A.1.3	Implementation	101
A.2	Alignment: quality control	101
A.2.1	Segmentation and skull-stripping	102
A.2.2	Residual affine component in a SVFs	103
A.3	Affine invariant LCC-demons registration	103
A.3.1	Local correlation coefficient	103
A.3.2	Evaluation on controlled experiments	104
	Bibliography	107

Context and objectives

Contents

1.1	Age related neurological disorders	1
1.1.1	Alzheimer’s disease	1
1.1.2	Toward a better understanding of disease progression	2
1.2	Brain morphometry	3
1.2.1	Segmentation-based approaches	3
1.2.2	Voxel-wise morphometry and deformation modeling	4
1.2.3	The stationary velocity field framework	4
1.3	General objective	5
1.4	Manuscript organization and contributions	6

1.1 Age related neurological disorders

Age-related neurological disorders are a growing public-health concern with the aging of the population ([National Institute of Aging, 2019](#)). Life expectancy is increasing and the risk of developing a neurodegenerative disease dramatically increases with age. For example, the proportion of people with Alzheimer’s disease (AD) doubles every 5 years above age 65 and about one third of people older than 85 may be affected.

1.1.1 Alzheimer’s disease

Alzheimer’s disease is the most common form of dementia for older adults. It progressively destroys memory, reasoning, language and eventually disrupts the ability to carry even the simplest tasks and to have normal social interactions. The symptoms of the disease were first described by Dr A. Alzheimer in 1906. Since then, many progresses have been done on the understanding of the disease and its characteristics ([Alzheimer’s Association and others, 2017](#)).

The first changes associated with AD are visible at the cellular level with, in particular, an abnormal accumulation of beta-amyloid 42 protein between neurons, forming plaques that can disrupt normal metabolism. In the mean time, another abnormal accumulation of a protein called tau is observed inside the cells. These accumulations in conjunction with other factors (such as inflammation or vascular

problem) further injure the neurons and disrupt the normal axonal and the synaptic communication. Finally, these dysfunctions directly result in a massive increase of the cell death rate.

These deaths and the structural changes affecting the neurons are visible at a macroscopic level with regions of the brain shrinking. This atrophy is widespread in the brain but not uniformly distributed. Beyond the most used atrophy markers that are generally measured in the hippocampi, the atrophy is also characterized in the enthorinal cortex and most of the neocortex along a temporo-parietal-frontal trajectory. The changes are also visible in the white matter (through white matter tract abnormality) and in subcortical structure such as the amygdala and the thalamus (Pini et al., 2016). However these symptoms and their progression can differ between patients and the structural changes are often used as a complementary diagnosis tool.

Finally, the degradation of the functioning of the brain is reflected in the cognition with memory decline, language impairment, and emotional and behavioral changes. As the disease progresses, the general condition declines making constant care often necessary, and leading to the death of the patient.

1.1.2 Toward a better understanding of disease progression

The causes and origins of the disease and the mechanisms involved are not fully understood yet. Several risk factors have been identified regarding the genetic (in particular for the early-onset familial AD or relatively the APOE gene for late-onset AD) or the general health and vascular conditions but the precise characterization of the proceedings of the disease is still under investigation. As of today, the only available treatments aim to slow down the progression of the cognitive symptoms for some patients. Eventually the objective would be to identify an eventual treatment or some prevention actions.

Time plays a central role in the development of neurodegenerative diseases. First, age can be considered as the first risk factor and it is associated with the general effect of time on the body and the mind. Second, these diseases develop over several decade and the evolution between the first signs and the terminal stages can last over 20 years. These aspects have emphasized the importance of the description of the evolution for the prognosis (the prediction of the evolution) and the prevention of the disease.

From the practical point of view, longitudinal studies have been conducted to follow a cohort of subjects over several years. The longitudinal follow-up gives an intrinsic control on the individual variability that can be leveraged to refine the cross-sectional observations. The understanding of the individual evolution is crucial to design potential treatment, assess their efficacy in clinical trial and eventually predict the future progression of the patient to propose a personalized response.

1.2 Brain morphometry

The brain is a complex organ that can be described at many scales and from many points of view. It contains billions of neurons with their own metabolism that communicate between them and with the rest of the body through electrical and chemical signals in a huge interconnected network. The brain is also the central place for the consciousness and more generally all the cognitive activities.

The study of the brain is made even more complicated by the fact that the object of interest is inside the skull of a living person and is by consequence not easily accessible. Post-mortem dissections and histological analyses were crucial to describe the anatomy of the brain and remain important to understand, for instance, the neural mechanisms involved in AD (Ikonovic et al., 2008). Animal model can also help understanding the brain (Roy et al., 2016). More significantly neuroimaging techniques make it possible to virtually explore the brain and give direct insights about its structure and its activity. Many different modalities are now available, each one informing us about different aspects. These modalities cover some metabolic function using Positron-Emission Tomography (PET), the functional activity of the cells for electroencephalography (EEG) or functional MRI, the tissue structure using diffusion imaging for example and the anatomy with specificity related to each modality.

In this work, we focus on the morphological changes affecting the brain. We will then introduce, in this section, morphometric approaches that deal with the assessments and the description of the shape of the brain and its evolution.

1.2.1 Segmentation-based approaches

The first step is to have a good description of the anatomy of the brain. Anatomical descriptions of the brain can be found since ancient Egypt. Nowadays, modern anatomical atlases built from histological sections are complete and precise (Talairach, 1988; Mai et al., 2015). These maps give accurate delineations or segmentations of cortical and subcortical area that can be adapted to subject-specific data.

Nevertheless the manual segmentation of the whole brain or even of a single structure is time-consuming. In the last decades segmentation softwares were proposed to automatically or semi-automatically perform this task usually using the manually defined atlas as a prior to their segmentation. The focus of segmentation software can be diverse: whole brain segmentation (Fischl et al., 2002; Manjón and Coupé, 2016), tissue classification with a focus on WM/GM (Ashburner and Friston, 2005) or specialized in a substructure such as the hippocampus (Carmichael et al., 2005; van der Lijn et al., 2008; Coupé et al., 2011).

From these segmentations it is easy to derive measurements that characterize these anatomical regions of interest (some measures can be computed without a real segmentation, for example with landmarks, but the idea is similar). The most common measures are arguably sub-cortical volumes and cortical thicknesses but more complex indicators can be designed such as the cortex curvature or folding (Pereira

et al., 2012).

Thanks to the availability of software providing brain analysis methods, these measures can be nowadays reproduced, evaluated and benchmarked in different applications. Their clinical use is still limited but these measures can be compared between subjects or groups and have been proven useful to monitor the brain evolution, characterize disease progression or to assist early-diagnostic of neurodegenerative diseases (Devanand et al., 2007; Colliot et al., 2008; Risacher et al., 2010; Coupé, 2018).

However, even if these approaches offer valuable quantitative assessments, they also show some limitations related to their design: some regions are harder to segment and are by consequence under-studied, scalar markers are unable to fully characterize the shape of a brain structure, and the resolution of the description is limited by the available segmentations. By consequence, many alternatives have been proposed in the last decades.

1.2.2 Voxel-wise morphometry and deformation modeling

The use of high-resolution digital images leads to a family of techniques that process the full images voxel-wise. These methods do not rely on a priori defined regions but aim for a description at the voxel level. The most important example is certainly the voxel-based morphometry (VBM) methods for the statistical comparison of gray matter density map (Ashburner and Friston, 2000), but similar approaches are commonly applied to various imaging data such as PET amyloid (Kemppainen et al., 2006) or fMRI (Liu et al., 2008).

In parallel of these methods focusing on the comparison of voxel intensities, some approaches aim at describing the morphological deformation that transform one brain into another one (Ashburner et al., 1998; Arsigny et al., 2006; Zhang et al., 2010). Deformation-based morphometry requires more accurate and higher resolution estimations of the deformations but is able to explicitly model the differences and to describe some subtle local changes.

The dichotomy between voxel-based and deformation-based methods is not absolute. Most VBM approaches use non-linear deformations to normalize the images (using modulation for example to take into account the deformation) (Hutton et al., 2009). Likewise, deformation-based approaches go beyond the local comparisons of deformation fields and are able to describe global shape changes (Cury et al., 2016) or to spatially normalize manifold valued data (Qiu et al., 2007; Koval et al., 2017).

1.2.3 The stationary velocity field framework

The work presented in this thesis is centered around deformation-based morphometry. More specifically, the deformation framework used is based on an parametrization of diffeomorphisms using stationary velocity fields (Arsigny et al., 2006). In this setting, smooth and reversible deformations that spatially align the anatomies are estimated through non-linear registration (Lorenzi and Pennec, 2014) between pairs

of images. These deformations are fully parametrized by the flow of a stationary velocity field (SVF). This approach is similar to the Large Deformation Metric Mapping (LDDMM) framework (Beg et al., 2005) but no metric between deformations is chosen *a priori* (even if we need a regularization criterion for the registration). To compute a deformation ϕ we integrate trajectories along the vector field v for a unit of time.

$$\phi(x) = \phi_1(x) = \int_0^1 v(\phi_t(x)) dt \quad \text{with} \quad \phi_0 = \text{Id.}$$

From these deformations it is possible to compute the corresponding volume changes locally in the brain (Lorenzi et al., 2013), but it is also possible to analyze the longitudinal intra-subject evolutions directly in the deformation space using the SVFs transported in a common reference space where they can be compared. A more complete description and discussion about the complete processing pipeline for the comparison of longitudinal evolution is described in (Hadj-Hamou et al., 2016). During this thesis, some practical contributions to this pipeline were developed; they are summarized in Appendix A.

1.3 General objective

With this thesis, we aim at contributing to the description of the longitudinal evolution of the brain in aging and in Alzheimer’s disease with a focus on the morphological changes. The description of AD symptoms shows a complex entangling of multiple mechanisms that unroll over several decades and some of them are reflected by morphological changes very early in the course of the disease. Moreover, mathematical and algorithmic tools in morphometry provide a large spectrum of well-established quantitative approaches that have been proven useful to study AD.

Our objective is to develop new approaches for the characterizations of the morphological changes that leverage on the versatility of deformation-based modeling to propose simple but extensive descriptions that go beyond simple volumetric measures. The idea is that the deformations could reflect cellular structural change that are not captured using simpler scalar assessments and that they consequently could help precise the clinical assessment of the disease progression.

We also consider that the relation with aging and the temporality of the disease are central key-points to the understanding of the brain evolution. To this end, we focus on longitudinal analysis and we draw on the existing databases of images acquired over several years to characterize the disease progression. Regarding the clinical applications, the problem is twofold: the description of the population-wise trends is interesting to understand the disease or to the conduct clinical trials but, in general, the clinical practice aims to take into account the specificity of the subject evolution in regard of the general model of the disease.

1.4 Manuscript organization and contributions

Each of the 4 chapters of this manuscript addresses different aspects of our main objective. We introduce here the focus and the contributions of each part.

First, in chapter 2, we consider the problem of the modeling of the morphological evolution. The proposed deformation-based generative model aims to jointly take into account normal aging and AD, and to fill the gap between the short-term individual observations and the long-term disease evolution.

Relatively to this descriptive model, the next chapters are more related to exploratory approaches. Statistical analysis of the images proved crucial to characterized the changes associated with Alzheimer’s disease. In chapter 3, we apply morphometric methods to the MAPT clinical trial to assess the potential effect of clinical interventions. The multivariate analysis of the deformations reveals an effect that is not visible with traditional approaches. This leads us to the general hypothesis testing framework for multivariate images that we present in chapter 4. This framework enables the statistical analysis of deformations or complex imaging data in various contexts.

This framework is in particular used in chapter 5 to explore the relations between brain morphology and cognition. This last chapter concludes our thesis by opening perspectives towards a more inclusive description of the brain changes that takes into account, in addition to the morphology, other clinically relevant variables.

We now summarize each chapter individually and recall their main contributions.

Modeling the effects of aging and Alzheimer’s disease on the brain morphology. The morphological aging of the brain and disease progression have generally been modeled independently even though they involve intertwined processes, and their consequences may affect the same structural features of the brain. In chapter 2, we introduce a deformation-based framework to jointly model the influence of aging and Alzheimer’s disease on the brain morphological evolution. The model is used to describe the spatio-temporal evolution of both processes and to define two cross-sectional progression markers: 1) a *morphological age* and 2) a *disease score*. The longitudinal evolution of these markers is then studied in relation with the clinical diagnosis of the subjects and used to generate possible morphological evolutions.

We show that the joint modeling of normal and pathological evolutions using deformations leads to promising results to describe age-related brain diseases over long time scales. Our generative model is simplistic but the markers are able to confirm the presence of an accelerated apparent aging component in Alzheimer’s patients and to highlight specific morphological changes that can help discriminate clinical conditions, even in prodromal stages.

The main contributions developed in this part are:

- We disentangle aging and disease progression for elderly patients. We propose a simple description of the brain morphological evolution using only two

interpretable image-based progression markers: the morphological age and a disease score.

- These markers are cross-sectional assessments and are consistent for intra-subject longitudinal analyses. They can be seen as alternative aging measurements compatible with ongoing biological processes.
- We show that the markers and the generative model can be used in a personalized image simulation setting. The framework allows us to generate smooth and realistic evolutions for several diagnosis conditions.

This chapter is based on the work published in *A model of brain morphological changes related to aging and Alzheimer's disease from cross-sectional assessments*, Raphaël Sivera, Hervé Delingette, Marco Lorenzi, Xavier Pennec, and Nicholas Ayache. *NeuroImage*, 2019.

Morphometry for the MAPT clinical trial. In chapter 3, a complete morphometrical analysis is conducted in the context of the MAPT study.

The Multidomain Alzheimer Preventive Trial (MAPT) was designed to assess the effect of an omega-3 supplementation and a multidomain intervention (physical activity, cognitive training and nutritional advice) on cognitive decline of people with subjective memory complaint. In term of cognitive testing, no significant effect on cognitive decline was found.

The objective of this chapter is to assess an eventual effect of the interventions on the evolution of the brain morphology using the MR images acquired during MAPT. The morphological changes are assessed from volume measurements of regions of interest and a voxel-wise deformation-based approach that was specifically developed for this problem.

We find that the multi-domain intervention is associated with a significant effect on the 3-year morphological evolution. The voxel-wise deformation-based approach shows that the differences are mainly located in the left peri-ventricular area next to the temporo-parietal junction. These morphological changes correspond to a slower morphological evolution and are correlated with a better performance in cognitive assessments. These results could not be observed using the volumetric morphometry approach. We argue that the use of neuroimaging could help better define early intervention strategies that are effective to delay cognitive decline and dementia.

The main contributions developed in this part are:

- A significant effect of the multidomain intervention on the brain morphology is found.
- We are able to relate this effect to a slower morphological evolution and better performance in cognitive assessments.
- From the methodological standpoint, this work illustrates the fact that the statistical analysis of longitudinal deformations is able to discriminate condi-

tions more precisely than segmentation-based approaches even without priors on the affected areas.

This work have been submitted and is under-review: *Voxel based assessments of treatment effects on brain evolution in the MAPT cohort*, Raphaël Sivera, Nicolas Capet, Valeria Manera, Roxane Fabre, Marco Lorenzi, Hervé Delingette, Xavier Pennec, and Nicholas Ayache. submitted to *Neurobiology of Aging*, 2019.

Statistical framework for the analysis of deformation. The statistical analysis of the deformations was central in the Chapter 3 to characterize the effect of the treatments in the MAPT study. However, the focus was on the description of the results and their clinical implications. The methodological contributions were then secondary. Nonetheless, because the causes of AD are not fully understood and because no treatment is available yet, the development of general approaches that can be easily adapted to new problems is a important challenge.

Voxel-wise statistical analysis for scalar images is a well-established domain in medical imaging. However, the usual approaches are limited in their characterization of the morphological changes and the analysis of the multivariate fields parametrizing the deformations raises new issues and challenges.

In this context, Chapter 4 describes the general mass-multivariate statistical testing framework for multivariate images that was used in Chapter 3. The framework is useful to characterize patterns at the population level and to highlight differences related to clinical variables. We present the theoretical foundations of the framework, detail the methodology and explore the importance of the main parameters of the approach. The question of the treatment effect in the MAPT study (introduced in Chapter 3) is used as the main illustration.

The main contributions developed in this part are:

- We present a complete framework for permutation-based hypothesis testing in multivariate images.
- We open several alternatives to the standard statistical analyses.
- No hypothesis on the data is required making it widely applicable in various context.

Relations between cognition and morphology. One of the main issue encountered in the first chapters is the interpretation of the observed morphological deformations. Measuring the deformations is a challenging task and the results may be method dependent. Then, the spatial interpretation is made even more complex by the inter-connectivity of brain areas and because the local information given by the displacements is not directly interpretable. If the deformations are not fully meaningful by themselves maybe the solution is to consider them in relation with other clinical variables.

To address this problem and open new perspectives on the description of the brain evolution, in Chapter 5, we explore the relations between the morphological

and cognitive evolution of aging subjects. The objective is double. First, aging and neurological disorders cannot be fully understood if we only consider one point view. The brain processes are intricated and their interactions are decisive to comprehend the whole evolution. This work is one step in this direction. Second, we want to use the cognitive information to inform our description of the morphological changes in order to interpret and eventually label specific patterns. In particular, the methods and the processing are similar to the ones used in Chapter 3 and this analysis may serve as a reference to better understand the observed treatment effect.

Two methodological approaches are proposed. The first one uses the statistical multivariate framework described in Chapter 4, the second relies on a cross-decomposition using partial least square (PLS). Spatial correlation patterns are described for several common cognitive scores and the main modes of the cognitive-morphological evolution are characterized in the MAPT data set.

The main contributions developed in this part are:

- We show significant spatial correlations between morphological changes in aging and the evolution of several cognitive variables.
- The spatial patterns are coherent with the assessed cognitive traits and the general knowledge of the brain.
- We highlight using PLS several modes of evolutions that can be interpreted as: 1. describing the general cognitive decline, 2. focusing on behavioral evolution, 3. specifically associated with the MDI treatment.

Modeling the effects of aging and Alzheimer’s disease on the brain morphology

Contents

2.1	Introduction	12
2.1.1	Modeling brain morphological aging	12
2.1.2	Disease progression modeling	13
2.1.3	Toward a joint model of brain aging and disease progression	14
2.2	Definition of the generative model	15
2.2.1	A space of reference morphologies	15
2.2.2	Modeling the morphological evolution of the template	16
2.2.3	Individual morphological variability and generative model	17
2.3	Estimation of the model parameters	18
2.3.1	Estimation of the template trajectories v_A and v_D in a given template space	19
2.3.2	Estimation of the template morphology T_0	20
2.3.3	Estimation of the subject’s parameters	21
2.4	Results	21
2.4.1	Experiments with synthetic data	21
2.4.2	Experiments on ADNI data	24
2.4.3	Generating diagnosis driven morphological evolution	33
2.5	Discussion and perspectives	36
2.5.1	Approximations in the model estimation	37
2.5.2	Limitations of the description	38
2.5.3	Perspectives	38
2.6	Appendix	39
2.6.1	Validation on a synthetic dataset	39
2.6.2	Bias on the estimated template trajectories	40

In this chapter, we introduce a deformation-based framework to jointly model the influence of aging and Alzheimer’s disease on the brain morphological evolution.

This framework is used to describe the spatio-temporal evolution of both processes and to define two cross-sectional progression markers: 1) a *morphological age* and 2) a *disease score*. The longitudinal evolution of these markers is then studied in relation with the clinical diagnosis of the subjects and used to generate possible morphological evolutions.

We show that the joint modeling of normal and pathological evolutions using deformations leads to promising results to describe age-related brain diseases over long time scales. Our generative model is simplistic but the markers are able to confirm the presence of an accelerated apparent aging component in Alzheimer’s patients and to highlight specific morphological changes that can help discriminate clinical conditions, even in prodromal stages.

This work was published in the following paper: *A model of brain morphological changes related to aging and Alzheimer’s disease from cross-sectional assessments*, Raphaël Sivera, Hervé Delingette, Marco Lorenzi, Xavier Pennec, Nicholas Ayache, NeuroImage, 2019 in press.

2.1 Introduction

Age-related diseases are a growing public health concern with the aging of the population. For this reason, a precise description of aging would be useful to predict and describe the evolution of these diseases. In complement to the *chronological age*, i.e. the time elapsed since birth, one would like to estimate a *biological age* that reflects the current physiological, functional or structural status of an organ relatively to the aging changes. However there is no unified way to describe aging in a clinical context since aging is a complex process which affects every part of the body with specific mechanisms and specific rate. As a consequence multiple theories of aging have been proposed (Medvedev, 1990), leading to the definition of surrogate age variables based on the quantification of biological changes.

2.1.1 Modeling brain morphological aging

In this chapter we focus on the aging of the brain based on the study of its shape evolution. The brain is not exempt from aging and a decline of cognitive processing speed, working memory, inhibitory function, and long-term memory is generally observed. This decline has been associated with neural activity changes (Park and Reuter-Lorenz, 2009) and it was also shown to be directly correlated with structural changes such as brain atrophy, cortex thinning and decrease of white matter integrity (Rosen et al., 2003; Rodrigue and Raz, 2004).

The normal brain morphology has been studied in image-based studies from the development stage to the most advanced ages. Measurements of brain structures (volumes, cortical thickness, etc.) have been performed for wide age ranges and the statistical analysis of the evolution of these measurements helps in providing an initial understanding of the normal brain shape evolution across life span (Good et al., 2001; Long et al., 2012). These descriptions have been used to estimate models

characterizing brain aging in order to highlight differences across brain areas (Hutton et al., 2009; Sowell et al., 2003). The inverse problem, i.e. how to associate an age to a brain image, was also addressed. Models have been designed to estimate the *chronological age* (Cole et al., 2017) from an image but they can also be used to characterize abnormal evolutions. For instance, a mean brain age gap estimate was highlighted for Alzheimer’s patients (Franke et al., 2010). In Huizinga et al. (2018), a morphological score is built to assess the effect of aging on the brain morphology and the authors show that AD subjects are associated with abnormally high scores. More generally, these surrogate brain age estimates have been associated with an increase of risk factors for several age-related disorders such as cardiovascular diseases (DeCarli et al., 2005; van Velsen et al., 2013). In a longitudinal setting, a brain age measurement could be used to compare the evolution of several clinical conditions (see Figure 2.1).

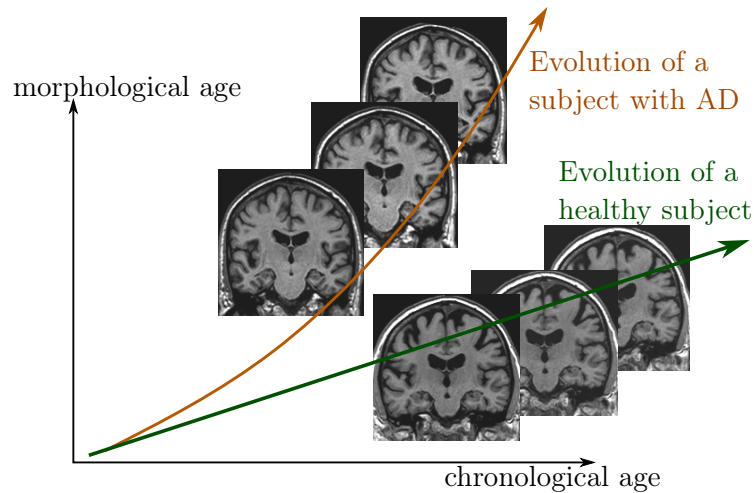


Figure 2.1: Schematic representation of two evolutions relative to an hypothetical *morphological age* reflecting the structural status of the brain relative to the aging process.

2.1.2 Disease progression modeling

Brain aging is often associated with the development of neurodegenerative pathologies. For example, it is estimated that one in three people over 85 have Alzheimer’s disease (AD), the most common form of dementia (Alzheimer’s Association and others, 2017). This disease comes with its own specific apparent brain morphological changes (Ohnishi et al., 2001) and computer-aided diagnosis techniques using neuroimaging features have shown promising results to classify and to predict clinical evolutions (Davatzikos et al., 2009; Klöppel et al., 2012; Schmitter et al., 2015). Longitudinal studies provide us with multiple acquisitions at different times for every subject but the disease affects patients over several decades, starting even before the first symptoms occur, and few studies follow a significant number of subjects

over such long times. Progression models have been developed to describe the global evolution and to put in relation the individual trajectories that could only be observed a limited number of times. They have been used to model the progression of biomarkers (Fonteijn et al., 2012; Donohue et al., 2014) but also directly brain shape (Cury et al., 2016) or spatio-temporal patterns in brain images (Koval et al., 2017; Schiratti et al., 2017). These models produce good representations of the disease progression and can combine a variety of available biomarkers for patient monitoring (Bilgel et al., 2016; Lorenzi et al., 2017).

2.1.3 Toward a joint model of brain aging and disease progression

The morphological aging of the brain and disease progression have generally been modeled independently. However we know that the structural features used in diagnosis (e.g cortical thickness or atrophy patterns) are also generally related to age. Indeed aging and neurodegenerative diseases involve intertwined processes with entangled consequences. Surrogate age measurements have been used to support the disease characterization (Franke and Gaser, 2012) or to put aside the aging part in order to focus on the disease specific changes. Lorenzi et al. (2015) proposed to model the normal aging evolution to separate the contribution of aging from a remainder that is not explicitly modeled. This remainder is then used to describe the pathological evolution. However, this method does not propose an intrinsic model of the disease progression making it difficult to describe and to characterize disease specific changes.

In this study, we propose a generative model of the brain morphological evolution that jointly takes into account the normal aging and the disease effects. Our model is based on the approach proposed by Lorenzi et al. (2015) and gives a deformation-based description of subject trajectories. It extends the original approach by explicitly modeling the disease specific brain morphological evolution. In addition to the apparent *morphological age* computed in the proposed approach, it allows us to get a *disease score*, thus providing two morphological imaging biomarkers accounting for the progression of the two main ongoing processes: normal aging and Alzheimer’s disease. These biomarkers are estimated cross-sectionnally from a single structural MRI but the model estimation exploits longitudinal data to match as accurately as possible the morphological evolution of the subjects.

We introduce in section 2.2 the generative model used to represent the brain morphology. Section 2.3 focuses on the estimation of the model parameters and on the inverse problem solved to compute the morphological age and the disease score of a subject. Experimental results are then presented in section 2.4 in order to evaluate our model and parameter estimation procedure. We illustrate how the model helps to describe the evolution of subjects at different disease stages using the ADNI database. Then we show how the two proposed markers can help to follow the evolution of elderly patients. Finally, limitations and perspectives are discussed in section 3.4.

2.2 Definition of the generative model

In the sequel, we quantify differences between morphologies by spatial deformations that can be estimated from magnetic resonance images (MRI) through non-linear image registration. A deformation represents either a morphological difference between the anatomies of two subjects or a longitudinal evolution of one subject-specific anatomy. Therefore deformation based frameworks are well suited to define a parametric model of the morphology. Our approach involves a population template morphology which is parametrized by two progression markers: the morphological age and the disease score.

In section 2.2.1, we expose the main ideas behind the definition of our reference parametric space. Then in section 2.2.2, we explain how the morphological evolutions are modeled in our framework. Finally in section 2.2.3, we show how we use the deformation framework to model individual morphologies relatively to this reference.

2.2.1 A space of reference morphologies

In deformation based morphometry, a single morphology is classically used to approximate and represent a population. For a set of images $\{I_k\}$, we define a common reference image T_0 , called template. The difference of morphology between the subject k and the template is modeled with a spatial deformation ϕ_k and an intensity noise ε_k is added in the subject space accounting for local intensity variability. Therefore the images are modeled as follows:

$$I_k = T_0 \circ \phi_k + \varepsilon_k.$$

In our approach, we want to take into account two major processes that affect the brain morphology over time: the normal aging and the disease evolution. To do so, we model the effects of these processes on the template using a deformation Φ parametrized by two variables that measure the progress of each process: the *morphological age* λ_{MA} and the *disease score* λ_{DS} . The two variables λ_{MA} and λ_{DS} can be seen as time variables and are scaled to correspond to years of standard evolution. In this model, $T_0 \circ \Phi(\lambda_{MA}, \lambda_{DS})$ represents the template morphology after λ_{MA} years of normal aging and λ_{DS} years of normalized disease progression.

In the ideal case, the morphological age is equal to the chronological age. Therefore, $T_0 \circ \Phi(t, 0)$ represents the morphology of a t years old healthy subject. Similarly $T_0 \circ \Phi(t, t')$ would be the typical morphology of a AD patient of age t with a disease duration of t' years. If one is able to associate an age λ_0 to the image T_0 , and assumes that this is the image of a healthy subject, then it is natural to enforce $\Phi(\lambda_0, 0) = \text{Id}$.

The parametric subspace of images generated like this will be used as a reference:

$$\mathcal{T} = \{T_0 \circ \Phi(\lambda_{MA}, \lambda_{DS}) \text{ for } \lambda_{MA}, \lambda_{DS} \in \mathbb{R}\}. \quad (2.1)$$

An initial model of the subject images is then:

$$I_k = T_0 \circ \Phi(\lambda_{MA}^k, \lambda_{DS}^k) \circ \phi_k + \varepsilon_k,$$

where λ_{MA}^k and λ_{DS}^k are the subject morphological age and disease score, while ϕ_k encodes differences specific to the subject morphology. Of course multiple options are available to combine the longitudinal deformation Φ with the subject specific changes ϕ^k . Here we write this operation as a right-composition but this choice will be discussed in section 2.2.2 and a similar but symmetrical operation will later be used.

Also, in the reference space, trajectories parametrized by time $t \mapsto T_0 \circ \Phi(\lambda_{MA}(t), \lambda_{DS}(t))$ give possible morphological evolutions where the morphological age and the disease score can be seen as reparametrization of time. In particular, it defines two archetypal trajectories (*i.e.* ideal models of evolution): the *normal aging template trajectory* $t \mapsto T_0 \circ \Phi(t, 0)$ and the *disease specific template trajectory* $t \mapsto T_0 \circ \Phi(\lambda_0, t)$.

Here, we assume that each progression of the two major processes (aging and disease) can be described with only one parameter. This implies that the evolution of healthy aging is similar for each subject, following the normal aging template trajectory, even if the speed of aging may vary from one subject to the other. Similarly, the disease progression is described using a single trajectory and we combine both template trajectories to model pathological evolutions.

2.2.2 Modeling the morphological evolution of the template

In a simplified approach, the template trajectories are assumed to be geodesics in an appropriate deformation space. Geodesics define continuous paths that can be easily parametrized and constrained to allow regularity in time. They can be used to interpolate between two anatomies or to approximate more complex trajectories (Christensen et al., 1994; Wang et al., 2007). In this work, we use the stationary velocity field (SVF) framework (Arsigny et al., 2006) for its ability to describe complex and realistic diffeomorphic (smooth and invertible) brain deformations in a straightforward manner (Lorenzi et al., 2013). In this framework, the observed anatomical changes are encoded by diffeomorphisms which are parametrized with the flow of SVFs. Within this setting, the metric between deformations is not chosen *a priori* even if we need a regularization criterion for the registration. To compute a deformation ϕ we integrate trajectories along the vector field v for a unit of time.

$$\phi(x) = \phi_1(x) = \int_0^1 v(\phi_t(x)) dt \quad \text{with} \quad \phi_0 = \text{Id}.$$

This relationship is denoted as the group exponential map $\phi = \text{Exp}(v)$.

By writing $\Phi(\lambda_{MA}, \lambda_{DS}) = \text{Exp}(\lambda_{MA}v_A + \lambda_{DS}v_D)$, we propose a linear model in the SVF space (*i.e.* the space of the parameter of the deformations) parametrized by two SVFs v_A and v_D . In particular, the two template trajectories are then

separately parametrized: v_A controls the normal aging template trajectory and v_D the disease specific template trajectory.

For each subject, the processes are meant to be intertwined and this can be modeled in different ways depending on the parametrization of the trajectories, for instance a right or a left composition. The proposed linear combination of the parameters provides us a middle ground. Indeed in the SVF setting, the relationship between composition and the linear combination of SVFs is given by the Baker-Campbell-Hausdorff formula (Bossa et al., 2007) and the linear combination of the SVFs is equivalent to alternate between right and left composition with infinitesimal steps.

To sum up, the longitudinal deformation Φ modeling the effects of aging and the disease on a reference morphology T_0 is parametrized by two SVFs: v_A and v_D . This ideal model generates a surface \mathcal{T} of possible images describing the evolution of the template morphology:

$$\mathcal{T} = \{T_0 \circ \text{Exp}(\lambda_{MA}v_A + \lambda_{DS}v_D) \text{ for } \lambda_{MA}, \lambda_{DS} \in \mathbb{R}\}. \quad (2.2)$$

2.2.3 Individual morphological variability and generative model

An individual image is modeled as follows:

$$I_k = T_0 \circ \text{Exp}((\lambda_{MA}^k - \lambda_0)v_A + \lambda_{DS}^k v_D) \circ \phi_k + \varepsilon_k, \quad (2.3)$$

where the choice of the intensity noise ε_k is implicitly related to the registration similarity metric. To specify the constraint on ϕ_k , we define a subject specific residual SVF w_r^k (r stands for residual) such that:

$$\text{Exp}((\lambda_{MA}^k - \lambda_0)v_A + \lambda_{DS}^k v_D) \circ \phi_k = \text{Exp}((\lambda_{MA}^k - \lambda_0)v_A + \lambda_{DS}^k v_D + w_r^k).$$

In this formula, $\text{Exp}(w_r^k)$ is approximately equal to ϕ_k given the first order of the BCH equation between composition and linear combination of SVFs. Moreover, we wish to have the subject specific deformation to encode what cannot be described using the template trajectories. That is why we impose w_r^k to be orthogonal to both v_A and v_D .

As we can see in Figure 2.2, the model parametrized by v_A and v_D allows us to characterize the subject morphology with two scalar variables, the morphological age λ_{MA}^k and the disease score λ_{DS}^k , and a SVF w_r^k for the subject-specific part. The orthogonality constraint makes the description of the subject uniquely defined. We denote by w^k the subject-to-template deformation SVF:

$$w^k = (\lambda_{MA}^k - \lambda_0)v_A + \lambda_{DS}^k v_D + w_r^k. \quad (2.4)$$

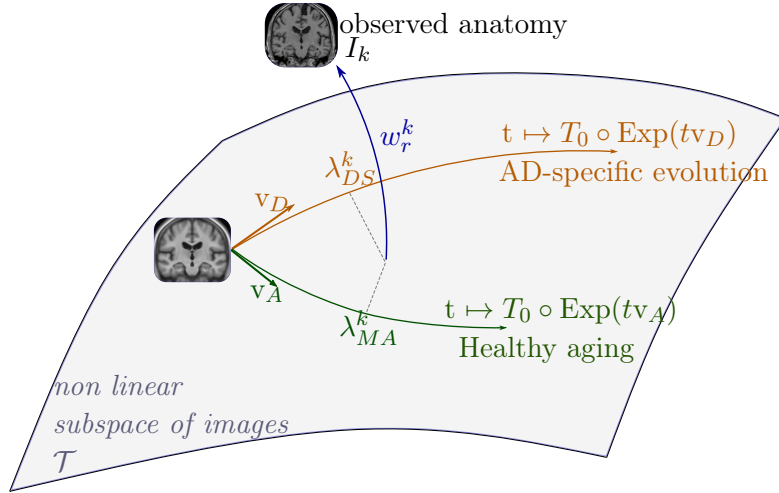


Figure 2.2: Our two-trajectory model. The template image T_0 , the normal aging template trajectory parametrized by v_A and the disease specific template trajectory parametrized by v_D define a subspace of possible morphologies of reference. An individual morphology is characterized by a morphological age λ_{MA}^k , a disease score λ_{DS}^k , and an SVF w_r^k modeling the subject-specific part. Each image can be projected onto the template subspace \mathcal{T} using a decomposition of the deformation between the image and the template T_0 .

2.3 Estimation of the model parameters

The model parameters are of two kinds: the population parameters (T_0 , v_A and v_D) and the subjects parameters (λ_{MA}^k , λ_{DS}^k and w_r^k). To tackle this joint estimation problem in a computationally efficient way, several assumptions are made:

1. When available, the longitudinal evolution in the template space can be approximated by the transported deformation estimated in the subject space. Parallel transport algorithms are commonly used in the geometrical analysis of longitudinal data. The use of geodesic parallelograms is in general an efficient way to bring individual trajectories in a common reference space (Lorenzi and Pennec, 2013). In practice, it allows us to work only with intra-subject deformation to estimate the model population parameters. It simplifies the optimization and is also more stable as the intra-subject variability is in general smaller than the inter-subject one.
2. We also assume, while estimating the population parameters, that the aging speed and the disease progression speed are constant for all the subjects in the training set.
3. Intra-subject deformations are relatively small and smooth. Consequently, the registration regularization has less impact on the estimated deformation. This consideration allows us to estimate these longitudinal evolutions independently

of the population model.

These assumptions allow us to efficiently decompose the problem of the parameter estimation. First, subjects with longitudinal data are processed independently and the intra-subject evolutions are modeled in the subject space. Then the population parameters (T_0 , v_A and v_D) are estimated using only intra-subject longitudinal evolutions. Finally, the subjects' parameters are estimated.

2.3.1 Estimation of the template trajectories v_A and v_D in a given template space

In this section we suppose that we know T_0 and that we can compute the subject-to-template deformation w^k for a reference time point. We also consider that we have longitudinal data for every subject.

First we address the inverse problem of estimating the intra-subject evolution parameters with the framework proposed by [Hadj-Hamou et al. \(2016\)](#). Images are preprocessed, rigidly aligned to the MNI-152 template and then longitudinally registered. Intra-subject deformations between follow-up images and the baseline image are computed using non-linear registration. The resulting intra-subject model in the subject's space is estimated using ordinary least square regression in the tangent space of SVFs. It is equivalent to the assumption that the deformation noises are centered, uncorrelated and have equal variance in the space of SVFs.

Then for a given template T_0 and subject-to-template deformation w^k , the intra-subject model can be transported using parallel transport in the template space to get v^k . This deformation can be decomposed along the template trajectories giving a morphological aging rate (noted s_{MA}^k), a disease progression rate (noted s_{DS}^k) and an orthogonal component (noted v_r^k):

$$v^k = s_{MA}^k v_A + s_{DS}^k v_D + v_r^k.$$

These progression rates are strongly related to the evolution of the morphological age and of the disease score (in a ideal euclidean case $s_{MA}^k = d\lambda_{MA}^k/dt$ and $s_{DS}^k = d\lambda_{DS}^k/dt$) and they will be used to normalize the speed of evolution. The estimation is done on two groups of subjects: a group \mathcal{G}_h composed by healthy subjects and a group \mathcal{G}_{ad} composed by patients diagnosed with AD. We assume that each healthy subject of \mathcal{G}_h is aging at normal speed $s_{MA}^k = 1, \forall k \in \mathcal{G}_h$ and does not have any evolution toward the disease $s_{DS}^k = 0, \forall k \in \mathcal{G}_h$. Similarly, each patient of \mathcal{G}_{ad} has a normal morphological aging rate $s_{MA}^k = 1, \forall k \in \mathcal{G}_{ad}$ and a constant unit disease progression rate $s_{DS}^k = 1, \forall k \in \mathcal{G}_{ad}$. Finally the subject specific components are assumed to be centered, uncorrelated and to have a fixed variance. The maximum likelihood problem writes:

$$\min_{v_A, v_D} \sum_{k \in \mathcal{G}_h} \|v^k - v_A\|^2 + \sum_{k \in \mathcal{G}_{ad}} \|v^k - v_A - v_D\|^2. \quad (2.5)$$

Data: an image I , a template image T and linear space of SVF \mathcal{T}

Result: two SVFs: $w_t \in \mathcal{T}, w_r \in \mathcal{T}^\perp$

$w_t = 0$;

repeat

$w_r = \text{registration}(T \circ \text{Exp}(w_t), I)$;
 $w_t = w_t + \text{proj}_{\mathcal{T}}(w_r)$;

until $w_r \perp \mathcal{T}$;

Algorithm 1: Iterative registration algorithm

The solution of the optimization problem is explicit:

$$\hat{v}_A = \frac{1}{|\mathcal{G}_h|} \sum_{k \in \mathcal{G}_h} v^k, \quad (2.6)$$

$$\hat{v}_D = \left(\frac{1}{|\mathcal{G}_{ad}|} \sum_{k \in \mathcal{G}_{ad}} v^k \right) - \hat{v}_A. \quad (2.7)$$

We should however note that $\|\hat{v}_A\|$ (resp. $\|\hat{v}_D\|$) is a biased estimator of $\|v_A\|$ (resp. $\|v_D\|$). We detail the bias estimation in the Appendix 2.6.2.

2.3.2 Estimation of the template morphology T_0

The population specific template morphology is computed using the algorithm proposed by Guimond et al. (2000) by alternating the registration of subject images to the template and the recomputation of the template intensities. However, in our approach, the subjects’ images do not need to be registered to T_0 directly but to their projection on the template space. To tackle this problem, we propose an iterative procedure where we register the image to its current projection on the reference space. Algorithm 1 details the procedure with simplified notations in the general case where w parametrizes the deformation between the image I and T . The reference linear subspace of SVF is denoted \mathcal{T} and w is decomposed accordingly $w = w_t + w_r$ with $w_t \in \mathcal{T}, w_r \in \mathcal{T}^\perp$. The registration regularization should only be applied to the residual part w_r . In the context of the LCC-demons registration algorithm, it boils down to the following minimization problem (see Lorenzi et al., 2013):

$$\min_{w_t \in \mathcal{T}, w_r \in \mathcal{T}^\perp, w'} \text{Sim}(I, T, \text{Exp}(w_t + w_r)) + \text{Dist}(w_r, w') + \text{Reg}(w').$$

The idea is to alternate between the optimization and the projection on the constraints.

Since we do not have any theoretical guarantee on the convergence of the algorithm, the stability and the convergence will be evaluated empirically. As the template estimation also involves iterative search, we can combine both algorithms for a faster optimization. The projection coordinates are kept from one iteration to the next and the images are registered to their estimated projections in the template

Data: a set of images (I_k) and a linear space of SVF \mathcal{T}

Result: a template image T , a set of pairs of SVFs (w_t^k, w_r^k)

$w_t^k = 0$ for all k ;

initialize T ;

repeat

$w_r^k = \text{registration}(T \circ \text{Exp}(w_t^k), I_k)$ for all k ;

$u = \text{mean}(w_r^k)$;

$T = \text{mean}(I_k \circ \text{Exp}(-w_r^k + u))$;

$w_t^k = w_t^k + \text{proj}_{\mathcal{T}}(w_r^k - u)$ for all k ;

until *convergence*;

Algorithm 2: Iterative template space estimation algorithm

space. The deformation update u is then computed and finally the new atlas image T and the estimated projections are updated (see Algorithm 2).

In this work, the intra-subject models are transported to the template space to update the template trajectories parametrizing \mathcal{T} at each iteration (using the approach described in the previous section 2.3.1). T is initialized using the MNI-152 template and the convergence is manually assessed comparing the template for successive iterations. At convergence, we obtain the template image T_0 and both template trajectories v_A and v_D .

2.3.3 Estimation of the subject's parameters

When the population parameters are learned, the estimation of the individual parameters for a new subject is relatively simple. The deformation w^k is computed by registration between a subject image and the template using Algorithm 1, and then linearly decomposed, $w^k = (\lambda_{MA}^k - \lambda_0)v_A + \lambda_{DS}^k v_D + w_r^k$, by solving the following linear system:

$$w^k \cdot v_A = \|v_A\|^2 (\lambda_{MA}^k - \lambda_0) + v_D \cdot v_A \lambda_{DS}^k, \quad (2.8)$$

$$w^k \cdot v_D = v_D \cdot v_A (\lambda_{MA}^k - \lambda_0) + \|v_D\|^2 \lambda_{DS}^k. \quad (2.9)$$

In practice, the estimation is not exact because we work with the noisy estimators \hat{v}_A and \hat{v}_D . The linear decomposition can also be computed locally or by using any voxel weighting scheme for the scalar product. When longitudinal data is available, this estimation is independently done for each time point.

2.4 Results

2.4.1 Experiments with synthetic data

We first evaluate our approach using synthetic data in order to assess the accuracy and the reproducibility of the biomarkers estimation. Realistic longitudinal MRIs are simulated using the software proposed by Khanal et al. (2016). The simulation

algorithm relies on a biophysical model of brain deformation and can be used to generate longitudinal evolutions with specific atrophy patterns. In this context, local atrophy is measured by the divergence of the stationary velocity field.

2.4.1.1 Simulated dataset

In this controlled experiments we choose to simulate two populations that are characterized by their atrophy patterns and that respectively emulate healthy controls and AD patients. Atrophy of the aging brain and the effect of AD have been extensively studied (Pini et al., 2016) and the atrophy measurements may vary depending on the methodology and the population studied. In this experiment, we choose to prescribe piecewise-constant atrophy map with constant value in brain areas delimited by the segmentation provided by *FreeSurfer* (Fischl et al., 2002). For every subject, the atrophy value of a region is sampled around a fixed population mean with an additive Gaussian noise of relative standard deviation of 5%. The healthy population is designed to have a small atrophy in the whole brain while the patients have a stronger atrophy especially in the hippocampal areas and the temporal poles. The means are chosen to give the order of magnitude of a one year evolution accordingly to what was reported in Fjell et al. (2010) for healthy aging and in Carmichael et al. (2013), with an additional scaling for the pathological evolution. We detail the exact regional values in Table 2.2 in the appendix (see 2.6.1).

Structural MRIs of 40 healthy subjects from the ADNI database are taken as input to the simulations. For every subject, deformations are simulated for both pathological and healthy settings. The deformation extrapolated 5 times is then applied to the original image in order to simulate 5 years long evolutions. We then have two matched populations of 40 pairs of images.

2.4.1.2 Model estimation

Individual longitudinal deformations are computed using registration, and the reference anatomy and the template trajectories are built using our framework. The divergence fields associated with these template trajectories can be compared to the prescribed atrophy. Figure 2.3 shows the average atrophy maps in the estimated template anatomy.

The estimated atrophy patterns are smoothed versions of the simulated ground truth. This effect was already observed (Khanal et al., 2016). First of all, the registration algorithm is unaware of the underlying simulation model and is unable to localize precisely the atrophy in homogeneous areas. Moreover, the spatial regularization of the registration and the parametrisation using SVFs also contribute to the smoothness of the estimated atrophy patterns. This is particularly visible in small (hippocampus) or thin regions (cortex). We therefore have a consistent bias when the atrophy measurements are integrated over the regions (see Figure 2.14 in appendix). Indeed, the local atrophy is affected by neighboring regions evolving in the opposite direction (the ventricles or the CSF for example). However, we can see that

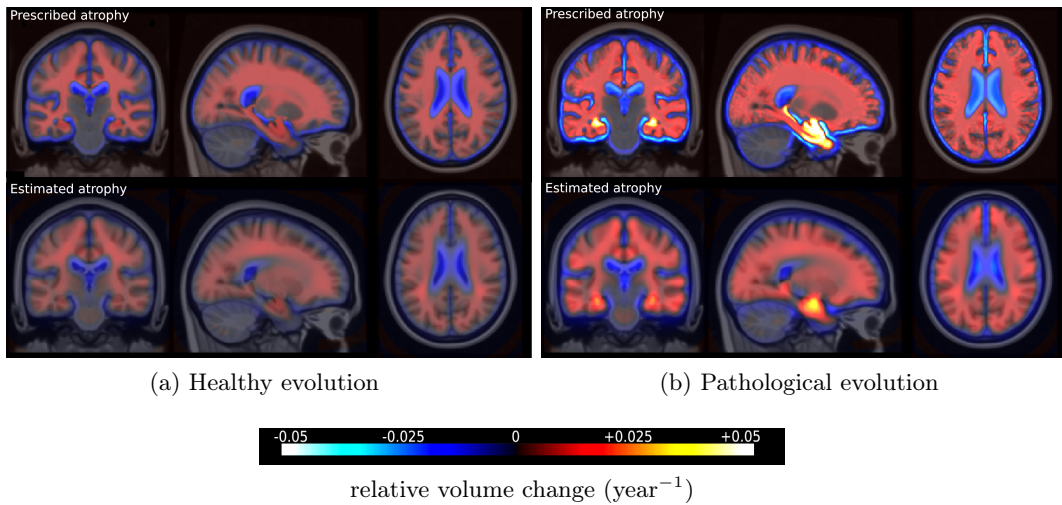


Figure 2.3: Simulated and estimated atrophy patterns in the template space. Top row: mean in the template space of the prescribed atrophy maps. Bottom row: atrophy maps estimated from the simulated images. Left: healthy simulations. Right: pathological simulations. The estimations are smoother but qualitatively similar to simulated maps.

the ratio between the pathological and healthy cases is conserved in every region. It was already noted that quantitative estimation using registration can be biased but can be more reproducible than the segmentation based approaches (Hadj-Hamou, 2016; Cash et al., 2015).

2.4.1.3 Imaging biomarkers estimation

The morphological age and the disease score were computed for each image. By construction there is no difference at $t=0$ (exactly the same images in the two groups). We compare the simulated differences at $t=5$ and the evolution of the cross-sectional assessments for each subject (see Figure 2.4a).

In this experiment, the initial anatomical variability is important, indeed the standard deviation at $t=0$ is equal to 48 years for $MA(0)$ and 24 years for $DS(0)$. At $t=5$, a difference is visible for the disease score but it is still diluted in the inter-subject variability. It is also possible to extrapolate the evolution to determine how many years of evolution are needed in order to get a significant difference between the healthy and the disease groups. For a significance level of 0.05, the disease score would be significantly discriminant after 13 years of evolution. This figure for synthetic data highlights the slow time-pace of the disease and the interest in modeling and extrapolating the evolutions.

Looking at the evolution of these cross-sectional biomarkers in Figure 2.4b), we see that the measures are relatively stable despite the large inter-subject variability

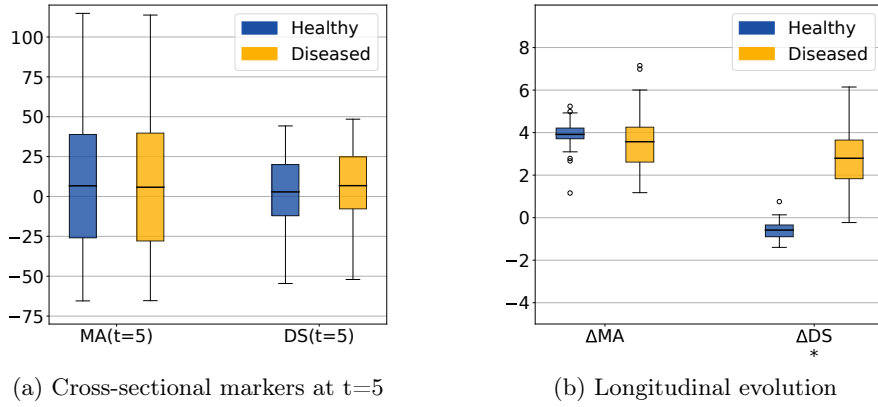


Figure 2.4: Evolution of the imaging biomarkers estimated on simulated data. MA=morphological age, DS=disease score. The longitudinal evolutions ΔMA and ΔDS are the differences between the two cross-sectional assessments i.e. $\Delta MA = MA(5) - MA(0)$. The star indicates that the difference between the healthy and the diseased subjects is significant (p-value < 0.05) for the unpaired t-test. By construction there is no difference between the two populations at t=0. The changes are generally underestimated but the longitudinal evolutions show the stability of the estimation despite a strong inter-subject variability.

and that the difference measured between two different time points gives a good estimate of the longitudinal evolution. In practice the estimations are slightly biased, for example the increase of morphological age $\Delta MA = MA(5) - MA(0)$ is expected to be equal to 5 for both populations while the mean of the estimation is equal to 3.92 for the healthy group and to 3.57 for the patients group. More importantly, the standard deviation is small in comparison to the standard deviation of the cross-sectional measurement ($\sigma_{\Delta MA} = 1.13$ while $\sigma_{MA(0)} = 47.8$). For the change of disease score ΔDS , the mean is equal to -0.6 for the healthy group and 2.8 for the patients group and it approximates the ideal expected values (respectively 0 and 5). The variance is also very small with respect to the cross-sectional one. In particular the difference between healthy and diseased subjects is clearly observed for the longitudinal evolution: the difference between the means is equal to 3.2 standard deviation for ΔDS . To summarize, the cross-sectional measurement gives a relatively stable assessment of the aging and disease progression and the markers’ evolution is strongly associated with the clinical diagnosis. Our generative model is able to explain most of the independently simulated changes.

2.4.2 Experiments on ADNI data

Longitudinal T1 sequences were obtained from the Alzheimer’s Disease National Initiative (ADNI) database. The ADNI was launched in 2003 with the primary goal to test whether magnetic resonance imaging (MRI), positron emission tomography

(PET), other biological markers, and clinical and neuropsychological assessment could be combined to measure the progression of mild cognitive impairment and early Alzheimer’s disease¹. Subjects are classified according to the evolution of their cognitive diagnosis. Three diagnoses are possible at each time point: normal, mild cognitive impairment (MCI) and Alzheimer’s disease. The subjects are also sub-classified according to the positivity of the beta-amyloid 1-42. We then have 6 distinct sub-groups: CN- (cognitively normal with negative $A\beta$), CN+ (cognitively normal with positive $A\beta$), MCIs- (MCI stable during the study time-window with negative $A\beta$), MCIs+ (MCI stable with positive $A\beta$), MCIC (MCI converter to AD) and AD (diagnosed with Alzheimer’s disease starting from the beginning). The table 2.1 sums up the demographic description of the population.

group	CN-	CN+	MCIs-	MCIs+	MCIC	AD
Number of subjects	108	69	96	120	228	203
Age at baseline	73.4 (5.6)	74.5 (6.5)	71.1 (7.7)	73.5 (6.6)	73.8 (7.1)	74.5 (7.7)
Gender (female)	47.2%	56.5%	47.9%	37.5%	41.7%	48.3%
Education (years)	16.4 (2.6)	16.2 (2.7)	16.2 (2.8)	16.4 (2.7)	16.0 (2.9)	15.0 (2.9)
ADAS13 at baseline	9.0 (4.0)	8.6 (5.0)	12.0 (4.9)	14.0 (5.4)	19.9 (6.7)	31.4 (7.3)

Table 2.1: Socio-demographic and clinical information of the study cohort. Standard deviations are shown in parentheses.

We estimate our template morphology and the template trajectories on a subset of subjects. In order to form this training set, we randomly selected 30 subjects from the CN- group and 30 from the AD group. To reduce the variability associated with the estimation of the model, these subjects were selected among the ones with strictly more than one followup acquisitions. In the following we distinguish between the *training* set of 30+30 subjects used to build the model and the remaining *testing* set (with in particular 78 CN- subjects and 173 AD subjects).

2.4.2.1 Estimation of the normal aging and the disease-specific template trajectories

The template anatomy is an average of the healthy subjects anatomies, so its age corresponds to the mean group age $\lambda_0 = 73.46$ y.o. The result of the estimation is shown in Figure 2.5. The estimated normal aging template trajectory is characterized mainly by ventricular expansion caused by the atrophy of the surrounding regions. Disease specific changes are widespread in the brain with a strong emphasis on the temporal areas.

The local atrophy can be measured by the divergence of the velocity field. The atrophy associated with each template trajectory is shown in Figure 2.6. For the normal aging, we see a well spread and mild atrophy pattern in the whole brain. The disease specific atrophy is particularly strong in the temporal area and is mainly

¹see www.adni-info.org for more information.

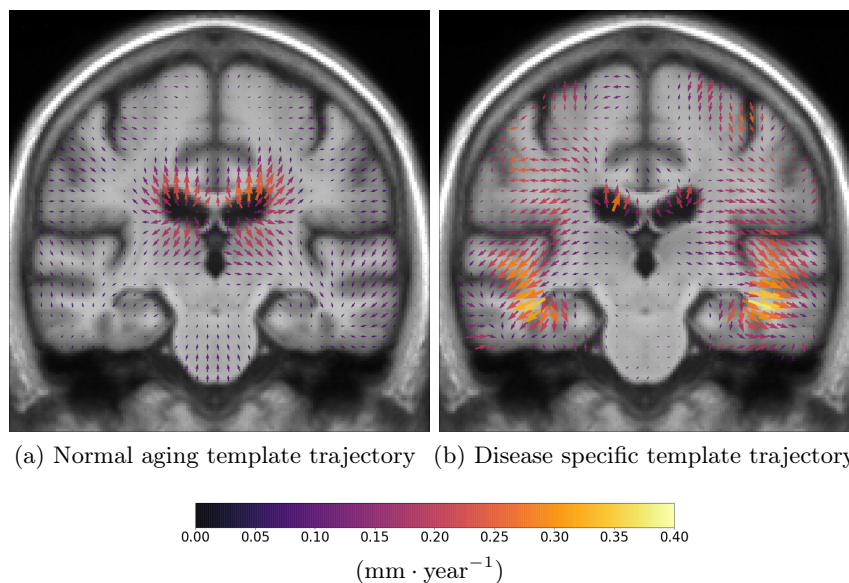
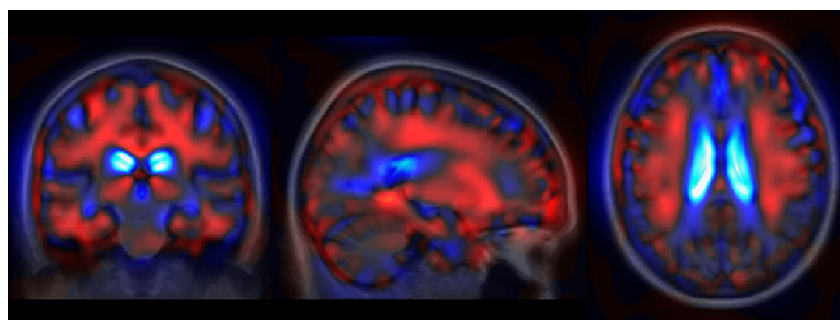


Figure 2.5: Template image and SVFs parametrizing the two template trajectories SVFs. Left: normal aging trajectory v_A showing a ventricular expansion related to a global atrophy. Right: disease specific trajectory v_D showing specific patterns, especially in the temporal lobes around the hippocampi areas. The color encodes the amplitude of the velocity at each position.

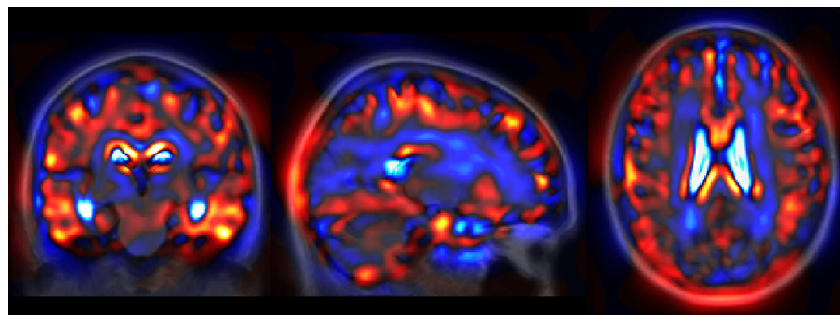
located near the cortical surface. The precise localization of the atrophy is always difficult with a morphometrical approach but the atrophy patterns are similar to what was already observed in the past for healthy subjects and AD patients (Hadj-Hamou, 2016).

These morphological evolutions can be compared to the normal aging model and the mean residual for the AD subjects in Lorenzi et al. (2015). The deformation characteristics and the magnitude of the atrophy are really similar. Our anatomy looks sharper which may be partially explained by the use of the template plane in the estimation. More importantly, differently from the work of Lorenzi et al. (2015), the joint estimation of a disease model in addition to the normal aging one provides us with a direct comparison between both processes.

The generative model can also be used to directly visualize the modeled morphologies. In particular the reference template plane \mathcal{T} described in Figure 2.2 can be sampled to shows the evolution of the template morphology in the two main directions. In Figure 2.7, we choose to represent the evolution over 20 years in both directions which is comparable to the longitudinal span of our data-set. Indeed the IQR of baseline age in the training set is equal to 9.3 years (and the total span is 29.9 years wide). The overlaid difference of intensity is used to show the changes at tissue boundaries. The global atrophy and the expansion of the ventricles is clearly visible for the aging evolution. The pathological changes are associated with smaller



(a) Atrophy along the normal aging template trajectory



(b) Atrophy along the disease specific template trajectory



relative volume change (year⁻¹)

Figure 2.6: Atrophy measured by the divergence of the SVFs parametrizing the two template trajectories.

structures but the shrinking of the hippocampi, the atrophy of the temporal lobes and also the widening of the sulci (related to the cortical thickening) are visible.

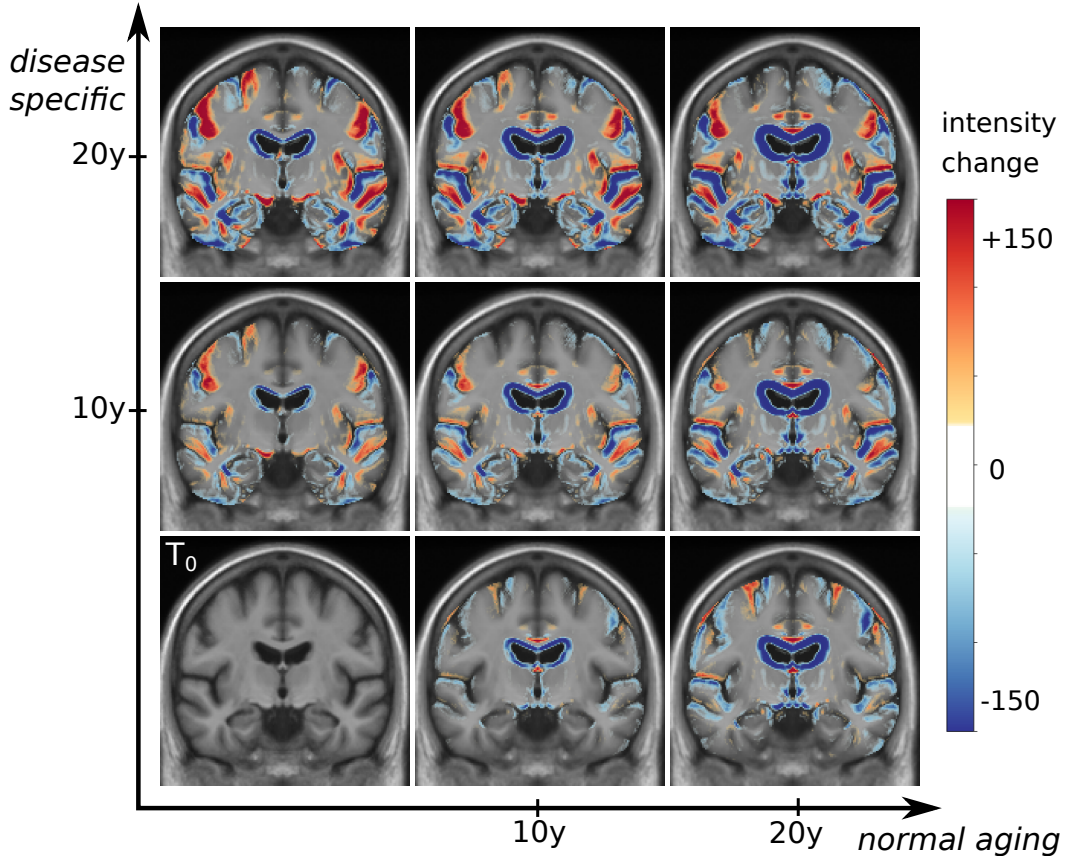


Figure 2.7: Representation of the 2D parametric template subspace generated using the template morphology T_0 and the two template trajectories v_A (horizontally) and v_D (vertically). In this figure, the bottom row correspond to a healthy evolution, and the diagonal (from bottom left to top right) to a typical pathological evolution. We also represent the voxel-wise intensity differences between the images and T_0 to highlight the boundary shifts between tissues and CSF.

2.4.2.2 Intra-subject variability of our progression markers

In [Lorenzi et al. \(2015\)](#), a morphological age similar to our measure was shown to be correlated with the chronological age and also that advanced AD stages were associated with “morphologically older” brains. To go further, we want to show that our proposed model represents also the aging at the individual level. For multiple acquisitions of a same subject, an aging measurement is expected to increase smoothly with time. If the subject is healthy, we can expect a linear increase with a slope of 1. We would also like to see an increase of disease score for the patients, while for the healthy subjects this marker should be stable and close to 0. Therefore,

the disease score is expected to specifically characterize the pathology (high and increasing disease score). The morphological age is expected to be more associated with the global evolution of the population over time and to be independent of the clinical condition.

The morphological age and the disease score are computed for each subject at each time point. Figure 2.8 shows the evolution of these cross-sectional measurements. First, at the population level, subjects are generally associated with a morphological age similar to their chronological age even though its variability is quite high. Second, for each subject, the evolution is mostly linear and the morphological age steadily increases. Third, the disease score is almost constant for healthy subjects and steadily increases for the AD subjects. Finally, we note that the AD subjects look older, age faster and, most of all, have a higher disease score than the healthy ones.

A linear random effects model can help us to quantify these observations. The model is fitted, for both morphological age and disease score, with fixed effects on age and sex and a random intercept and slope for each subject. The focus is set on the analysis of the regression for the CN and the AD groups. For each coefficient of the regression, the confidence intervals are given for one standard deviation of the estimation.

The model is first fitted to the morphological age measures in the CN group leading to a coefficient of 0.26 ± 0.11 for the fixed effect of age while the mean subject slope is 0.10 ± 0.02 . Both are significantly positive. In comparison to the same model without the random slope the relative improvement brought by the intra-subject linear evolution is significant by a large margin (p-value inferior to 10^{-6} for the likelihood ratio test). The regression has also a positive (but not significantly) coefficient for male subjects (1.81 ± 1.2) meaning that male morphologies looks older (similar to a 7 years shift). Concerning the disease score, we also observe a relatively good fit of the linear model. The evolution is generally slower with 0.12 ± 0.1 for the fixed effect of age and 0.12 ± 0.01 for the mean individual slope.

For the AD group, the linear model is also well adapted (p-values inferior to 10^{-6} for the likelihood ratio test). The main remark is probably that the intra-subject slopes are in average more important than for healthy subjects (around 0.52 ± 0.06 for the morphological age and 0.71 ± 0.05 for the disease score) while the fixed effect related to age of 0.23 ± 0.07 (for the morphological age) and 0.14 ± 0.04 (for the disease score) are more similar to the one observed previously.

2.4.2.3 Cross-sectional discriminating power

We want to study the relation between the observed disease progression and the proposed markers. We start with a discriminant analysis using only the first image available for each subject.

Figure 2.9 shows the distribution of the estimations for each group. These results are related to global differences in brain shape observable cross-sectionally between clinical groups. We see a gradual increase of both markers towards more advanced

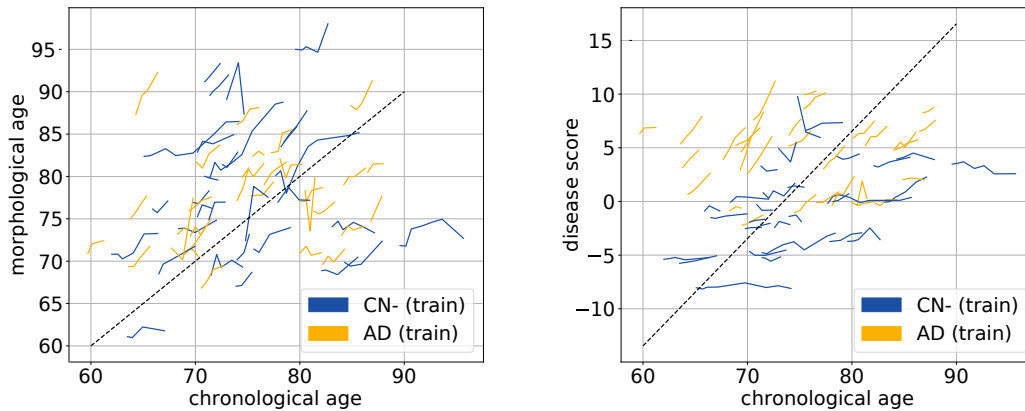


Figure 2.8: Evolution of cross-sectional markers for every subject of the two training sets. [Left] chronological age, the dashed line corresponds to the expected evolution of healthy subjects i.e. the morphological age is equal to the chronological age. [Right] disease score, the dashed line is the expected pathological evolution.

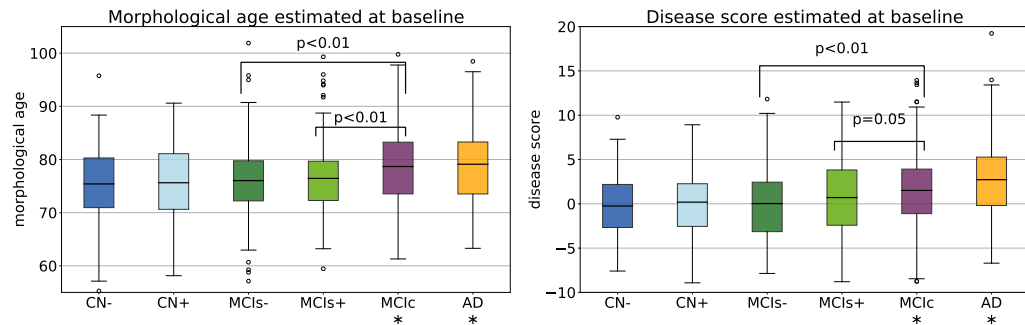


Figure 2.9: Box-plot of the group-wise markers estimated at baseline for the clinical groups. Stars below the name of the group indicate a significant difference to the CN- group for a t-test at the level 0.05. Both markers gradually increase towards more advanced disease states.

disease states. Significant differences in morphological age and in disease score are observed between the control group CN-(train) and both the MCIc and AD groups. Moreover, the difference between the MCI stable and the MCI converters is stronger for the morphological age while the disease score better differentiates MCI- and MCI+. As such, the morphological age is more associated with the general cognitive degradation while the disease score seems more correlated with more AD-specific biomarkers.

We also perform a simple linear classification task between the MCIs and MCIc groups using this two cross-sectional markers. A SVM linear classifier is fitted to the full data-set to perform the binary classification². The error penalty weights are

²The classification task was performed using the SVC module of the scikit-learn python li-

adjusted between the two classes to balance the trade-off between false positive and false negative rate. The mean classification accuracy using a 10-folds cross-validation scheme is equal to 0.59. The linear decision function is equivalent to the projection on the SVF $v_A - 0.003v_D$, so the differences between MCIs and MCIc subjects is, in our model, only associated with the aging trajectory while the disease specific changes do not seem to have an impact before the conversion. For comparison, for the same experiment between CN and AD, the linear classifier corresponds to the projection on $v_A + 0.49v_D$ so approximately $(v_{bn} + v_{ad})/2$, i.e. the mean evolution of the whole population. Of course, in both cases we do not reach the performance of state-of-the-art dedicated algorithm but it allows us to see how both markers are associated with the diagnosis. Moreover this discriminant approach could be extended by using information in a subset of targeted areas.

2.4.2.4 Regional analysis of the progression

In the context of Alzheimer’s disease, specific morphological changes are known to be non-uniformly distributed. This spatial information can be taken into account in our model using a regional segmentation. In this section we focus on the temporal area which is often associated with Alzheimer’s disease (Double et al., 1996). Using the AAL atlas (Tzourio-Mazoyer et al., 2002), we segment the temporal lobe of our template anatomy. The mask is then used to compute the regional morphological age and disease score for each subject. These markers only encodes the morphological differences in this specific area visible at the time of acquisition (i.e. at baseline in this case). Results are shown in Figure 2.10.

The region is more adapted to the disease score than to the morphological age model. Indeed, for a healthy subject the deformations in this area are really small. However the choice of a disease-adapted region is improving the performance of the disease score. It is now able to capture early specific changes and the difference between CN- and CN+ is significant.

2.4.2.5 Longitudinal evolution of the markers

To explore more in details the longitudinal evolution of these markers, a linear model is fitted to the individual evolutions. The intercept can be interesting as it aggregates the measure at every time point and helps reduce the noise but more importantly the slope can be very informative. Results are shown in Figure 2.11 for the whole brain markers.

A progressive evolution, from CN- to AD, is visible for the morphological age with subjects evolving faster and faster. Concerning the disease score, the evolution is almost negligible for CN- and MCIs- and relatively slow for CN+ and MCIs+ while the changes are clearly visible for MCIc and AD. Significant differences are visible between healthy subjects (CN-) and MCIc or AD subjects or even between MCI stable and MCI converters, but also between MCIs- and MCIs+ (or more generally

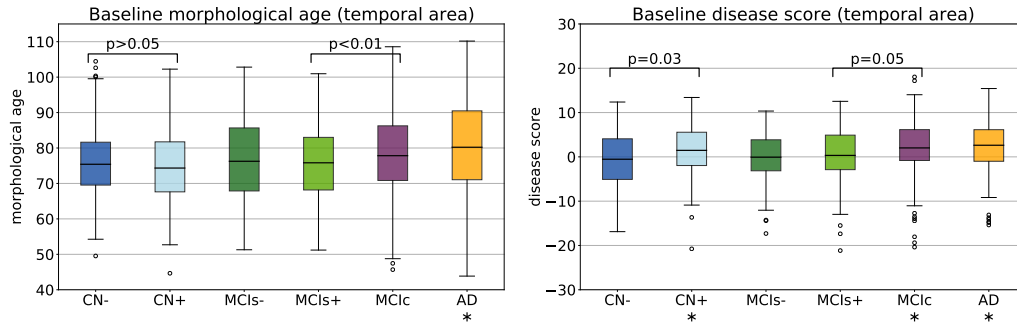


Figure 2.10: Box-plot of the markers estimated in the temporal lobes. Stars below the name of the group indicate a significant difference to the CN- reference for a t-test at the level 0.05. The area known to be related to the AD makes the disease score estimation less sensitive to the overall noise.

between subjects with negative amyloid or positive amyloid marker). This may indicate that our measures are able to capture the global progression of the disease. The changes are larger for diagnosed patients but similar patterns of evolution are observed in the early stages of the disease. A significant difference is also observable between the CN- and CN+ group for the temporal disease score slope.

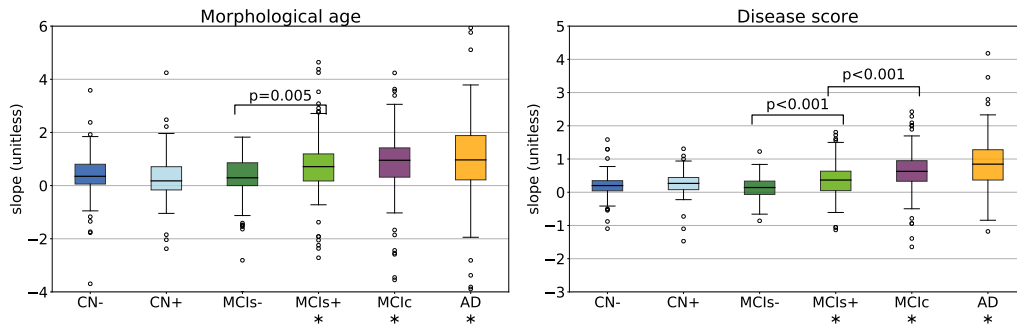


Figure 2.11: Box-plot of the rate of evolution of the markers computed using individual linear regressions. Stars indicate a significant difference to the CN- reference for a t-test at the level 0.05. Top row shows the results for the whole brain while the bottom row shows the result for the temporal lobe only. A gradation is visible from the CN to the AD subjects.

The disease score evolution is close to zero for the healthy subject and close to one for the pathological one but more generally, the slopes are in average smaller than their expected values. For example the average disease score evolution in the AD group is equal to 0.82 and this discrepancy is particularly important for the morphological age slope of the CN- group that is only equal to 0.33 (instead of 1). This bias was already observed previously and can be in part explained by the estimation procedure (see 2.6.2).

2.4.3 Generating diagnosis driven morphological evolution

One of the main advantages of our model is its ability to be generative. From a pair of biomarkers (morphological age, disease score) we can generate a corresponding morphology and to deform a specific subject anatomy in the directions defined by the template trajectories. In this section, the model is used to generate plausible morphological evolutions of a subject, for several diagnosis condition, and compare them to the observed one.

2.4.3.1 Modeling the group evolution

In what follows, each subject is associated with a pair (position, speed) in the marker plane using the coefficient of the linear regression. For each group, a continuous vector field is regressed from these observations. This field estimates the local average speed at every point. In order to keep a local consistency in the markers space, we use a kernel ridge regression with an RBF kernel. The spatial scale is set to 10 years, for both the morphological age and the disease score axis, to get large scale patterns despite the high inter-subject variability. The regularization weight, which does not seem to have a large effect on the result, is set to 1. Results are presented in Figure 2.12 for the CN (i.e CN-, CN+ together), MCI, MCIc and AD groups. The figure is centered on the high data density domain (as expected the extrapolation can be less reliable in lower density sectors).

Differences in amplitude, i.e. speed of evolution, and orientation are clearly visible between the groups and are in agreement with the linear regression results shown in the previous section. In particular we see a progressive amplitude increase from the CN group to the AD group.

These diagrams also help to describe the variability within the same clinical group. For the CN group we can distinguish between the low disease score and low morphological age area (in the bottom left) where in average the changes are negligible, and the rest where there is a slow horizontal evolution. This pattern suggests that the healthier and younger subjects are morphologically stable and do not show the same visible aging process. The MCI stable evolution is relatively uniform and in average with slightly larger amplitudes but overall similar to the CN one.

The MCI converters however show a stronger and more vertical evolution. We should also note that subjects with high morphological age and low disease score (bottom right) seem to follow a different, more horizontal evolution implying a fast morphological aging but less important disease specific changes. The AD group confirms this trend and in fact MCIc and AD look very similar. The mean evolution is strong and more vertical. A main evolution is visible from bottom left to top right with a slightly more horizontal part in the middle giving this global tangent-like aspect. Beside, a horizontal evolution, similar to what was observed for the MCIc model, is also visible in the bottom right. This difference of evolution suggests a possible stratification of the disease in two sub-categories.

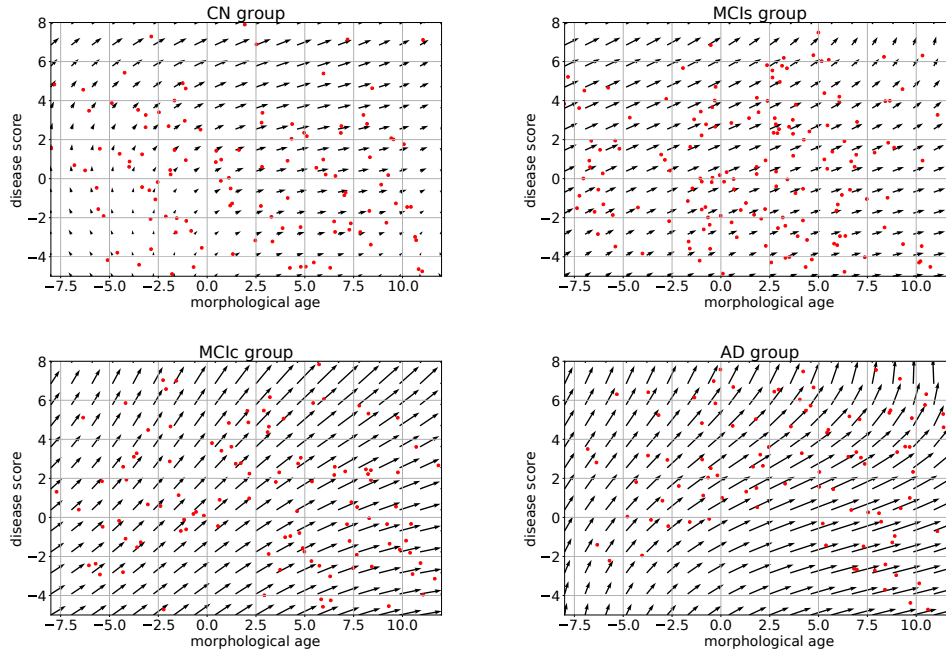


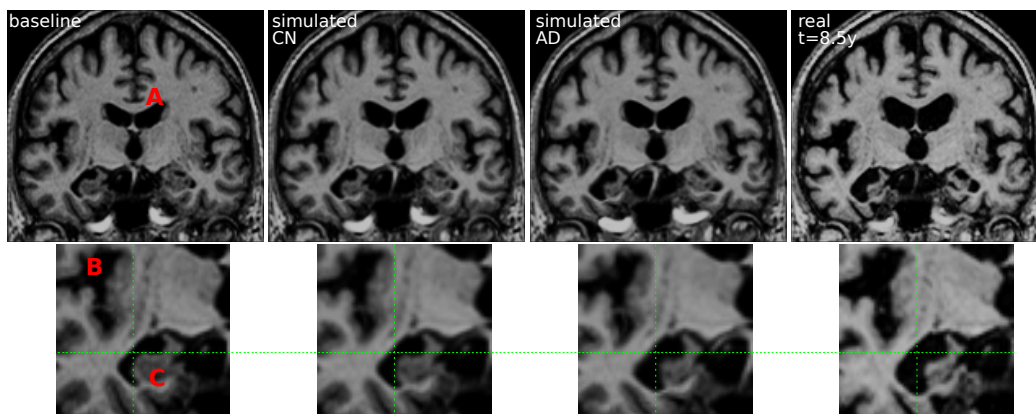
Figure 2.12: Results of the kernel ridge regression for the markers’ evolution for the CN, MCIs, MCIc and AD groups. It shows the regressed vector field with data points shown in red. Amplitude and orientation variability is visible between the groups (with stronger and more vertical evolution for AD than for CN) but also within each group giving non-linear morphological evolutions.

2.4.3.2 Generating a subject specific evolution

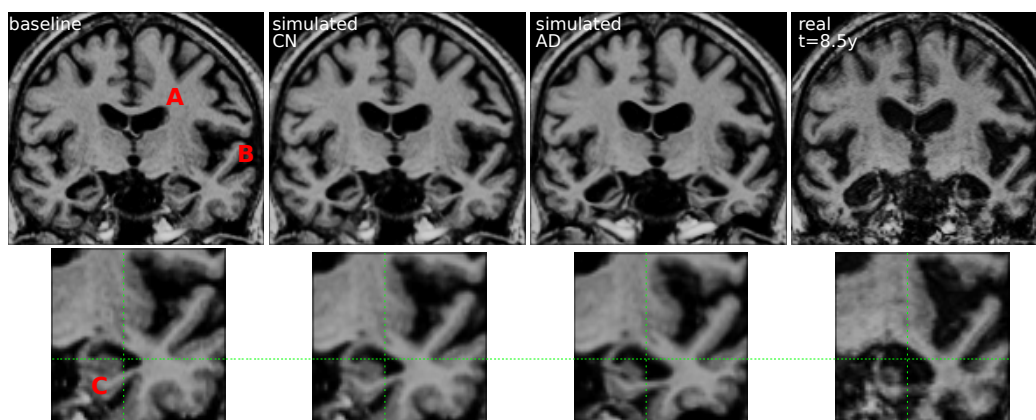
The regression model can then be used to simulate the evolution of a given subject. We choose here to model the evolution of MCIc subject in order to predict the changes subsequently observed around the time of diagnosis. Several evolutions were computed for the subject 0361. This subject is chosen among the MCIc subjects because it has the longest time interval, here 8.5 years, between the first and the last acquisition. From the starting point, the markers at $t=0$, we integrate a trajectory using the speed given by the regression model at each point. Each clinical group is associated with its own model and then a different trajectory is computed for each diagnosis.

These markers changes can be directly translated in brain images to visualize the morphological evolution. Figure 2.13 shows the results for the end point of the trajectories for the CN and AD models. Images are generated by deforming the baseline image using the simulated deformation transported in the subject space. A bias correction is applied to the markers’ estimation using the value estimated on the synthetic dataset (the measured changes only correspond to 80% of the expected value).

For both subjects, the cortical atrophy, in particular in the temporal lobe, is



(a) Subject 0361



(b) Subject 0566

Figure 2.13: Evolution of the two subjects over 8.5 years. Both subjects were diagnosed MCI (for 5y and 3y) then AD. From left to right: (1) real image at baseline, (2) simulated image from the baseline image using the healthy (CN) evolution, (3) using the pathological (AD) evolution, (4) real image at $t=8.5y$. The second row zooms in the most interesting areas between the ventricle (A), the lateral sulcus (B) and the hippocampus area (C). In both cases, even if the simulated changes do not match the full extend of the real case, the atrophy is visible in the sulcus and the hippocampus and there is a difference in shape and size of the ventricles.

clearly visible in the real image and to a lesser extent in the AD simulation. The expansion of the ventricles is always clearly visible, but even in the AD simulation, the volume change is inferior to the observation. Subject 0566 is, from the baseline, generally more atrophied and by comparison it makes some real changes less visible (in particular the sulci widening). This difference in baseline anatomy is interesting here for two reasons: first the deformations match correctly the anatomical structures (the hippocampi for example) despite the anatomical variability, second it generated two different evolutions for the same diagnosis. For the AD simulation, the deformation seems stronger in the temporal lobe and relatively weaker in the ventricles for the second case (0556). These evolutions are learned from population trends and even if they are not predictive for a particular subject, they are not aberrant in comparison with the real evolutions. Overall, the morphological changes simulated looks realistic even if they do not match perfectly the observed changes. Other aspects of the evolution are hard to quantify and often poorly documented. For example the evolution of the shape of the ventricles of subject 0361 is different in the three images. It may be related to different spatial distribution of the degeneration in the brain and inherent mechanical constraints. We also observe a global motion towards the bottom of the temporal lobe and a local rotation in the image for the real case and the AD simulation.

2.5 Discussion and perspectives

In this work, we proposed a novel deformation-based approach to measure the progression of normal and pathological processes from their effects on brain morphology. In the context of Alzheimer’s disease, it provides a simple description of the brain morphological evolution for elderly patients using only two degrees of freedom: an aging measurement and a disease score. The advantages come from three main properties.

First, we disentangle the aging and the disease progression using interpretable image-based biomarkers. Second, these markers are cross-sectional assessments and are consistent for intra-subject longitudinal analyses. They can be seen as alternative aging measurements compatible with ongoing biological processes. In particular the disease specific evolution appears to be associated with a positive amyloid marker even in prodromal stages. Third, we show that the markers and the generative model can be used in a personalized image simulation setting. It allows us to generate smooth and realistic evolutions for several diagnosis conditions.

Biological (or here morphological) age estimates were proven to be interesting to analyze the patient condition; here, in addition, the disease score is used to get a simple marker of the disease progression. The joint modeling gives a more complete description of the disease progression than a brain-age metric. The evolution is not seen as a simple accelerated aging process or the divergence from the normal evolution. On the contrary, it seems possible to capture the general worsening with the morphological age while the disease score measures an additional evolution that

is more specific to the disease. Both patterns of evolution appeared to be related with the development of Alzheimer’s disease and this approach provides an intuitive interpretation and a simple decomposition of the morphological changes observed in mass-univariate morphometric approaches. Further analysis should be conducted to analyze the relation of these complementary patterns to clinical and cognitive variables.

Moreover the ideas behind the decomposition and the longitudinal/cross-sectional estimation of the parameters could be used in a different setting: the geodesic description is much more general than the SVF framework applied to structural MRIs. Other geodesic parametrizations could be more practical in another setting and even for the same image data, an LDDMM approach, for example, would generally concentrate the deformation more on the high intensity gradient interfaces while the SVFs model more spatially diffuse deformations.

Several approaches have been developed to model the longitudinal shape evolution using geodesic modeling (Singh et al., 2013). Here we suggest that decoupling the estimation of individual trends using parallel transport could allow a more accurate and non-linear description of the disease while keeping a similar hierarchical mixed-effect structure.

As could be expected, a single variable is not enough to precisely describe the morphological changes that can occur while aging. Likewise, the inter-subject variability limits the accuracy of the modeling of the evolution of individual subjects. Some limitations come from the error and the approximation in the estimation of the model, others are related to aspects that are not taken into account and would require to modify the approach.

2.5.1 Approximations in the model estimation

One of the limitations of our approach is the difficulty to accurately estimate the markers. For example, some time points look like outliers when we perform the estimation for successive images of the same subject. This might be due to MR distortions that cannot be completely corrected by the registration. As a result, some artifacts are still visible in the estimated deformations. A better understanding of the effects of these distortions on the registration results would be useful to improve our additive deformation noise model. A first step could be to work with data where the ground-truth deformation is more controlled (scan-rescan images for example).

In this work, we used an orthogonal projection using the L_2 scalar product in the SVF space to define the subject specific deformation. This choice is arbitrary and using another region of interest could change the result. A possible alternative would be to match the subject with the closest morphology in the model for a metric more adapted to deformations. We could also try to decorrelate the two information, using ICA for example.

We saw, using synthetic and longitudinal data, that we were under-estimating the evolution speed. This bias is partially explained by the estimation procedure that does not take into account the uncertainty on the population parameters of the

model (see 2.6.2). This problem cannot be easily solved for two reasons: first the bias caused by the estimation of the template anatomy using the healthy controls groups is hard to quantify, second an unbiased estimation would be far more complex and would not be possible without the full knowledge of the training set for every new subject. Another source of error can be the registration algorithm: the inter-subject registrations are larger than the intra-subject ones and no transitivity is guaranteed.

2.5.2 Limitations of the description

Probably the main limitation of our approach is the inter-subject variability of the markers. The markers are very stable in longitudinal but the anatomical variability is not completely absorbed by the biomarker estimation and they are sensitive to the subject differences. The use of an explicit or even implicit model of the shape of a normal brain would help to makes them more specific to time related changes.

We could also note that the template anatomy is estimated using healthy subjects only. This modeling choice may introduce a bias towards the healthy population. We argue that the explicit modeling of the disease in our model could help mitigating this issue. Moreover, this choice makes the origin of the morphological age and the disease score simpler to define. Finally, pulling all the subjects could also enable a more accurate estimation (Marchewka et al., 2014) but the relatively large and homogeneous populations involved does not make it critical here.

In this context, the use of a single reference anatomy to parametrize the template space could also be discussed. Here, for example, it introduced a bias toward a certain age because of the way we composed deformations. A multi-atlas approach could be a better solution if a single anatomy is not enough to describe our population (Blezek and Miller, 2007; Sabuncu et al., 2009). And we could do something similar to what was proposed by Rohé et al. (2016) to intricate the registration and the template subspace prior.

Regarding the estimation of the template trajectories using individual longitudinal SVFs transported in the template, potential bias associated with the asymmetrical use of the baseline image (in processing and in modeling) has been highlighted (Reuter et al., 2012) and the imbalance in number of time points or total longitudinal span should be taken into account for a more reliable estimation. However, Hadj-Hamou et al. (2016) has shown that the asymmetry caused by the non centrality of the time point in the longitudinal sequence is not completely relevant in the stationary velocity field framework because the SVF are expressed in Eulerian coordinate and should be identical at all time points.

2.5.3 Perspectives

The use of segmentation to compute the progression markers in multiple regions would be another way to extend the description. We showed that using a segmentation of the temporal lobes could tighten the link between the morphological markers and early clinical conditions. However the question of the regional interactions is

not addressed in this work and the spatial analysis of brain deformation remains a research topic.

Our model of the morphological evolution is generic and of low dimensionality. Consequently it only partially captures the changes related to AD and, in practice, the subject specific field w_r^k is actually encoding changes that are related to the disease and its variability but are not currently modeled. To enhance the description, the model should integrate the observed and latent clinical condition of the subjects. A description that goes beyond the global cognitive diagnosis would be interesting to pursue in the temporal description of the morphological changes in order to better describe the disease progression and capture this evolution from a healthy state to a pathological one.

Beside, the observed longitudinal evolution is also influenced by various factors such as sex, genetic, other pathologies or even the image acquisition protocol. The ability to handles these covariates has shown its utility to better describe the brain morphological evolution (Muralidharan et al., 2016).

These extensions would also improve the generative aspect. Coupling our approach with a proper disease progression model, and using a mixture model for the trajectories, would enable the generation of morphological trajectories in a more diverse setting to explore and sample the range of possible evolutions.

2.6 Appendix

2.6.1 Validation on a synthetic dataset

The regional values of atrophy set are given in Table 2.2. For each subject, the values are sampled around these means with a 5% standard deviation.

brain area	mean pathological (in %)	mean healthy (in %)
white matter	1.0	0.8
cortex	3.0	0.4
hippocampi	5.2	1.0
amygdalae	5.2	1.0
entorhinal cortex	6.5	0.7
temporal poles	6.2	0.6
other areas	0.0	0.0

Table 2.2: Specified mean regional atrophy for the healthy and the pathological evolutions. The goal is to get simple but realistic atrophy patterns. It should be noted that the atrophy is specified using the divergence of the SVF in the area. The local volume changes are computed using a spatio-temporal integration scheme.

The comparison with the values obtained after simulation or estimated through registration are shown in Figure 2.14. The estimation can be biased by the spatial regularization and the loss of information in intensity homogeneous areas. The

relative changes is however similar between the two populations and the method is by consequence adapted to compare the two evolutions.

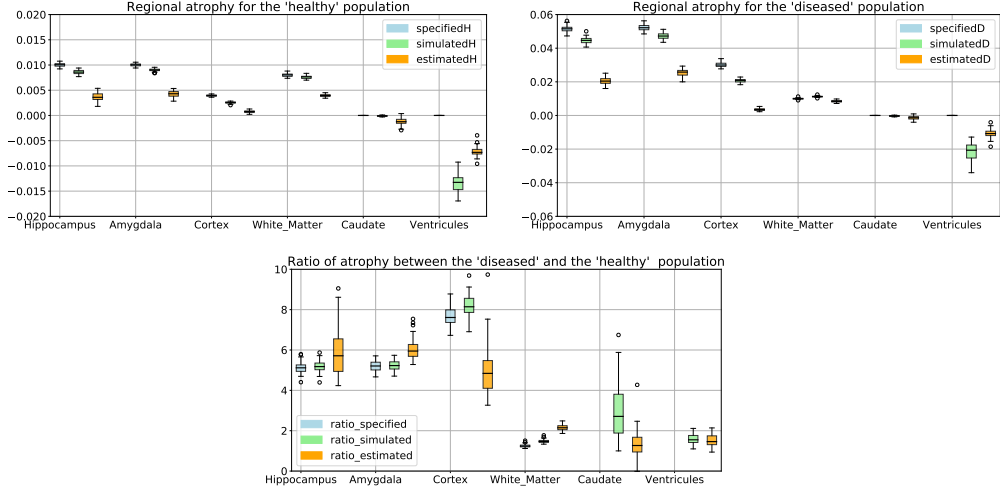


Figure 2.14: Comparison of prescribed, simulated and estimated atrophy values for several brain regions. The difference between prescription and simulation can be explained by numerical approximations in the biophysical model while the estimation is biased due to spatial regularization. However the estimation bias is consistent across regions and subjects.

2.6.2 Bias on the estimated template trajectories

The norms of the SVFs parametrizing the template trajectories have an effect on the normalization of the individual biomarkers estimation. We estimate here the bias on the the norm of the estimation relatively to the estimated norm.

For the normal aging trajectory we have:

$$\begin{aligned}
 E(\|\hat{v}_A\|^2) &= \frac{1}{|\mathcal{G}_h|^2} E(\langle \sum_{k \in \mathcal{G}_h} v^k | \sum_{k \in \mathcal{G}_h} v^k \rangle) \\
 &= \frac{1}{|\mathcal{G}_h|^2} \sum_{i,j \in \mathcal{G}_h} E(\langle s_{MA}^i v_A + s_{DS}^i v_D + v_r^i | s_{MA}^j v_A + s_{DS}^j v_D + v_r^j \rangle) \\
 &= \frac{1}{|\mathcal{G}_h|^2} \sum_{i,j \in \mathcal{G}_h} E(s_{MA}^i s_{MA}^j) \|v_A\|^2 + E(s_{DS}^i s_{DS}^j) \|v_D\|^2 \\
 &\quad + E(s_{MA}^i s_{DS}^j + s_{DS}^i s_{MA}^j) \langle v_A | v_D \rangle + E(\langle v_r^i | v_r^j \rangle)
 \end{aligned}$$

Assuming that the subjects are independent and identically distributed:

$$\begin{aligned} E(\|\hat{v}_A\|^2) &= \|v_A\|^2 + \frac{1}{|\mathcal{G}_h|^2} \sum_{i \in \mathcal{G}_h} \text{Var}(s_{MA}^i) \|v_A\|^2 + \text{Var}(s_{DS}^i) \|v_D\|^2 \\ &\quad + 2E(s_{MA}^i s_{DS}^i) \langle v_A | v_D \rangle + E(\|v_r^i\|^2) \\ &= \left(1 + \frac{\text{Var}(s_{MA}^h)}{|\mathcal{G}_h|}\right) \|v_A\|^2 + \frac{\text{Var}(s_{DS}^h)}{|\mathcal{G}_h|} \|v_D\|^2 + \frac{E(\|v_r^h\|^2)}{|\mathcal{G}_h|} \end{aligned}$$

And similarly for $\|\hat{v}_D\|$:

$$\begin{aligned} E(\|\hat{v}_D\|^2) &= E\left(\left\|\frac{1}{|\mathcal{G}_{ad}|} \sum_{k \in \mathcal{G}_{ad}} v^k - \hat{v}_A\right\|^2\right) \\ &= E(\|\hat{v}_A\|^2) + E\left(\left\|\frac{1}{|\mathcal{G}_{ad}|} \sum_{k \in \mathcal{G}_{ad}} v^k\right\|^2\right) - 2E\left(\left\langle \frac{1}{|\mathcal{G}_{ad}|} \sum_{k \in \mathcal{G}_{ad}} v^k | \hat{v}_A \right\rangle\right) \\ &= \left(1 + \frac{\text{Var}(s_{MA}^{ad})}{|\mathcal{G}_{ad}|}\right) \|v_A\|^2 + \left(1 + \frac{\text{Var}(s_{DS}^{ad})}{|\mathcal{G}_{ad}|}\right) \|v_D\|^2 + 2 \langle v_A | v_D \rangle \\ &\quad + \frac{E(\|v_r^{ad}\|^2)}{|\mathcal{G}_{ad}|} + E(\|\hat{v}_A\|^2) - 2(\|v_A\|^2 + \langle v_A | v_D \rangle) \\ &= \left(\frac{\text{Var}(s_{MA}^{ad})}{|\mathcal{G}_{ad}|} + \frac{\text{Var}(s_{MA}^h)}{|\mathcal{G}_h|}\right) \|v_A\|^2 + \left(1 + \frac{\text{Var}(s_{DS}^{ad})}{|\mathcal{G}_{ad}|} + \frac{\text{Var}(s_{DS}^h)}{|\mathcal{G}_h|}\right) \|v_D\|^2 \\ &\quad + \frac{E(\|v_r^{ad}\|^2)}{|\mathcal{G}_{ad}|} + \frac{E(\|v_r^h\|^2)}{|\mathcal{G}_h|} \end{aligned}$$

The terms $\text{Var}(s_{MA})$, $\text{Var}(s_{DS})$ and $\|v_r\|^2$ are related to the individual variability that is not modeled and to the noise in the estimated deformations. We empirically estimate the variances in the training population and we get that, for our training dataset, $\|\hat{v}_A\|^2 \approx 1.65\|v_A\|^2$ and $\|\hat{v}_D\|^2 \approx 1.21\|v_D\|^2$. We should however note that the same subjects from \mathcal{G}_h are used to estimate the template anatomy T_0 and the normal aging template trajectory v_A . Therefore, we are certainly underestimating the bias coming from the intra-subject morphological variability for this population.

This bias has a direct influence on the markers estimations. Indeed for a subject k , the markers are solution of the linear system 2.9 involving $\|v_A\|^2$, $\|v_D\|^2$ for which the biased estimators are used.

Exploratory morphometry for the MAPT clinical trial

Contents

3.1 Introduction	44
3.1.1 The MAPT study	44
3.1.2 Measuring the evolution of the brain morphology	45
3.2 Method	46
3.2.1 Participants	46
3.2.2 Regional atrophy and cortical thickness assessments	46
3.2.3 Statistical analysis of the deformations	46
3.3 Results	49
3.3.1 Cognitive, behavioral and functional assessment	49
3.3.2 Segmentation-based statistical analysis of the interventions' effect	49
3.3.3 Deformation-based statistical analysis of the interventions' effect	51
3.3.4 Description of the MDI effect on the brain morphology	51
3.3.5 Relation to cognitive and behavioral scores	55
3.4 Discussion	56
3.4.1 Interpretation of the morphological changes	57
3.4.2 Limitations	57
3.4.3 Perspectives	58
3.5 Conclusion	58

Chapter 2 was dedicated to the description of morphological changes associated with known clinical conditions. In this chapter, deformation modeling is used in more exploratory perspective: the assessment of changes associated with a clinical variable. This work is conducted in the context of the Multidomain Alzheimer Preventive Trial (MAPT).

The MAPT was designed to assess the effect of omega-3 supplementation and a multidomain intervention (physical activity, cognitive training and nutritional advice) on cognitive decline of people with subjective memory complaint. In term of cognitive testing, no significant effect on cognitive decline was found. In this work, we evaluate the effect of the interventions on the evolution of the brain morphology

using the MR images acquired in the study. The morphological changes are assessed from volume measurements of regions of interest and a voxel-wise deformation-based approach.

We find that the multi-domain intervention is associated with a significant effect on the 3-year morphological evolution. The voxel-wise deformation-based approach shows that the differences are mainly located in the left peri-ventricular area next to the temporoparietal junction (TPJ). These morphological changes correspond to a slower morphological evolution and are correlated with a better performance in cognitive assessments. These results could not be observed using the volumetric morphometry approach. We argue that the use of neuroimaging could help better define early intervention strategies that are effective to delay cognitive decline and dementia.

This chapter is based on the following article submitted to *Neurobiology of Aging* (under review): *Voxel based assessments of treatment effects on brain evolution in the MAPT cohort*, Raphaël Sivera, Nicolas Capet, Valeria Manera, Roxane Fabre, Marco Lorenzi, Hervé Delingette, Xavier Pennec, Nicholas Ayache, Philippe Robert

3.1 Introduction

3.1.1 The MAPT study

Subjective memory complaint (SMD) in individuals with unimpaired performance on cognitive tests is regarded as a possible risk factor of Alzheimer's disease (AD) (van Oijen et al., 2007). For this reason, people with SMD represent a promising population to put in place early interventions, and prevention strategies (Andrieu et al., 2008). Current evidence suggests that multi-domain interventions targeting nutrition, physical exercise, cognitive activity and social stimulation may improve cognitive health (Kivipelto et al., 2018) in older adults with SMD (Ngandu et al., 2015). For instance, an increase in omega-3 polyunsaturated fatty acid intake has been proposed as a possible intervention for preventing or delaying age-related cognitive decline (Dacks et al., 2013).

The Multidomain Alzheimer Preventive Trial (MAPT) was a 3-year multi-centric randomized placebo-controlled trial designed to assess the effect of omega-3 supplementation and a multidomain intervention (MDI) on cognitive decline of patients SMD. The multidomain intervention consisted of regular group sessions focusing on physical activity, cognitive training, nutritional advice and social stimulation and an individual preventive consultation to follow each participant. The supplementation in omega 3 polyunsaturated fatty acids was tested in comparison with a placebo. 1680 Subjects were randomized in four parallel groups including three intervention groups (omega-3 alone, MDI plus placebo, MDI plus omega-3) and a placebo control group. All participants in the MAPT study underwent extensive and longitudinal clinical testing, including classical clinical scales and tasks to assess the physical and cognitive functioning, the autonomy in activities of daily living, and the presence cognitive and behavioral disturbances. Furthermore, a sub-sample of participants

underwent structural magnetic resonance imaging (sMRI) both at baseline and at the end of the 3-year trial. A complete description of the trial is published in [Vellas et al. \(2015\)](#).

[Andrieu et al. \(2017\)](#) studied the effects of the treatments on the cognitive evolution, as indexed by a combination of classical clinical and cognitive tests, and found no significant differences between any of the three intervention groups and the placebo group. In this work, we want to investigate the effect of the interventions on the brain morphology.

3.1.2 Measuring the evolution of the brain morphology

Structural magnetic resonance imaging (sMRI) has been proven helpful to characterize neuropsychiatric conditions ([Bron et al., 2015](#); [Jack et al., 2016](#); [Lorenzi et al., 2017](#)) or to assess the efficacy of specific treatments ([Lieberman, 2005](#)).

In particular, knowledge of the spatial pattern and rate of decline of cerebral brain volume provide valuable information for detecting early neurodegenerative disease such as Alzheimer’s disease at different stages of disease progression ([McDonald et al., 2009](#)). For example, atrophy in Alzheimer’s disease is not uniform across brain regions, nor does it follow a linear trajectory. The atrophy rate is greater in early disease stages within the medial temporal cortex, and in later stages within prefrontal, parietal, posterior temporal and cingulate cortex ([Risacher et al., 2010](#)). This non linear behavior of the atrophy is also characterized by an acceleration phase during the early stages of the disease, specifically hippocampal and entorhinal atrophy in mild cognitive impairment ([Sabuncu et al., 2011](#); [Devanand et al., 2007](#); [Colliot et al., 2008](#)). However, to our knowledge, there is no strong evidence that measure of regional brain atrophy is relevant to predict symptomatic change in patients with cognitive complaints and normal neuropsychological tests ([Saykin et al., 2006](#); [Sluimer et al., 2008](#)).

These usual morphometric features are generally obtained from *a priori* defined regions of interests (ROIs) such as the hippocampi or the ventricles. They offer valuable quantitative assessments but can be limited and many alternatives have been proposed during the last decades to assess the local differences in brain tissues composition and shape ([Ashburner and Friston, 2000](#); [Qiu et al., 2007](#); [Hutton et al., 2009](#)).

Deformation based approaches describe the voxel-wise shape remodeling ([Ashburner et al., 1998](#)). Anatomical differences and morphological evolutions are estimated using registration and modeled by diffeomorphic deformations. These transformations of the space give us direct information on the visible evolution of the brain. These approaches can be powerful describe the aging brain evolution ([Sivera et al., 2019b](#)) and to explore the potential morphological changes to suggest morphological correlates ([Schmidt-Wilcke et al., 2009](#)).

3.2 Method

3.2.1 Participants

Subjects of the MAPT cohort with two MRI acquisitions at baseline and 36 months are included. The total sample includes 377 subjects, randomized in the MDI plus omega-3 group (N=98), the MDI plus placebo group (N=100), the omega-3 alone group (N=87), or the placebo control group (N=92). The demographic features and the genetic, cognitive and functional profile of subjects at baseline are reported in Table 3.1. One subject was later excluded during processing due to a failure of execution of the processing pipeline.

In order to verify that the 3-year evolution of the subjects cognitive, behavioral and functional profile in our sub-sample (N=377) is comparable to the original study population (N=1680), we submitted the evolution of the scores (results at M36 minus results at baseline) of each of the available tests to separate ANOVAs with Group (MDI plus omega-3, MDI plus placebo, omega-3 alone, placebo) as between-subject factor.

3.2.2 Regional atrophy and cortical thickness assessments

First, the longitudinal morphological evolution is evaluated using cortical thickness and sub-cortical volumetric measures. The data is processed using the longitudinal pipeline available in the FreeSurfer software (Reuter et al., 2012). We decided to focus our analysis on the longitudinal evolution of the regions of the brain well known to be atrophying in dementia from early stages (MCI) to severe stages (Whitwell et al., 2007). The evolution is measured between the baseline and the 36-month follow-up.

For the volumetric information 24 regions are selected and total left plus right changes are computed for 9 of these regions. The cortical thickness is evaluated in 34 regions of each hemispheres. Similarly the left/right averages are added to the analysis. We also evaluate the effect on the total white and grey-matter volume and the brain parenchyma fraction (BPF). In total the evolution is characterized by 142 morphological variables (the extensive list is available in supplementary material) in order to get a exploratory overview of the possible morphological changes.

The difference between the treatment groups are tested using an ordinary linear model with no cross effect. A non-parametric re-randomization test is performed to correct for multiple comparisons.

3.2.3 Statistical analysis of the deformations

3.2.3.1 Modeling the brain evolution

The deformation-based approach used in this work is based on an efficient parametrization of diffeomorphisms using stationary velocity fields (Arsigny et al., 2006). In this setting, smooth and reversible deformation that spatially align the anatomies are estimated through non-linear registration (Lorenzi and Pennec, 2014)

	Multidomain plus polyunsaturated fatty acids (n=98)	Multidomain plus placebo (n=100)	Polysaturated fatty acids (n=87)	Placebo (n=92)	Overall (n=377)
Age, years	74.96 [4.22]	74.34 [3.84]	74.67 [3.68]	74.48 [4.03]	74.61 [3.95]
Female sex	59 (60.2)	69 (69.0)	57 (65.5)	57 (62.0)	242 (64.2)
Education					
No diploma or primary school certificate	16 (16.5)	14 (14.0)	14 (16.3)	16 (17.58)	60 (16.0)
Secondary education	38 (39.2)	27 (27.0)	23 (26.7)	26 (28.57)	114 (30.5)
High school diploma	16 (16.5)	13 (13.0)	14 (16.3)	10 (11.0)	53 (14.2)
University level	27 (27.8)	46 (46.0)	35 (40.7)	39 (42.9)	147 (39.3)
APOE ε4 (n=23 missing data)	19 (20.7)	26 (28.3)	16 (19.5)	21 (23.9)	82 (23.2)
Mini Mental State Examination	28.13 [1.53]	28.38 [1.43]	28.39 [1.48]	27.99 [1.56]	28.22 [1.50]
Clinical dementia rating					
0	57 (58.2)	54 (54.0)	49 (56.3)	44 (47.8)	204 (54.1)
0-5	41 (41.8)	46 (46.0)	38 (43.7)	48 (52.2)	173 (45.9)
Free and Cued Selective Reminding Test					
Free recall*	27.70 [5.51]	28.26 [6.88]	29.16 [5.47]	27.36 [6.78]	28.10 [6.22]
Total recall*	45.45 [3.08]	45.73 [3.22]	46.37 [2.40]	45.30 [4.33]	45.70 [3.35]
Trail Making Test					
Part A	44.11 [16.71]	44.27 [16.28]	43.71 [13.58]	43.22 [12.91]	43.84 [14.98]
Part B (n=2 missing data)	107.41 [46.70]	106.91 [42.35]	110.94 [43.39]	110.40 [53.83]	108.82 [46.59]
Verbal fluency					
Category Naming Test	28.17 [6.70]	28.11 [7.98]	28.44 [6.92]	28.58 [7.37]	28.32 [7.24]
Controlled Oral Word Association Test	22.16 [5.78]	22.44 [7.54]	22.05 [5.89]	22.46 [5.81]	22.28 [6.30]
DSST (WAIS—R coding)	40.59 [10.29]	41.10 [10.48]	39.30 [8.91]	39.36 [8.16]	40.13 [9.55]
Memory functioning†	48.64 [16.17]	48.57 [16.98]	47.39 [15.94]	49.42 [18.85]	48.53 [16.97]
Consequences of everyday life‡	39.77 [24.72]	36.14 [23.57]	37.94 [24.32]	38.37 [21.95]	38.04 [23.61]
Geriatric Depression Scale (n=2 missing data)	2.80 [1.94]	3.28 [2.44]	3.41 [2.46]	3.18 [2.82]	3.16 [2.43]
Short Physical Performance Battery	11.00 [1.44]	11.14 [1.26]	10.76 [1.28]	11.01 [1.39]	10.98 [1.35]

Data are mean (SD), or n (%), unless otherwise specified. Percentages were calculated on the basis of the number of participants for whom data were available for each criterion. DSST=Digit Symbol Substitution Test. WAIS—R=Wechsler Adult Intelligence Scale—Revised. ADCSADLPI=Alzheimer's Disease Cooperative Study Activities of Daily Living Prevention Instrument. *48 is the total possible score. †Measured on a visual analogue scale.

Table 3.1: Baseline characteristics of the population

between pairs of images. These deformations are fully parametrized by the flow of a stationary velocity field (SVF). From these deformations it is possible to compute the corresponding volume changes locally in the brain (Lorenzi et al., 2013), but it is also possible to analyze the longitudinal intra-subject evolutions directly in the deformation space using the SVFs transported in a common reference space where they can be compared (Hadj-Hamou et al., 2016).

3.2.3.2 Assessing the effect of the treatments on the brain evolution

We want to evaluate the effect of both interventions on the morphological evolution. To do so, we test the H_0 hypothesis that the treatments are not correlated with the vector fields parametrizing the deformations. However these fields are high dimensional and repeating the test for every voxel will result in a large number of correlated tests. To address this problem, we use a non-parametric re-randomization test to control for family-wise error rate (FWER). This statistical approach does not require strong assumptions on the data distribution, especially when the patients are in a randomized trial, and is well fitted to the neuroimaging settings (Nichols and Holmes, 2002).

Several statistics can be used at the voxel level; most of them are based on the comparison of the residual variance matrix and the model variance matrix. Here, the voxel statistic is based on the likelihood ratio to compare the goodness of fit between the linear model taking into account the treatments and the constant model. No additional covariate are included in our model. The maximum of this statistic is then used to summarize the result over the whole brain. This maximum statistic assesses the effect of the treatments globally and is not associated to any specific brain region.

For each random permutation of the subjects, the maximum of the likelihood-ratio statistic map is computed on the whole brain area. These values give an empirical distribution of the maximum statistic under H_0 . This empirical distribution allows us to compute, for the real group assignment, the resulting p-value.

In order to get a reliable p-value estimation, $N = 2000$ permutations are computed. Indeed, the standard deviation of the p-value estimation at a significance level of $\alpha = 0.05$ is equal to:

$$P = \sqrt{\alpha(1 - \alpha)/N} \approx 0.005$$

Additional likelihood ratio analyses are performed to assess the effect of both the omega-3 and the multidomain intervention variables individually. For these analyses the same permutations of the 376 subjects are used, the only difference is that the linear model tested only takes into account one variable of interest.

To complement these results about the existence of a global difference in morphological evolution, we are also interested in the interpretation of the eventual difference and their relation to an evolution toward more advanced diseased stages. To achieve this we first visualize the voxel-wise z-value maps. They give insights on where the difference are correlated with the treatments. Then, the mean evolution

is computed for each subgroups and the local differences are highlighted. It allows us to make hypotheses about the properties of these differences.

3.3 Results

3.3.1 Cognitive, behavioral and functional assessment

Results regarding the 3-year evolution of the subjects cognitive, behavioral and functional profile in our sub-sample are reported in Table 3.2. The evolution of the scores (results at M36 minus results at baseline) of each of the available tests is evaluated using ANOVAs with Group (MDI plus omega-3, MDI plus placebo, omega-3 alone, placebo) as between-subject factor.

Converging with previous reports (Andrieu et al., 2017), no significant effect of the intervention was found on the evolution of the overall cognitive functioning (Mini-Mental State Examination (Folstein et al., 1975)), depression (Geriatric Depression Scale (Yesavage et al., 1982)), and physical fitness (Short Physical Performance Battery (Guralnik et al., 2000)). Similarly, no consequences were found on everyday life (Alzheimer’s Disease Cooperative Study Activities of Daily Living Prevention Instrument (Galasko et al., 2006)).

Concerning the evolution of the cognitive profile, no effect of group was found on the evolution of memory (as indexed by classical neurophysiological tests – Free and Cued Selective Reminding Test (Grober et al., 1988) – and self-report – Visual-analogue scale EVA (McNair and Kahn, 1983)) attention and executive functions (Digit Symbol Substitution Test (Wechsler, 1981) and TMT-A and TMT-B (Reitan, 1958)) or verbal letter fluency (Controlled Oral Word Association Test (Cardebat et al., 1990)). A significant effect on semantic verbal fluency (Category Naming Test) was found ($p=.027$), but post-hoc comparisons (Bonferroni corrected) revealed no-significant difference between the three intervention groups and the control group. These results argue that our sub-sample ($N=377$) is comparable to the original study population ($N=1680$).

3.3.2 Segmentation-based statistical analysis of the interventions’ effect

These results rely on the longitudinal processing done using the FreeSurfer software. One subject was excluded from this analysis due to a processing failure. All the results for the evolution of every volumes and cortical thickness are reported in supplementary material. Table 3.3 summarizes the global result. No effect is observable in any regions, including the total white-matter, gray-matter or brain volume to intra cranial volume, for both interventions. This absence of observable effect is not related to the large number of variables selected in this exploratory approach. Indeed the strongest difference are measured in the corpus callosum posterior and the left thalamus and their respective p-values without correction are respectively 0.036 and 0.043. By consequence no meaningful observation or hypothesis can be

Difference between (score at 3 years - score at baseline)	Multidomain plus polyunsaturated fatty acids (n=98)		Multidomain plus placebo (n=100)		Polyunsaturated fatty acids (n=87)		Placebo (n=92)		p-value ^a
	Mean [SD]	Mean [SD]	Mean [SD]	Mean [SD]	Mean [SD]	Mean [SD]			
Mini Mental State Examination	0.14 [1.69]	-0.14 [1.89]	-0.11 [1.89]	-0.28 [2.40]	0.525				
Free and Cued Selective Reminding Test									
Free recall ^{†*}	1.93 [4.95]	1.32 [5.68]	1.59 [5.48]	1.49 [5.61]	0.886				
Total recall ^{†*}	0.27 [2.75]	-0.40 [4.90]	-0.32 [2.99]	-1.12 [6.44]	0.218				
Trail Making Test									
Part A	-3.32 [12.30]	-1.93 [17.77]	-1.62 [15.82]	-1.46 [12.52]	0.818				
Part B	-1.56 [34.20]	-2.10 [35.37]	-3.78 [36.03]	-2.32 [58.45]	0.988				
Verbal fluency									
Category Naming Test	-1.06 [6.10]	1.05 [6.72]	0.90 [8.55]	-1.38 [6.67]	0.027				
Controlled Oral Word Association Test	0.36 [5.37]	1.13 [5.69]	1.65 [6.43]	-0.20 [6.38]	0.167				
DSST (WAIS—R coding)	0.26 [5.09]	1.13 [5.44]	1.12 [6.42]	-0.31 [6.21]	0.256				
Memory functioning^{†‡}	-0.78 [17.62]	-0.27 [17.98]	1.48 [17.25]	1.36 [22.82]	0.800				
Consequences of everyday life^{†‡}	-5.84 [29.48]	-4.76 [23.98]	-3.27 [23.31]	3.13 [26.48]	0.085				
Geriatric Depression Scale	0.01 [2.48]	-0.41 [2.53]	0.13 [2.83]	-0.03 [2.43]	0.494				
Short Physical Performance Battery	-0.16 [1.68]	-0.10 [1.37]	-0.12 [1.68]	-0.41 [1.62]	0.534				

^aANOVA
 DSST=Digit Symbol Substitution Test. WAIS—R=Wechsler Adult Intelligence Scale—Revised. ADCSADLP=Alzheimer’s Disease Cooperative Study Activities of Daily Living Prevention Instrument. *48 is the total possible score. †Measured on a visual analogue scale.

Table 3.2: Evolution of the cognitive assessments. Differences between (score at 3 years) - (score at baseline)

drawn from this experiment.

max Fisher statistic	univariate parametric p-value	corrected p-value (2000 perm.)
3.35	0.036	0.80

Table 3.3: Statistical analysis of morphological assessments relatively to treatment groups. Differences are evaluated using a simple linear model with no cross-effect. We report the global statistic over all the 142 variables. The maximum is attained on the volume of the posterior corpus callosum but the result is not significant after correction for multiple tests.

3.3.3 Deformation-based statistical analysis of the interventions' effect

A significant effect of the treatments was found on the morphological evolution over the 3 years follow-up (p-value = 0.018). This result for the primary efficacy analysis jointly consider both treatments (MDI and omega-3) and is corrected for the whole brain voxel-wise comparisons. The localization of this effect is visible in Figure 3.1(a). The effect is stronger in the peri-ventricular area in the left hemisphere near the temporoparietal junction (TPJ). Few differences are also visible in the right temporal lobe but are not statistically significant.

Further testing is done to evaluate if the effect is associated with the omega-3 or the MDI. The statistical maps are shown in Figure 3.1(b-c). The linear model associated with the omega-3 is found to be not significantly better than the constant model (p-value = 0.927) while the one with the MDI shows a significant correlation (p-value = 0.026). Furthermore, the model taking into account both variables do not perform significantly better than the one taking into account the MDI only (p-value=0.933). Moreover the pattern observed for the MDI is really similar to the one found for the joint modeling.

To summarize, the observed changes in the morphological evolution are only associated with the MDI. Consequently, in the following we focus on the description of the MDI effect only.

3.3.4 Description of the MDI effect on the brain morphology

In this section we visualize the effect of the MDI. First, the full 3-dimensional statistical z-map is represented in Figure 3.2. The positive cluster extent in the left posterior peri-ventricular area seems to indicate a relatively localized effect. However we should keep in mind that the displacements which are visible in this area could be induced by morphological changes in other parts of the brain. Moreover the high z-values in the corresponding area in the right hemisphere and in the right temporal lobe may indicate the existence of changes in these regions too.

These results mean that the mean evolutions for the group undergoing MDI and the other one are different. Figure 3.3 represents the mean velocity fields parametriz-

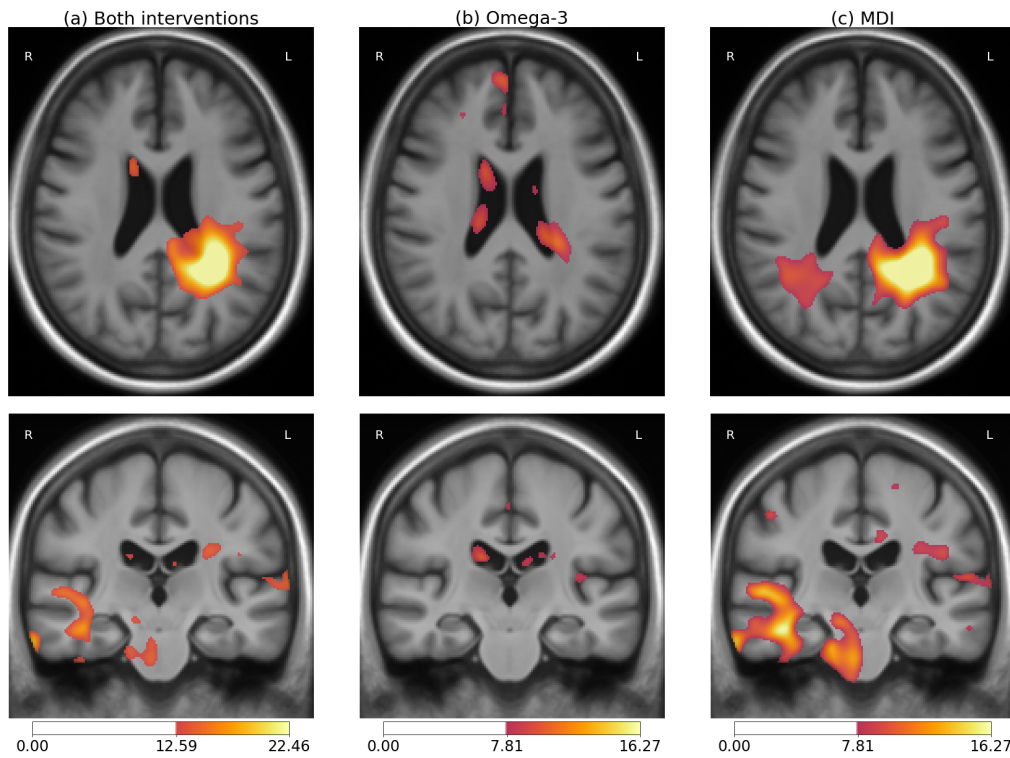


Figure 3.1: Voxel-wise z-values of the likelihood ratio test for the linear model taking into account: (a) both categorical variables associated with the omega-3 supplementation and the multidomain intervention, (b) omega-3 only, (c) multidomain intervention only. High values indicate a good fit of the model relative to the constant model. Intensity thresholds correspond to the independent parametric testing at level 0.05 (lower threshold) and 0.001 (higher threshold).

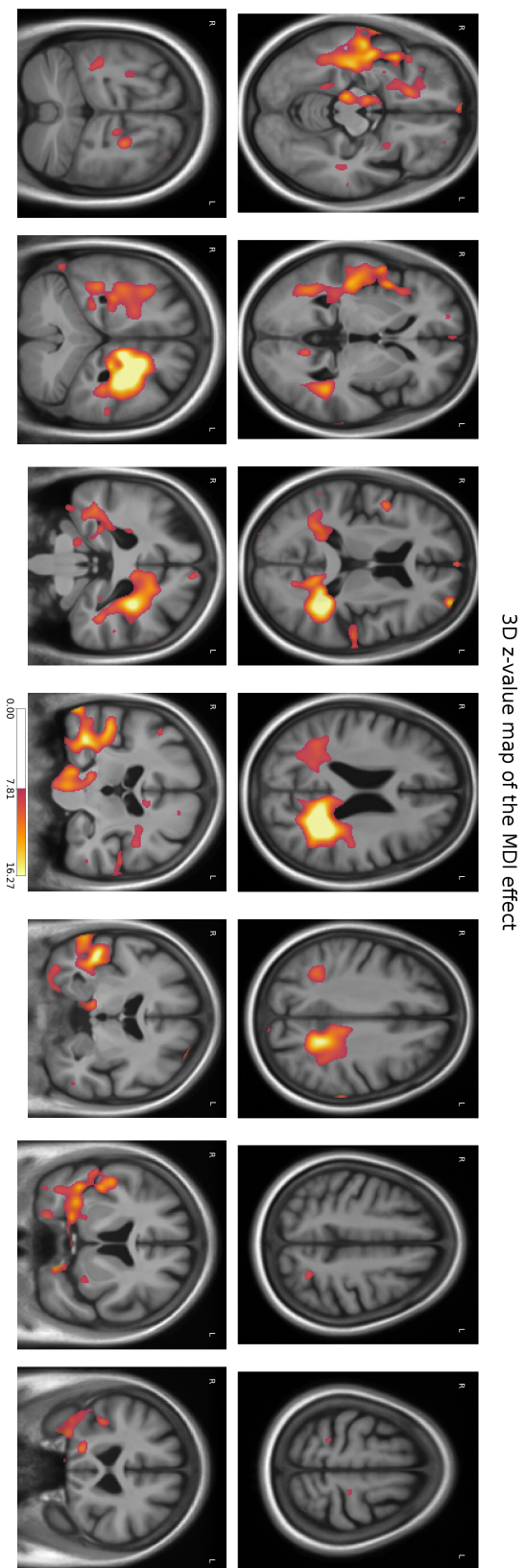


Figure 3.2: Voxel-wise z-values of the likelihood ratio test for the multidomain intervention effect. Intensity thresholds correspond to the independent parametric testing at level 0.05 (lower threshold) and 0.001 (higher threshold). Slice are sampled regularly every 12 mm in the longitudinal axis and every 15 mm along the axial axis. The central column corresponds to the results shown in Figure 3.1(c).

ing the longitudinal deformation for the subjects with or without the MDI. The omega-3 status is ignored to increase sample sizes as no morphological effect was found.

At first glance, both evolutions are similar and correspond to a normal aging process. The difference in the peri-ventricular area is visible and corresponds to a deformation of smaller amplitude. The main effect is visible in the TPJ with a slight asymmetry and a stronger difference on the left. We can also note a reduction of the deformation in both temporal lobes for subjects undergoing the MDI.

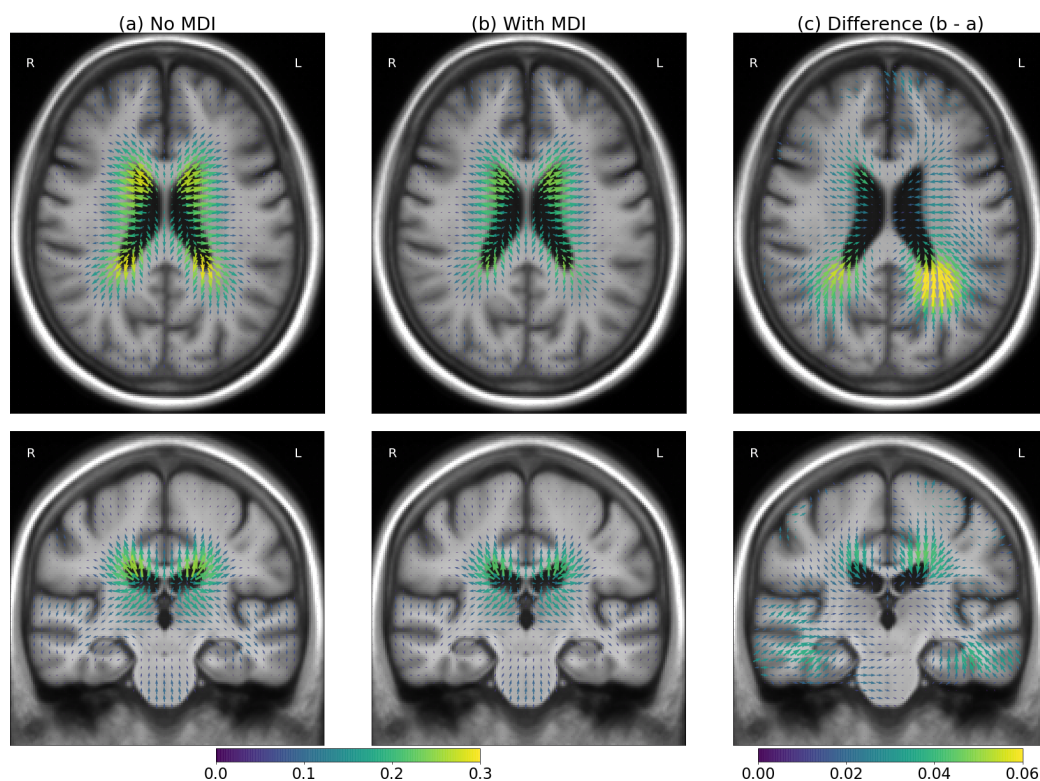


Figure 3.3: Mean velocity fields parametrizing the deformation for the subjects without MDI (left) and with MDI (center). The difference (right, scaled $\times 5$) highlights the changes associated with the intervention. The two means are similar, the main difference is a smaller amplitude of deformation in the peri-ventricular area.

It is also possible to visualize the atrophy associated with these deformations. Indeed any spatial transformation can induce local volume changes. In our model, the relative volume change in each voxel is given by the log-Jacobian map (see Figure 3.4). Here the difference clearly shows a reduction of the atrophy in the whole brain (less atrophy in the parenchyma and less expansion in the CSF). However the pattern previously observed on the deformation was more localized. It suggests that the difference of evolution is associated with a global difference in volume loss but also a more complex reconfiguration of the morphology.

Integrating the log-Jacobian using a predefined segmentation can be used to assess the volume change of specific brain areas. Here, using a segmentation of 44 regions provided the FreeSurfer software, the main regions affected by the MDI are the left lateral ventricle, the left inferior ventricle and the right hippocampus. No region shows a significant effect when correcting for multiple comparisons for this volumetric approach. This confirms that the difference is only partially due to volume changes and is associated to a more complex morphological reconfiguration.

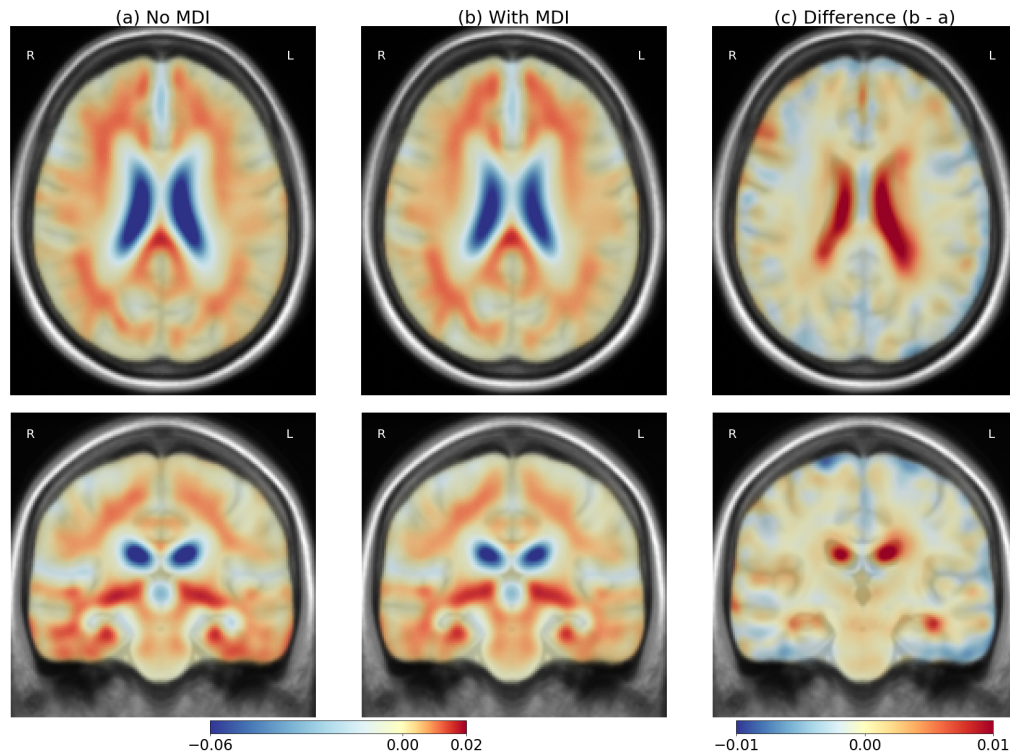


Figure 3.4: Mean atrophy for the subjects without MDI (left) and with MDI (center) and the difference (right) to highlights the changes associated with the intervention. The two means are similar, the difference correspond to a global slower atrophy pattern.

3.3.5 Relation to cognitive and behavioral scores

No direct cognitive effects of the treatment was observed. However it is possible to relate the observed morphological differences to clinical assessments in order to better interpret the results. In this section we compare the morphological changes that are MDI related to the cognitive evolution measured by clinical scores.

In the previous section, a regression model is used to evaluate the effect of the MDI. Our objective is to interpret the SVFs that parametrize this model (those shown on Figure 3.3). To this end, these SVFs are used to linearly decompose the

individual morphological evolutions. The coefficients given by this decomposition are, by construction, not independent between subjects, but they give information about these specific morphological changes in the population.

These coefficients are then analyzed in relation with the evolution of four cognitive and behavioral scores. The four selected variables, previously introduced, are the Mini Mental State Examination (MMSE), the Geriatric Depression Scale (GDS), the free recall test and the category naming verbal fluency test. The evolution is measured by the difference in score between the 3-year followup and the baseline. The analysis is done using the whole population (i.e. the 376 subjects).

The correlation between the coefficients of the morphological decomposition and each clinical variable is assessed using a Fisher test. The results are presented in Table 5.1. The intercept measures the average evolution in cognitive score in the whole population while the coefficient corresponds to the average change for a difference in morphological evolution equal to the difference between the control and the MDI group.

We find that the morphological changes that were associated with the MDI are also statistically associated with positive cognitive and behavioral effects: decrease in the GDS total score and increase of the memory recall test. For comparison, no significant longitudinal evolution in GDS is observed for the control group while a general improvement in recall performance is observed for the overall population (the morphological changes of interest being associated a better amelioration). The effect on MMSE and fluency is not significant but, in both cases, it corresponds to a slowing of the cognitive decline.

	F-test p-value	intercept [CI]	coefficient [CI]
MMSE	0.078	-0.14 [-0.35, +0.066]	0.097 [-0.011, +0.2]
GDS	0.011	0.013 [-0.25, +0.28]	-0.18 [-0.32, -0.042]
Recall	<0.001	1.3 [+0.71, +1.8]	0.55 [+0.26, +0.84]
Fluency	0.44	-0.2 [-0.94, +0.55]	0.15 [-0.23, +0.54]

Table 3.4: Test for correlation between the MDI morphological effect and the cognitive or behavioral clinical assessments. The intercepts correspond to the average 3-year evolution, the coefficients correspond to the score change for a change in morphological evolution equal to the difference between the control group and the MDI group.

3.4 Discussion

The statistical analysis of the brain deformations estimated from structural MRIs shows a possible effect of the MDI on the morphological changes. Differences are observed in peri-ventricular area and in the temporal lobes in an asymmetrical pattern. The effect is stronger in the left hemisphere near the pre-cuneus and using a non-parametric analyses, we show that this effect is statistically significant. No

visible effect was found for the omega-3 supplementation relatively to the placebo.

3.4.1 Interpretation of the morphological changes

First we need to note that locally and globally the MDI is also associated with a general pattern of reduction of deformation and atrophy in the whole brain. It would mean that the evolution is similar but slower thanks to the interventions. This evolution is mainly visible through the ventricular expansion and would explain that the effect is visible at the ventricle boundary. However the difference are not uniformly spread in the brain and in particular the left hemisphere is more affected.

The localization of the effect in the deep white matter may also indicate a direct effect on the myelinised axons in this area. Physiological changes in these area could have complex implication on the general cognitive condition. It is particularly interesting as the damage on the white matter tracks is not necessarily correlated with cortical atrophy as shown by [Agosta et al. \(2011\)](#) for amnesic patients with mild cognitive impairment. These observations would give early signs of the decline before the cortical atrophy.

The asymmetry could be related to the selection of exercises proposed in the MDI. A focus on tasks strongly relying on language would be associated with a stronger effect on the left hemisphere. One interpretation is that the intervention would increase or contribute to maintain the cognitive reserve in these areas.

Similarities can be found with early Alzheimer atrophy patterns where cortical atrophy is first observed in the temporal lobes, in the inferior parietal and posterior cingulate ([McDonald et al., 2009](#)). The difference of atrophy measured between controls and MDI subjects is smoother and less localized but is not incompatible with this pattern. The atrophy in one region can create a deformation that propagate to neighboring regions where the displacements may be more visible.

Finally, the correlation of the morphological effect with the cognitive and behavioral evolution suggests a possible long-term improvement of the subjects cognitive and behavioural profile. Moreover, the amplitude of the effect could be comparable to the normal speed of the cognitive decline.

3.4.2 Limitations

The selected morphometric approach gives a rough localization of visible structural displacements in the brain. The localization incertitude is reinforced by the fact that registration algorithms are not able to accurately estimate deformations in intensity homogeneous area and by consequence, the method can be more sensitive in high contrast area and miss some information elsewhere. Beside we saw that these deformations may be not directly related to local volume changes. The use of a full diffeomorphic deformation is powerful to detect subtle changes but the increase in complexity and generality makes the interpretation of the results more difficult.

Moreover, it is difficult to interpret the observed differences because we can only consider the joint effect of the MDI which regroups a large spectrum of activ-

ity (social interactions, physical cognitive exercises, nutritional advice, etc.). The interpretation is also made difficult because no direct effect of the treatment on the cognitive evolutions has been shown. And, in this work, the relation between these morphological changes and the cognitive condition studied in section 3.3.5 is evaluated on the same dataset and could be sensitive to hidden exogenous variables.

Finally, the implementation of the statistical model has its own limitations. It makes the GLM model applicable to multivariate fields but does not include alternative approaches such as mixed-effect modeling or computational optimizations. It would also be interesting to try to replicate these results with a different morphological modeling and statistical approach as the model choice and the software may influence the findings (Bowring et al., 2018).

3.4.3 Perspectives

Further analyses should be done to better understand the nature of the changes, their localization and the long-term consequence on the cognition.

An analysis of the deformation over a longer time interval could increase the accuracy of the estimation. It would also make it possible to link the local deformations to the cognitive evaluation despite the intrinsic variability of these assessments. Then, in conjunction with structural images, metabolic or functional imaging would give complementary information on the localization and the nature of the changes. Diffusion imaging could help validating the hypothesis concerning the effect of the interventions on the white matter tracks.

In order to design better protocols, it would be interesting to know the relative contribution of each intervention within the MDI. We should however note that the observed joint effect is relatively weak and it would not be an easy task.

Finally, the long-term effect on cognition should be assessed. It could be interesting to focus on cognitive tests evaluating functions which rely on the highlighted areas. The morphological effect of the multi-domain intervention goes against changes that are usually associated with the aging of the brain and the decline of the cognitive functions. However even if these changes are correlated, we have no proof that the MDI has a long-term protective effect that could be used for the prevention of the cognitive decline and of an eventual dementia.

3.5 Conclusion

In the context of the MAPT study, a significant effect of the multidomain intervention on the brain morphology was found. The voxel-wise deformation-based approach highlights differences in the white matter near the left temporoparietal junction. These morphological changes correspond to a slower morphological evolution and are correlated with better performance in cognitive assessments. These results are in agreement with previous studies suggesting a positive effect of a similar intervention on cognitive health in older adults (Kivipelto et al., 2018). To go

further, the long-term effects of the intervention on the cognitive and behavioral condition should be evaluated.

From the methodological standpoint, the results of the present study suggest that the statistical analysis of longitudinal deformations may be a powerful tool to assess the morphological changes and the effect of a treatment even without priors on the affected areas. In this work, we were unable to observe these differences using the volumetric approach, which suggest that the observed effect is associated with a complex morphological reconfiguration that cannot be reduced to simple local volume changes. This work thus suggests that the morphological changes may help to link physiological changes and changes in cognitive functions, even before changes in cognitive functions are clinically relevant.

Statistical framework for hypothesis testing on deformations

Contents

4.1	Context and objectives	62
4.1.1	Deformation-based morphometry	62
4.1.2	Clinical motivation: the MAPT case study	63
4.1.3	Chapter overview	63
4.2	Method	64
4.2.1	Multivariate General Linear Model	64
4.2.2	Likelihood ratio test	65
4.2.3	Design for non-parametric statistical testing on images	65
4.3	Experimental results in the MAPT study	67
4.3.1	Results of the multivariate analysis	67
4.3.2	Comparison to the atrophy-based assessments	68
4.4	Alternative methodological choices	70
4.4.1	Voxel statistic	70
4.4.2	Summary statistic	72
4.5	Discussion	76
4.5.1	Statistical analysis parameters	76
4.5.2	Interpretation of morphological differences	77
4.6	Conclusion	78
4.7	Appendix: the Likelihood ratio test and Aikake's weights	78

The statistical analysis of the deformations is a central element of Chapter 3 where it is used to characterize the effect of the treatments in the MAPT study. However the focus of Chapter 3 was on the description of the results and their clinical implications. Nonetheless, the development of general approaches that can be easily adapted to new problems is an important challenge too.

Voxel-wise statistical analysis is a well-established domain in medical imaging in particular for scalar images. However, the usual approaches are limited in their characterization of the morphological changes while the analysis of the multivariate fields parametrizing the deformations raises new issues and challenges.

In this context, this chapter presents the general mass-multivariate statistical testing framework for multivariate images behind the analysis in Chapter 3. This framework builds on deformation-based morphometry (DBM) to be able to characterize patterns and highlight differences related to clinical variables on multivariate images. This chapter presents the theoretical foundations of the framework, details the methodology and explores the importance of the main parameters of the approach. The question of the treatment effect in the MAPT study (introduced in Chapter 3) is used as main illustration.

4.1 Context and objectives

4.1.1 Deformation-based morphometry

Neuroimaging opens a window on ongoing processes in the brain that are otherwise not quantifiable. It can be used to assess the effect of a treatment, a specific clinical condition or the reaction to any kind of stimulation. Each imaging modality brings its own specificity but the spatial processing and the general problematic are generally similar. However high-dimensionality of the data makes quantitative assessments difficult.

In this context, the first attempts to produce voxel-specific statistical inferences were developed for metabolic images (PET). The multitude of required statistical tests raises the need to control the false positive rate. Parametric solutions were proposed based on the random fields theory. They lead to the development of multiple methodological and computational tools for voxel-based morphometry (VBM) and a rich literature addressing this problem and its applications (see [Ashburner \(2012\)](#) for an overview of these developments).

However, these parametric approaches rely on hypotheses about the spatial consistency of the voxels that can lead to error in false positive rate estimation ([Eklund et al., 2016](#); [Cox et al., 2017](#)). Non-parametric statistical framework have been proposed in particular with [Nichols and Holmes \(2002\)](#). They usually require less hypotheses to be valid at the expense of additional computing resources ([Winkler et al., 2014](#)).

Finally, the analyses are often restricted to scalar values leading to strong limitations for the analysis of multivariate images such as diffusion tensor images (DTI) or deformation fields. In DTI, scalar quantities such as fractional anisotropy or trace are generally used (see [Smith et al. \(2006\)](#) for example) even if such summaries do not capture all the information available. Concerning deformations, most approaches rely on scalar measures of the atrophy and multivariate statistical analyses are often restricted to low dimensional representation of the deformations ([Ashburner et al., 1998](#); [Teipel et al., 2007](#)).

Some works have (partially) addressed this problem. [Schwartzman et al. \(2005\)](#) proposed a statistical model for the direction of the diffusion (first eigenvector of the diffusion matrix) based on the bipolar Watson distribution. A t2 Hotelling statistic can also be used when the objective is to compare two groups of subjects.

This approach is used in [Hadj-Hamou et al. \(2016\)](#) on stationary velocity fields parametrizing longitudinal deformations, but this approach is limited to a two-sample discriminant analysis. Using a different approach, [Maris and Oostenveld \(2007\)](#) handles multi-sensors electroencephalographic (EEG) data using a clustering strategy that combine the results obtained using univariate statistics. To our knowledge, there is no general framework available for the analysis of multivariate images (such as vector fields) even if the conceptual background is well established.

4.1.2 Clinical motivation: the MAPT case study

The Multidomain Alzheimer Preventive Trial (MAPT) presented in Chapter 3 is a randomized placebo-controlled trial designed to assess the effect of omega-3 supplementation and a multidomain intervention (MDI) on cognitive decline of patients subjective memory complaint (SMD) ([Vellas et al., 2015](#)).

Our objective is to assess the effect of the treatments on the morphological evolutions that are modeled using a longitudinal deformations parametrized by a SVFs estimated through non-linear registration. The study is conducted on the 376 participants with a structural MRI acquisition at baseline and at the end of the 3-year trial. For each subject, a stationary velocity field (SVF) parametrizing the longitudinal deformation in a common reference anatomical space is computed using the framework proposed in [Hadj-Hamou et al. \(2016\)](#). The resulting experimental setting is then quite simple and standard: each subject is characterized by a 3D vector field (the SVF) and two treatment variables.

4.1.3 Chapter overview

The objective of this chapter is to describe a complete non-parametric statistical framework for multivariate images. In this framework, the General Linear Model (GLM) is used to design statistics that are suited for the multivariate analysis at the voxel level. The multiple comparisons problem raised by the high-dimensionality of the images is handled using a permutation-based approach that requires few hypotheses and is well-suited for the analysis of medical images.

The first part introduces the method. We summarize the problem of hypothesis testing for the GLM with the elements that will be needed in this work. Then, the key steps of the method are presented. The presentation is practical and the approach is illustrated by the MAPT clinical trial. In the second part, the results of the multivariate analysis of the MAPT are compared to the results obtained using a scalar atrophy measurement. Finally, we evaluate possible alternatives to the design of the statistical analysis on two aspects: the choice of the voxel statistic, and the use of a super-threshold summary statistic.

4.2 Method

This section is written as a practical guide to conduct a multivariate statistical analysis using our problem as a case study. What follows describes how to assess a correlation between a treatment or a clinical variable and, the longitudinal morphological evolution measured by image registration. We first introduce the elements of hypothesis testing that will be used at the voxel level.

4.2.1 Multivariate General Linear Model

The main objective of this work is to be able to analyze multivariate images in a general context. To do so, the idea is to use the generic General Linear Model (GLM) framework (that includes t-test, t2-test, ANOVA, etc.) that is easily applicable to multivariate data at each voxel.

We consider the problem with n subjects, d explanatory variables and p responses. We denote Z the (n, d) design matrix, Y the (n, p) response matrix, B the (d, p) coefficients, and e the (n, p) residual vector. The linear model writes:

$$Y = ZB + e. \quad (4.1)$$

Under the hypothesis of Gaussian residual, $e \sim \mathcal{N}(0, \Sigma)$, the maximum likelihood estimators (MLE) write:

$$\hat{B} = (Z'Z)^{-1}Z'Y, \quad (4.2)$$

$$\hat{\Sigma} = \frac{1}{n}(Y - Z\hat{B})'(Y - Z\hat{B}) \quad (4.3)$$

$$= \frac{1}{n}Y'(I_n - P_Z)Y, \text{ where } P_Z = Z(Z'Z)^{-1}Z'. \quad (4.4)$$

In the context of hypothesis testing, the objective is generally to assess if the model is a good explanation of the responses. The statistics that are used to reject the null hypothesis are often based on a decomposition of the variance between the part explained by the model and the variance of the residuals. In particular, using the predicted values $\hat{Y} = Z\hat{B}$ and the residuals $\hat{e} = Y - \hat{Y}$, we show that we have:

$$Y'Y = \hat{Y}'\hat{Y} + \hat{e}'\hat{e}.$$

This equation is directly seen as the decomposition of the variances between the model variance matrix H and the error variance matrix E :

$$Y'Y = H + E \text{ with } H = \hat{Y}'\hat{Y} \text{ and } E = \hat{e}'\hat{e},$$

These two parts are (p, p) covariance matrices because of the multivariate response. Then the usual statistics are based on the eigenvalues of these matrices, or directly on the matrix $A = HE^{-1}$. For instance, the Wilk's lambda: $\Lambda^* = |E|/|E + H|$, Pillai's trace: $\text{tr}(H(E + H)^{-1})$ are commonly used in the context of MANOVA.

4.2.2 Likelihood ratio test

In this work, our statistic derives from the likelihood ratio test. This approach relies on model comparison between two sets of predictors, for example between the model evaluated and the constant model.

Let $L(B, \Sigma)$ be the likelihood of the linear model with Gaussian residual, we have:

$$\log L(B, \Sigma) = -n \frac{\log |\Sigma|}{2} - n \frac{p}{2} \log(2\pi) - \sum_i e_i^T \Sigma^{-1} e_i \quad (4.5)$$

$$= -n \frac{\log |\Sigma|}{2} - n \frac{p}{2} \log(2\pi) - \text{trace}(\Sigma^{-1} (Y - ZB)' (Y - ZB)) \quad (4.6)$$

In general, the likelihood ratio criterion used to compare two sets of predictors is defined by:

$$\Lambda = \frac{\max_{H_0} L(B, \Sigma)}{\max_{H_1} L(B, \Sigma)} \quad (4.7)$$

Using the maximum likelihood estimators from 4.4 in the likelihood 4.6, the formulation simplifies and the criterion is only based on the error matrices of both linear models:

$$\log \Lambda = \log \frac{L(\hat{B}_0, \hat{\Sigma}_0)}{L(\hat{B}_1, \hat{\Sigma}_1)} = \frac{n}{2} \log \left(\frac{|\hat{\Sigma}_0|}{|\hat{\Sigma}_1|} \right).$$

When the objective is to assess the complete model in comparison to the null model, the statistic can equivalently be written using the model variance matrix H and the error variance matrix E previously introduced:

$$\Lambda = \left(\frac{|E|}{|E + H|} \right)^{n/2}. \quad (4.8)$$

This statistic should be large when the model is unexplanative (i.e we reject H_0 if Λ is small). More generally, this approach can be used to compare two sets of predictors or to include covariates in the analysis. We refer to [Fujikoshi \(2016\)](#) for more details. We can remark that this test is related to AIC model selection (see [Appendix 4.7](#)).

4.2.3 Design for non-parametric statistical testing on images

A permutation-based scheme is build around this voxel statistic. We based the description of our method on the step highlighted by [Nichols and Holmes \(2002\)](#). They detail 6 keys steps to perform a permutation analysis for neuroimaging:

1. the null hypothesis,
2. the exchangeability constraints,

3. the choice of two statistics: voxel-level statistic and summary statistic,
4. the relabeling algorithm,
5. the permutation distribution,
6. the significance level.

The application of this organization for the assessment of the treatment effects in the MAPT leads to the following points.

Null hypothesis: Our H_0 hypothesis is that the treatment variables have no effect on the SVFs parametrizing the longitudinal intra-subject evolution.

Exchangeability: The MAPT study was randomized. Under H_0 , using another random assignment of the subjects regarding the treatments would lead to the same data distribution. Every subject can be assigned to every treatment group as long as we match the original randomization. In particular, the number of subjects undergoing each treatment (and each combination of treatments) is conserved.

Voxel-level statistic: The likelihood ratio test for the GLM is computed for each voxel independently (see formula 4.8). No covariates are included in our analysis; we compare the linear model regressing the local velocity vector using the two treatment variables (error matrix E_0) to the constant model (error matrix E_1).

In order to improve the direct interpretability, the result is scaled depending on q the number of predictors of the model (here $q = 2 + 1$) and r the number of predictor of the reduced model ($r = 1$). Indeed for large n we can show using Wilk's theorem that:

$$-(n - r - 1 - \frac{1}{2}(p - r + q + 1)) \ln\left(\frac{|E_0|}{|E_1|}\right) \sim \chi_{p(r-q)}^2, \quad (4.9)$$

where the dimension p of the output is here equal to the dimension of the voxels of the SVF, i.e. $p = 3$.

Summary statistic: The summary statistic aims to combine all the voxel statistic to give a global assessment image wise. We use the maximum statistic (single threshold for the rejection of the *omnibus hypothesis*). We also evaluated the other approach described by [Nichols and Holmes \(2002\)](#) (ie suprathreshold cluster). The central idea behind the maximum statistic is that, quoting: “for the single threshold test to be equally sensitive at all voxels, the (null) sampling distribution of the chosen statistic should be similar across voxels”.

Relabeling: For combinatorial and computational reasons, it is impossible to do the exact permutation test and compute the statistic for every admissible permutations. Consequently we randomly draw N valid permutations, the standard deviation of the p-value estimation at a significance level of $\alpha = 0.05$ can be estimated using:

$$P = \sqrt{\alpha(1 - \alpha)/N}$$

We choose to draw $N = 2000$ giving a standard deviation $P \approx 0.005$. The permutation sampling is performed using the numpy randperm function.

Permutation distribution: For each permutation, the maximum statistic is computed independently. From the computational point of view, this can be expensive but is easily parallelized on a computer cluster.

Significance threshold: We use the 95th percentile to determine significance. The hypothesis can be rejected if the maximum statistic from the true assignment is above this threshold. More generally empirical p-value can be computed using the sorted distribution of the maxima obtained for every permutations.

4.3 Experimental results in the MAPT study

4.3.1 Results of the multivariate analysis

A significant effect of the treatments on the morphological evolution is found (p-value = 0.018). This result for the primary efficacy analysis jointly considers both treatments (MDI and omega-3) and is corrected for the whole brain voxel-wise comparisons. The localization of this effect is visible on Figure 4.1 that reproduces the z-value maps for the likelihood ratio tests for each intervention. The effect is stronger in the peri-ventricular area in the left hemisphere near the temporoparietal junction.

Further testing is done to evaluate if the effect is associated with the omega-3 or the MDI. Additional likelihood ratio analyses are performed to separately assess the effect of each predictor variable. The z-value maps are reported on Figure 4.1. The linear model associated with the omega-3 is found to be not significantly better than the constant model (p-value = 0.927) while the one with the MDI shows a significant correlation (p-value = 0.026). Furthermore, the model taking into account both variables do not perform significantly better than the one taking into account the MDI only (p-value=0.933) and the pattern observed for the MDI is similar to the one found for the joint modeling.

These results of the analysis of the MAPT are fully described and discussed in [Sivera et al. \(2019a\)](#).

4.3.2 Comparison to the atrophy-based assessments

The results of the multivariate analysis of the deformations can be compared to similar analyses on a scalar (ie univariate images) measure of the atrophy derived from the deformations. Here, the atrophy is quantified by the log-Jacobian of the SVFs (Lorenzi et al., 2013). Our question regarding the potential effect of the MAPT interventions on the brain morphometry can then be addressed using the atrophy maps. This volumetric analysis, that is possible using existing software such as SPM (Friston, 2007), will be compared to the statistical analysis on the SVFs directly. Of course, our framework can also be applied to univariate images as it is only a particular case of multivariate analysis.

Figure 4.2 shows the results obtained using our framework on the log-Jacobian maps. No significant effect is detected. A cluster of high z-value is visible in the same area than for the SVFs but the empirical p-values are lower, in particular when considering the MDI alone. These results would not be conclusive by themselves.

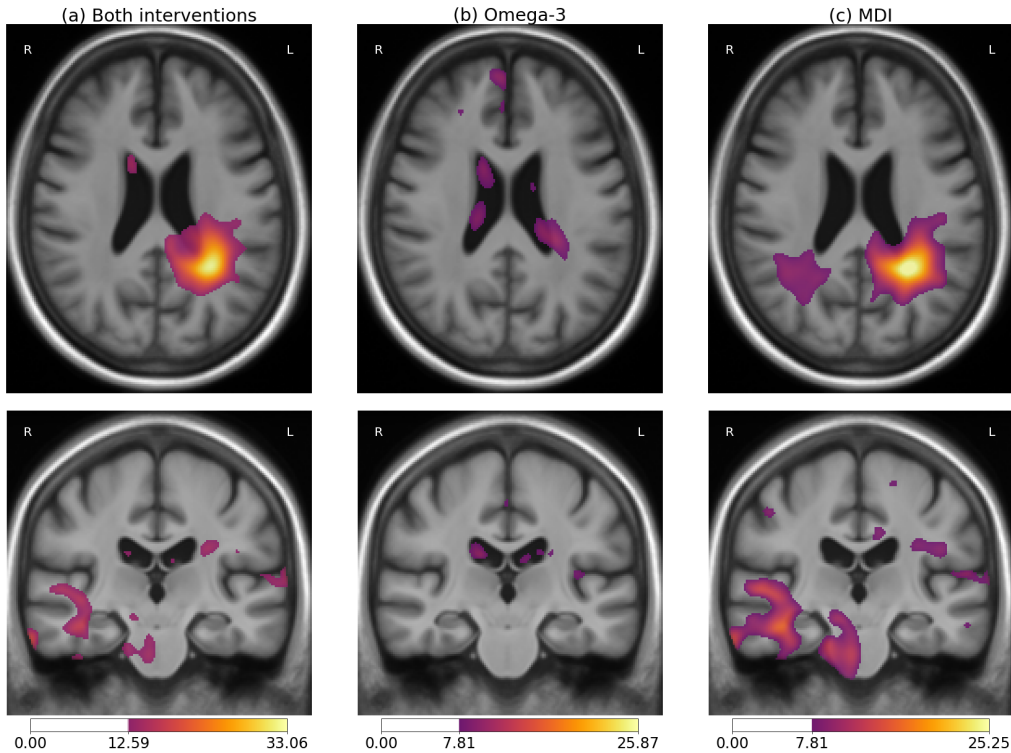


Figure 4.1: Voxel-wise z-values of the likelihood ratio test for the multivariate linear model on the SVFs taking into account: (a) both treatment variable, (b) the omega-3 only, (c) the multidomain intervention only. High values indicate a good fit of the model relatively to the constant model. Intensity thresholds correspond to the level $p=0.05$ for the non-corrected voxel parametric test (lower threshold) and for the empirically corrected using permutations test (higher threshold).

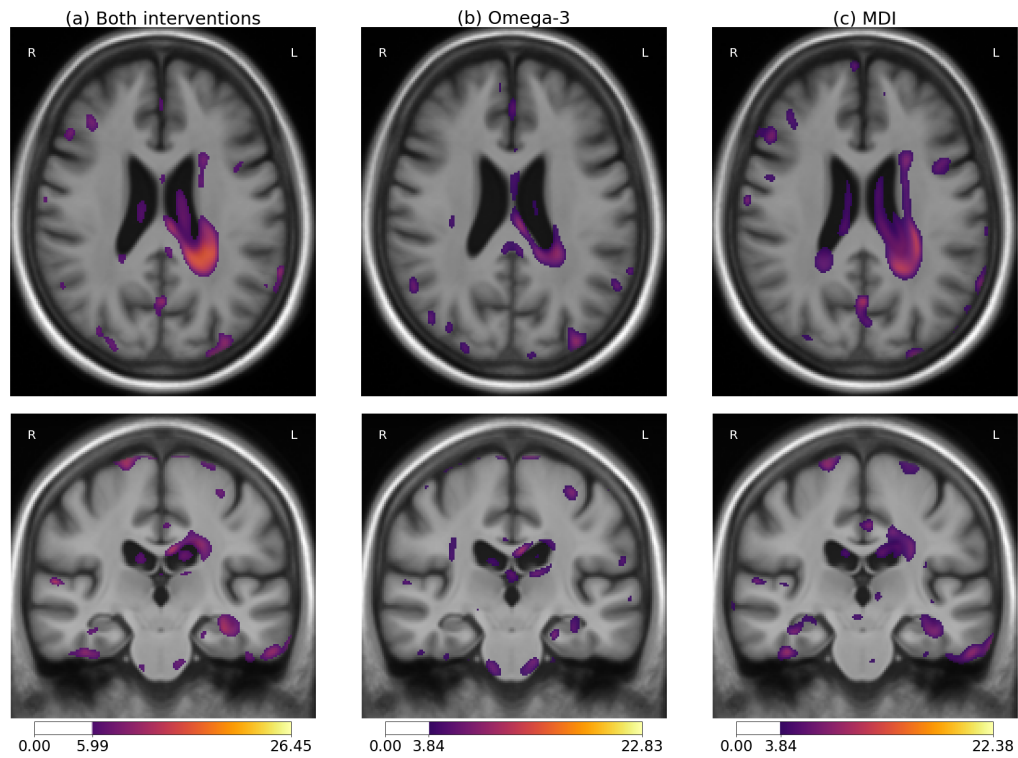


Figure 4.2: Voxel-wise z-values of the likelihood ratio test for the univariate linear model on the atrophy maps taking into account: (a) both treatment variable, (b) the omega-3 only, (c) the multidomain intervention only. High values indicate a good fit of the model relatively to the constant model. Intensity thresholds correspond to the level $p=0.05$ for the non-corrected voxel parametric test (lower threshold) and for the empirically corrected using permutations test (higher threshold).

A similar factorial design is used in SPM to regroup the subjects depending on their MDI and their omega-3 supplementation status. We assess the effect of both interventions on the atrophy maps. The clusters are computed on the whole brain with a height threshold of $p=0.001$. The results are presented in Figure 4.3. Not many clusters are visible and they are all small. The results are not significant.

4.4 Alternative methodological choices

In this section, we explore possible alternatives to the design of the statistical analysis on two aspects. First we consider other possible voxel statistics. Then, we evaluate the use of a super-threshold summary statistic.

4.4.1 Voxel statistic

The first alternatives to the likelihood ratio statistic are other statistics usually used to test the same hypothesis in the literature. Here we consider the Lawley-Hotelling's trace, the Pillai's trace and Roy's maximum root. All these tests are related to the eigenvalues λ_k of the matrix HE^{-1} . Here, for sake of simplicity, we do not rescale them. The definitions used are:

- Likelihood ratio: $S = \log\left(\frac{|H|}{|E|}\right) = \log(|HE^{-1}|) = \sum_k \log(\lambda_k)$,
- Lawley-Hotelling's trace: $S = \text{trace}(HE^{-1}) = \sum_k \lambda_k$,
- Pillai's trace: $S = \frac{V}{p-V}$ where $V = \text{trace}(H(H+E)^{-1}) = \sum_k \frac{\lambda_k}{1+\lambda_k}$ and $p = 3$,
- Roy's maximum root: $S = \max_k \lambda_k$.

The results comparing these statistics are presented in Figure 4.4. They are really similar; only the Roy's root shows a visible difference with more contrast between the significant area and the rest of the brain. The amplitude of the statistics also highlights the weakness of the model explanation. The *explained variance* is relatively small (we can say that it explains at best 10% of the variability in the best area) and the matrix HE^{-1} is close to the identity. This observation can also explain why the first three statistics are similar. Indeed if all the λ_k are close to 1 the three formulas give similar results. The Roy's root may be highlighting the anisotropy of the data and could lead to complementary results.

To explore more widely the problem of the choice of the voxel-statistic, we can take a more open approach. Indeed any metric or pseudo-metric on symmetric positive matrices can be used. We propose here three additional alternatives of the form $f(H) - f(E)$ inspired by the the likelihood ratio statistic that can be written $\log(|H|) - \log(|E|)$. We use:

- the trace: $\text{trace}(H - E)$
- the sum of coefficients: $\sum_{i,j} H_{ij} - E_{ij}$

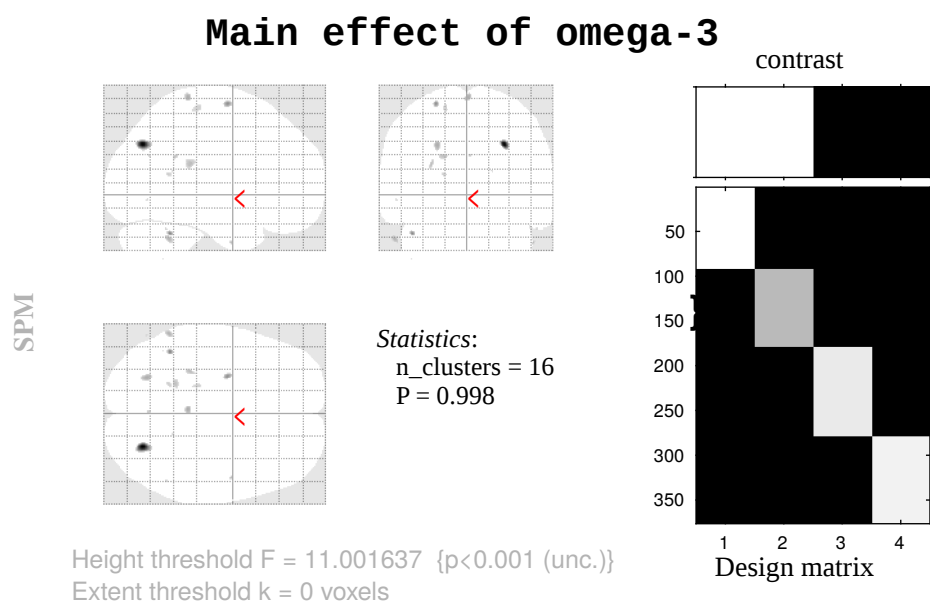
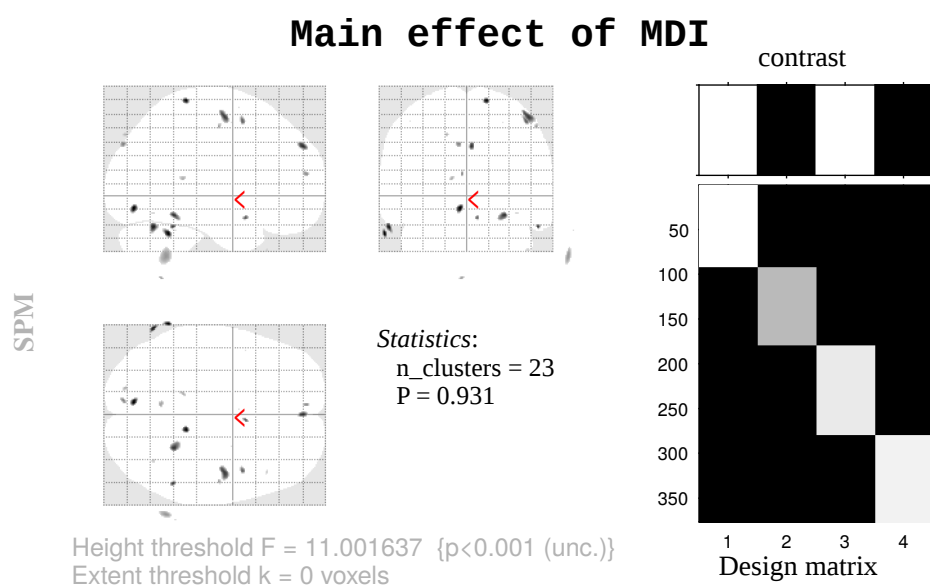


Figure 4.3: SPM results for the voxel-wise statistical analysis of the effect of both MAPT interventions on the log-Jacobian maps. Top: multidomain intervention, bot: omega-3 supplementation. No significant result is highlighted.

- the sum of squares: $\sum_{i,j} H_{ij}^2 - E_{ij}^2$

The results are shown in Figure 4.4. The highlighted voxels are localized in relatively similar areas. However their extents and relative importance do vary. For example for the sum of squares the peri-ventricular pattern is much more symmetric with counterpart in the right hemisphere clearly visible, and the differences are not visible in the right temporal lobe while the left hippocampus seems to appear.

4.4.2 Summary statistic

The most common alternative to the maximum statistic is the size of supra-threshold clusters. The statistical map is thresholded at a given threshold. Then the size of the connected components is used to determine the significance of the result: a large cluster being associated with a positive result.

The main question when using a cluster based approach is how to create these clusters. Here we use a simple thresholding reducing the problem to the choice on one single parameter: the threshold value. The choice of the threshold may influence the result. A too low value will create huge blobs that will be uninterpretable and probably not meaningful. A too high value will be too strict to detect potential results.

For the sake of interpretability, the threshold is often chosen to correspond to a specific p-value for the individual voxel statistic. The threshold t is then parametrized by the level α of the single voxel parametric test using the inverse cumulative distribution function: $t = \text{ICDF}(1 - \alpha)$. The level $\alpha = 0.001$ is often recommended; higher thresholds may be too strict, more liberal lower thresholds may create large (ie uninterpretable) and unreliable clusters (Woo et al., 2014). In our case, the threshold is considered relatively to the theoretical H_0 distribution of the voxels: a χ^2 distribution with 6 degree of freedom.

The H_0 distribution of supra-threshold cluster sizes can be computed in the context of the statistical field theory. However, we use in the work the permutation-based non-parametric scheme introduced earlier.

This section compares the results for the maximum statistic and for the supra-threshold statistic for several thresholds.

4.4.2.1 Supra-threshold clusters

In Figure 4.5, we explore the empirical distribution of the max cluster size and the number of clusters for several values of α . As expected the cluster size increases when α increases (i.e. the threshold decreases). We also see that the number of cluster increases. Eventually the clusters will merge (when the threshold comes close to 0) but, for the values considered here, the clusters are generally well separated and the apparition of new clusters is clearly the dominant trend.

The main interesting point is that the result is significant (red line above the dark gray zone showing the confidence interval) only for $\alpha \leq 0.001$. Above this limit, the variability in cluster sizes explodes. This evolution is not directly reflected in

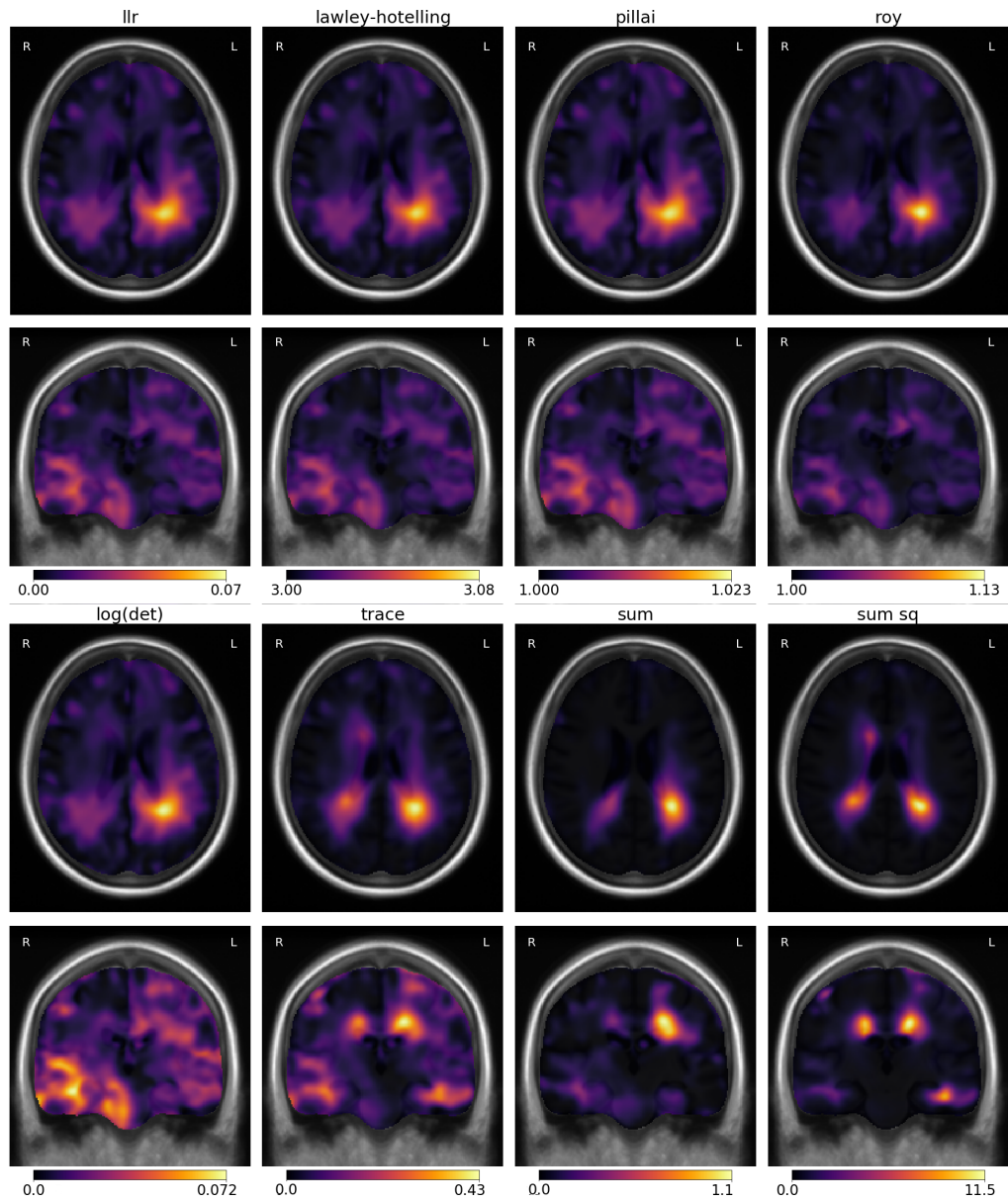


Figure 4.4: Alternatives to the likelihood ratio statistic illustrated on the MDI/no MDI test. Top: statistical test, from left to right: log likelihood ratio, Lawley-Hotelling's trace, Pillai's trace, Roy's maximum root. Bottom: pseudo-metrics, from left to right: $\log(\det)$, trace, L_1 difference, L_2 difference.

the number of clusters. It increases, for both the actual data and the permuted versions, but the number is always quite similar (even if we observe in average a bit more clusters for the permutations).

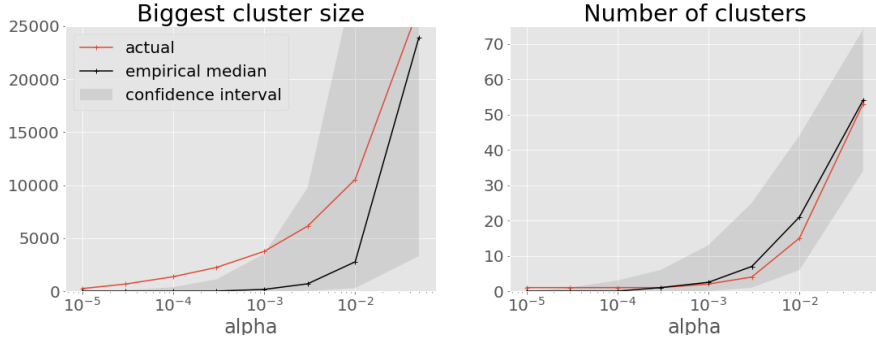


Figure 4.5: Size of the biggest cluster and number of clusters in function of the threshold. The red line shows the actual values. The black line (median) and the gray area (inter-quantile interval $[q_{0.05}; q_{0.95}]$) show the empirical distribution over 2000 permutations. The result is significant when the red line is above the dark gray area (i.e. $\alpha \leq 0.001$ for the cluster size).

In Figure 4.6, we visualize the actual clusters for several level of α . Without surprise, the main area that was observed previously observed (see Figure 4.1) is clearly visible. Moreover, at every level, no other cluster has a comparable size. We see that the extent of this cluster is strongly dependent on the choice of the parameter α and this can make the anatomical interpretation even more difficult.

4.4.2.2 Empirical distribution of max cluster size

In this paragraph, we discuss the difference between the empirical distributions estimated using the maximum or the cluster size statistic. The distributions are reported in Figure 4.7. The maximum statistic is unimodal with a mode around 25 and slightly skewed to the right but the tail is relatively short. The cluster size statistic is strongly concentrated near 0. Indeed, by definition, every sample on the left on the threshold line for the maximum leads to a maximum cluster of size 0. On the other side, the distribution has a very long and thin tail, the 0.9 quantile is equal to 1816 while the 0.99 one is equal to 10198 (the observed maximum was 68421). For comparison these quantiles are equal to 31.1 and 37.2 for the maximum statistic.

Finally, the estimated p-value is lower for the maximum statistic than for the cluster size (respectively 0.018 and 0.044). However we have seen previously that the significance may depend on the thresholding parameter α .

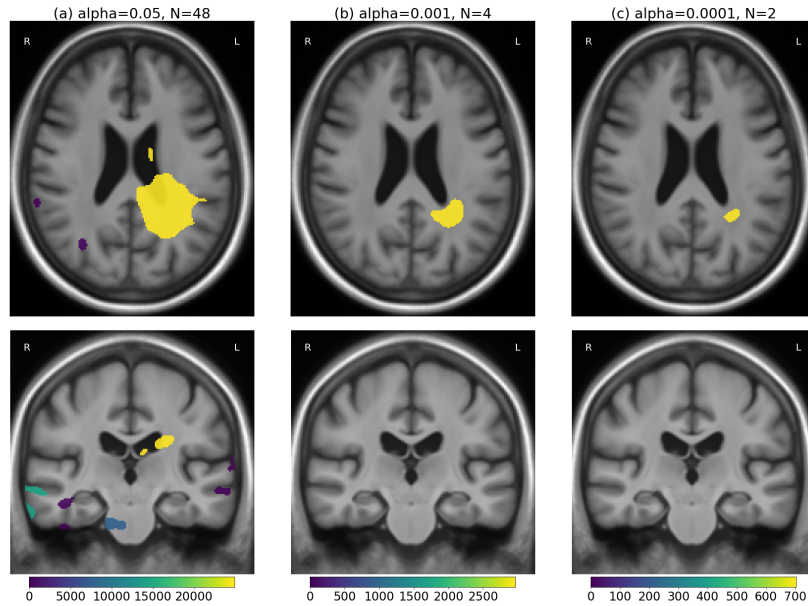


Figure 4.6: Clusters for the MDI test for different thresholding level. The color encodes the size of the clusters (several cluster can then have similar color). For a low threshold ($\alpha = 0.05$) many clusters are visible ($N=48$) but many of them are small with only a single large one (in yellow). These smaller clusters disappear when the threshold increases and the main cluster shrinks.

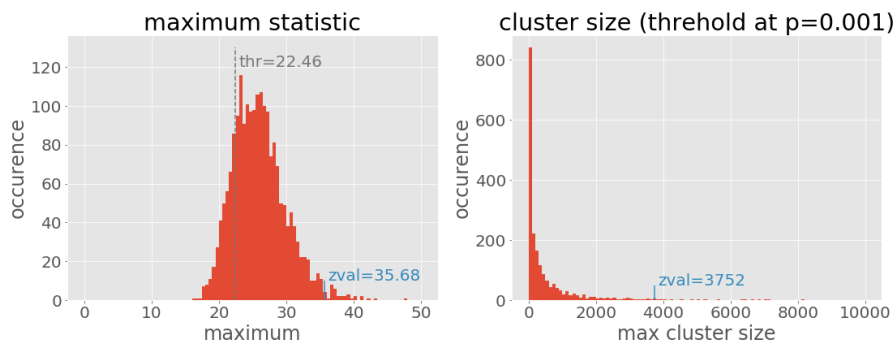


Figure 4.7: Empirical H_0 distribution of the two statistic (max cluster size and maximum) for 2000 permutations. The actual z-value is reported in blue, the corresponding empirical p-values are respectively equal to 0.018 for the maximum and 0.044 for the cluster size. The dashed vertical line one the right plot shows the level used for clustering (on the left)

4.5 Discussion

4.5.1 Statistical analysis parameters

The three main parameters of our model are the choices of the regression model, the voxel-wise statistic and the summary statistic. In this section we discuss their importance and possible alternatives.

4.5.1.1 Regression model

In this work we use a linear regression model but more complex models could be considered. First, the model could include, in the linear combination, a quadratic term or more generally any non-linear function of the predictive variables. On one hand, a simple model does not take into account the interaction between the variables. On the other hand, the addition of non-linear terms should be well-thought to justify for the additional complexity. In particular, in this experiment, the signal is already very weak compared to the total variability and eventual second order effects would be hard to observe.

Second, the model could regress the deformation at a larger scale (and not a voxel only) or using any basis decomposition of the SVFs. This approach would require a more precise understanding of the nature of the deformation and of the observation objectives.

Third, the model could add a regularization term on the input or the output variables. In particular, a sparse penalty could be interesting when the number of variables becomes large. In our case the regression problem is of low dimension and a regularization does not seem necessary.

4.5.1.2 Voxel-wise statistic

The likelihood ratio statistic belongs to a larger family of statistics on the generalized linear model. We showed that a model selection approach using AIC would result in similar findings. We found that the usual MANOVA statistics, such as the Lawley-Hotelling's or the Pillai's trace, produce similar results. These statistics rely on a slightly different metric on the covariance matrix and are equivalent for such small changes.

More generally any metric on HE^{-1} or even functions on H and E could be used. The definition of metrics and pseudo-metrics on symmetric definite matrices and their relations is an interesting topic that goes beyond our practical application. Using a few examples, we found that the results are generally similar even if the relative importance of the voxel significance can change. Finally these examples highlight the limitations of the current approach that focus on a specific measure of the differences.

Another approach to this problem is to directly consider the metric used on the deformation. The L^2 metric that corresponds to the Gaussian residual used here is not optimal to compare deformations. The geometry of the underlying images and

of the mean evolution make the problem strongly anisotropic. This aspect should be explored in future works.

4.5.1.3 Summary statistic

We evaluated the two solutions proposed in [Nichols and Holmes \(2002\)](#) (i.e. maximum statistic and supra-threshold cluster size) and they produced similar results. The regularity of the field is such that a high maximum correspond to a large cluster. An advantage of the maximum statistic is that it does not need an arbitrary threshold parameter. The cluster approach may be less sensitive to spatial inhomogeneity in voxel distribution. However no such problem was observed here (and, for example, a visual inspection shows the locations of maximum for every permutations were uniformly distributed over the whole brain).

From the computational point of view, the clustering and the size estimation step adds an additional computational cost that can becomes prohibitive when iterating over lots of permutations.

Overall, the maximum statistic seems easier to use, more interpretable and more reliable. We argue that these observations are probably dependent on the kind of data that we use. The fields are smooth and even the variability and the noise is strongly spatially correlated. Moreover the cluster aspect is not clearly meaningful for a continuous representation of a deformation. We expect the results to be different for fMRI for example.

We note that the maximum statistic on region (the whole brain for example) can be relaxed to a more general family that allow to weight differently every voxel. If S is the statistical map and m a weight map in $[0, 1]$ then we are looking for a threshold z such that $P(\exists i, m_i S_i > z) \leq \alpha$. If $m=1$ (or at least is binary) it is equivalent to the maximum statistic.

Other alternatives are possible: sum over the image, full likelihood (ie product) etc.. However they often loose the spatial information and are by consequence less informative and less sensitive to a change localized in a small region.

4.5.2 Interpretation of morphological differences

Being able to highlight morphological correlates to a clinical changes is only half the battle. The deformation estimation is not homogeneously accurate in every part on the brain. The differences can be measured in one area while they are hidden in the noise in another one. Moreover, the p-values are a poor indicator of the importance of the changes. This results that are not always interpretable. For example, using a similar approach, testing the effect of Alzheimer's disease (large morphological changes) shows differences in the whole brain giving no spatial information about the disease.

Even when the changes are well localized, it is not obvious to compare the effect of several variables in a specific regions. Are they concurrent? opposite? do they have a catalyst effect? These interactions have to be studied in additional work

using cross-decomposition or joint modeling analyses.

4.6 Conclusion

In this chapter, we presented in detail a general statistical approach to hypotheses testing on deformation fields. We also discussed possible design choices and their influence on the results. The method was illustrated by an application on the description of the intra-subject longitudinal evolution.

Overall the non-parametric approach opens numerous alternative to traditional statistics but more importantly they can be used to challenge the parametric approaches that have been proven to be susceptible to bias and error in false positive rate estimation (Silver et al., 2011; Scarpazza et al., 2015; Cox et al., 2017).

This work expands the methodological perspectives and clarifies the nature of the findings presented in the article *Voxel based assessments of treatment effects on brain evolution in the MAPT* by Sivera et al. (2019a).

4.7 Appendix: the Likelihood ratio test and Aikake's weights

The likelihood ratio test is strongly related to the AIC model selection. Indeed the AIC is defined as $2k - 2\ln(L)$ where k is the number of parameters and L the likelihood. For the linear model with Gaussian residuals:

$$AIC = 2p(q + 1) + pn \ln(2\pi) + n \ln(|\hat{\Sigma}|)$$

The Akaike weights are defined by:

$$w_i = \frac{\exp(-1/2\Delta_i)}{\sum_k \exp(-1/2\Delta_k)} \text{ with } \Delta_i(AIC) = AIC_i - \min AIC$$

Assuming that the constant model has the minimal AIC we get (otherwise exchange the indices i and 0):

$$\begin{aligned} \Delta_i(AIC) &= AIC_i - AIC_0 = 2p(q + 1) + pn \ln(2\pi) + n \ln(|E|) - 2p(r + 1) - pn \ln(2\pi) - n \ln(|H|) \\ &= 2p(q - r) + n \ln(|E|) - n \ln(|H|) = 2p(q - r) + 2 \ln(\Lambda) \end{aligned}$$

and then:

$$\begin{aligned} w_i &= \frac{\exp(-1/2\Delta_i)}{1 + \exp(-1/2\Delta_i)} \\ &= \frac{1}{1 + \exp(p(q - r))\Lambda} \end{aligned}$$

Remark: more generally the ratio of the weights is:

$$\frac{w_i}{w_j} = \frac{L_i}{L_j} \exp(k_j - k_i)$$

Relations between cognitive assessments and brain morphological evolution in the MAPT cohort

Contents

5.1 Introduction	80
5.2 Data and Method	81
5.2.1 Participants: the MAPT cohort	81
5.2.2 Modeling the brain evolution	81
5.2.3 Statistical analysis of the correlation	81
5.2.4 Cross decomposition of the cognitive and morphological evolution	82
5.3 Correlation between morphological changes and cognitive and behavioral scores	82
5.4 PLS between cognitive variables and morphological evolution	84
5.4.1 Dimension reduction with PCA	86
5.4.2 Cross decomposition with PLS	86
5.5 Discussion	88
5.5.1 Correlations with cognitive scores	88
5.5.2 Cross-decomposition of morphological and cognitive evolutions	88
5.5.3 Limitations and perspectives	88

The interpretation of the observed morphological deformations was a recurrent issue in the previous chapters. The precise localization may be method dependent, the spatial interpretation is made complex by the inter-connectivity of brain areas and the local information in the displacements is not directly interpretable.

In this chapter, we explore the relations between the morphological and cognitive evolution of aging subjects. The objective is twofold. First, aging and neurological disorders cannot be fully understood if we only consider a single point of view. The brain processes are intricate and their interactions are crucial to apprehend the whole evolution. Second, we want to use the cognitive information to inform

our description of the morphological changes in order to label and interpret specific patterns. In particular, methods and processing steps are similar to the ones used in Chapter 3 and this analysis may serve as a reference to better understand the observed treatment effect.

Two methodological approaches are proposed. The first uses the statistical multivariate framework introduced previously and described in Chapter 4, the second relies on a cross-decomposition using partial least squares (PLS). The results show realistic evolution patterns that open perspectives on the description of aging from the cognitive and morphological standpoint.

5.1 Introduction

Aging is associated with many neurological processes from cognitive decline (Levy et al., 1994) to functional (Bishop et al., 2010) and structural (Good et al., 2001; Long et al., 2012) brain changes. These changes are often described separately but the description of their interactions have also led to a better understanding of the ongoing brain mechanisms (Park and Reuter-Lorenz, 2009).

In this work we are more specifically interested in the relations between the cognitive evolution and the brain morphological changes. Indeed, on one hand, the cognitive state can be partially assessed using clinical tests but these tests are partial, can be time-consuming and the followup assessments can be biased due to repetition (Collie et al., 2003). On the other hand, morphological changes can be measured precisely and with reproducibility. They give (partial) information on the whole brain with high resolution but this information may not be directly interpretable.

Correlations between morphological measurements and cognition have already been established for healthy older adults (for example between memory and hippocampal volume (Rosen et al., 2003; Bjørnebekk et al., 2010)). However these analyses are often restricted to a specific brain area or cognitive change. Moreover these structure-cognition associations for healthy subjects are not easily replicated and appear sensitive to the studied sample and choice of cognitive measures (Raz and Rodrigue, 2006).

Voxel-wise correlations were studied using voxel-based morphometry for various pathological and clinical condition: the relation with verbal and non verbal memory in fibromyalgia patients (Luerding et al., 2008), the CERAD test and other dementia related assessments in the context of Alzheimer’s disease (Santos et al., 2011) or the correlation between chemotherapy-induced structural changes and cognitive impairment using DTI (Deprez et al., 2011).

Deformation-based approaches have also been proven useful to describe morphological evolution and signs of correlation between specific patterns of evolution and cognitive changes have been highlighted (Sivera et al., 2019a). In this work we want to go further in this direction and use a deformation-based approach to propose a comprehensive voxel-wise description of the general evolution of the brain

morphology in relation to several clinical assessments. The analysis is performed on a cohort with subjective memory complaints from MAPT. We then focus on the relations between the evolution of the patient clinical scores and the morphological evolution over the 3 years of the study.

The analysis is divided in two steps. The first uses the statistical multivariate framework introduced in Chapter 4 to assess for correlations between cognitive evolution and local morphological changes. The second relies on a cross-decomposition using PLS to describes modes of longitudinal evolution observed in the MAPT cohort. The first section introduces the MAPT cohort, the image processing and both statistical analysis methods. The results are then presented for the correlation assessments and the cross-decomposition. Finally the interpretations of the results, their implications and their eventual limitations are discussed in the final section.

5.2 Data and Method

5.2.1 Participants: the MAPT cohort

The same cohort of subjects from MAPT than in Chapter 3 is used. The total sample includes 376 subjects that belong to the MDI plus omega-3 group (N=98), the MDI plus placebo group (N=100), the omega-3 alone group (N=87), or the placebo control group (N=91). Every subject has two MRI acquisitions at baseline and 36 months. In addition to imaging data, the longitudinal followup of the participants includes classical clinical evaluation scales to assess the physical and cognitive functions.

5.2.2 Modeling the brain evolution

This work aims to complete and to extend our previous work presented in Chapter 3. In particular, the image processing steps and the deformation modeling are identical to the ones used in that study. Our deformation-based approach relies on a description of the longitudinal evolution using stationary velocity fields (SVFs) estimated through non-linear registration between pairs of images. The individual longitudinal intra-subject SVF are then transported in a common reference space where they can be compared (Hadj-Hamou et al., 2016).

5.2.3 Statistical analysis of the correlation

In order to relate the morphological changes to clinical variables, we chose to assess the correlation between the deformation and several cognitive scores. The statistical analysis is based on a voxel-wise likelihood ratio test for the general linear model. The statistic compares the predicting power of the linear model relatively to the constant mean. The model and its limitations are more extensively described in Chapter 4.

For the MAPT cohort, the clinical variables evaluated are the evolution of five clinical scores: the Mini Mental State Examination (MMSE), the Grober free recall

test, the category naming verbal fluency test, the Wechsler Adult Intelligence Scale (WAIS) and the Geriatric Depression Scale (GDS). A subset of the GDS items (denoted GDS-apathy) that is expected to be more related to the motivation of the patient is also included (items 2, 9 and 13).

5.2.4 Cross decomposition of the cognitive and morphological evolution

To explore more in depth the relations between morphological changes and cognitive evolution, a cross-decomposition is performed by a projection on latent structures (PLS) algorithm (Wegelin, 2000) using the scikit-learn library (Pedregosa et al., 2011). For computational reasons, PLS is performed on projection of the data on reduced dimensional space computed by PCA. We keep only the first modes of the PCA such that they represent more than 99.9 percent of the variance. It would have been possible to keep all the modes (ie n=376 the number of subjects) to get exact results but with no meaningful differences. Moreover we use a more restrictive brain mask that excludes the cerebellum and the brain stem where the more important registration artifacts are found.

The clinical variables included are the MDI treatment and the evolution of the five clinical scores previously introduced (MMSE, Grober free recall, category naming verbal fluency, WAIS and GDS). A z-score is computed for each variable in order to normalize the inputs.

To evaluate the stability of the PLS coefficients, a population subsampling is performed. We repeat the estimation with two-thirds of the population randomly selected.

5.3 Correlation between morphological changes and cognitive and behavioral scores

The relation with the 3-year morphological changes is independently tested for each clinical scores using the likelihood ratio test for multivariate images introduced in chapter 4. The z-value maps are presented in Figure 5.1 and the corresponding p-values are reported in Table 5.1. These results were obtained without taking into account both treatments. Similar results are found when using a linear correction for the MDI.

test	fluency	MMSE	GDS (apathy)	GDS	Grober	WAIS
p-value	0.117	0.001	0.003	0.126	<0.001	0.014

Table 5.1: P-values for multivariate tests of the relation between morphological changes and cognitive or behavioral clinical assessment. The p-value are estimated with 2000 permutations in order to get a standard deviation of 0.005 on the estimation at the 0.05 level. The significance, in bold, takes into account the Bonferroni correction for the 6 tests.

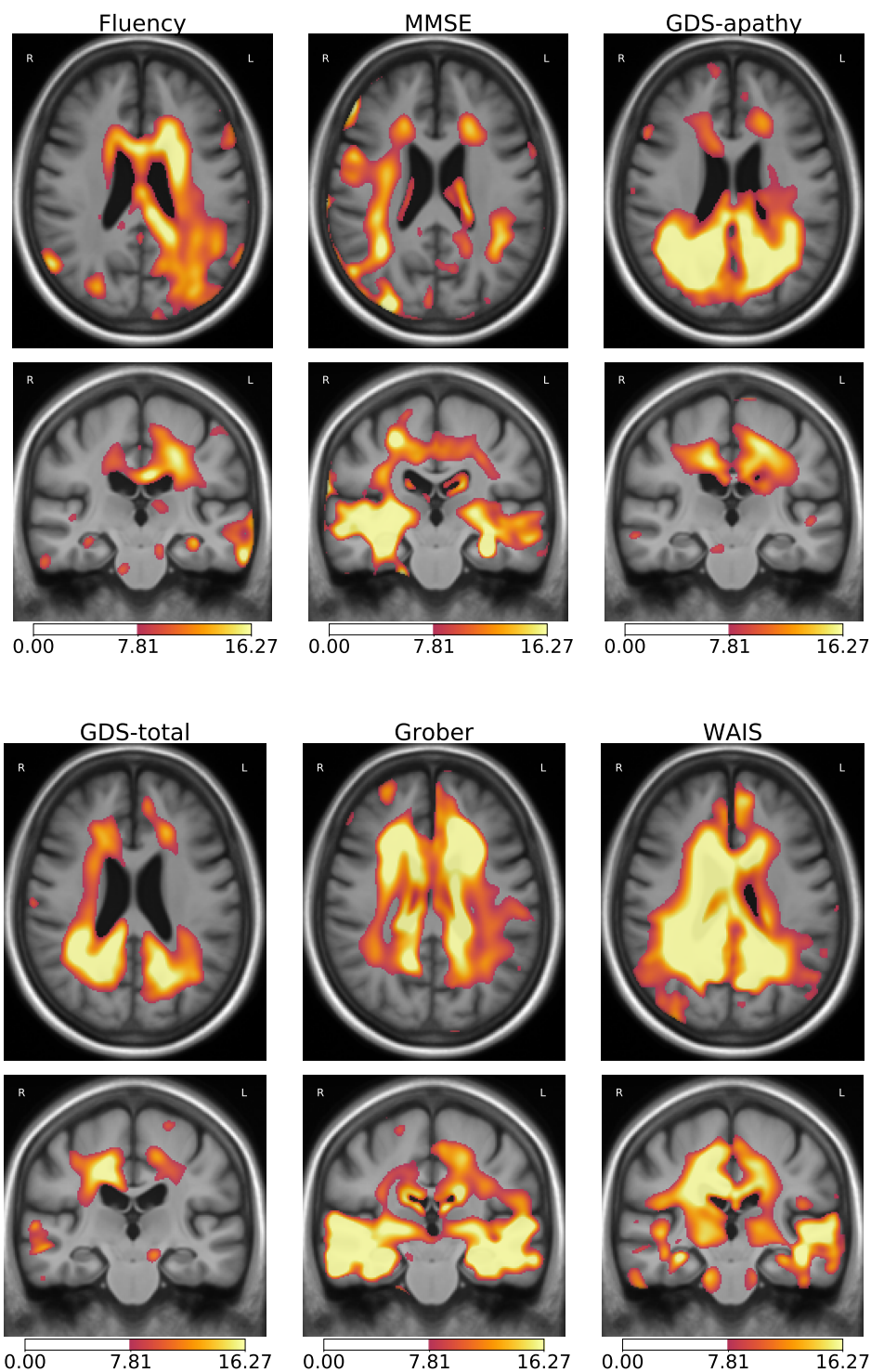


Figure 5.1: Voxel-wise z-values for the test of the correlation between the cognitive score and the SVF parametrizing the deformation. From top left to bottom right: fluency, MMSE, GDS-aphathy, GDS, Grober and WAIS tests. Intensity thresholds correspond to the independent parametric testing at level 0.05 (lower threshold) and 0.001 (higher threshold).

First we can note that the effect is significant for several clinical assessment (see Table 5.1). This means that we are able to find correlations with most of the usual cognitive measures. In addition to this global finding, the localization of the differences can be observed. The fluency is associated with changes in the peri-ventricular area with an emphasis on the left hemisphere. MMSE is strongly correlated with changes in the hippocampi and medial temporal area, GDS with parietal areas close to the pre-cuneus. The Grober test is associated with changes in the temporal lobes but also in the ventricles probably sign of the global atrophy. Finally, the WAIS test correlates with changes in a large region of the brain in the white matter and the ventricles from the frontal to the parietal lobes with an emphasis on the right hemisphere.

These changes can be interpreted regarding the cognitive function assessed. It is expected to find language related tasks more associated with the left side and memory exercises associated with temporal areas. This consistency gives a good idea of the accuracy and the resolution of the localization of the morphological changes even without any ground-truth on the actual deformations in each area.

In this regard, the localization of the changes associated with GDS is the one that is most similar to the one observed for the MDI (see Figure 3.1) even if the correlation in the right hemisphere is relatively stronger here.

Beyond highlighting areas of difference, we can also represent the difference of longitudinal evolution correlated with these scores (see Figure 5.2). The localization of the deformations matches the localization of the differences shown in Figure 5.1. This would indicate that the variance of the noise is roughly uniform and the statistical differences are only visible where the amplitude of the signal is stronger.

The second observation concerns global atrophy: every effect goes in the expected direction. An increase of any score (or a decrease for GDS) correspond to a decrease of the ventricular expansion.

Beyond this agreement, each field shows its own specificity. For the fluency, the deformation is localized mainly exclusively in the lower part of the frontal lobe. The deformation in the temporal lobe associated with MMSE seems to be focused on the hippocampi while the rest of the lobe is more affected for Grober. GDS shows a more specific change next to the temporo-parietal junction and the effect is really stronger when considering only the apathy-specific items.

Finally, we note that the patterns are generally symmetric. A clear asymmetry is only visible for the fluency (mainly on the left frontal) and for WAIS (mainly on the right frontal lobe and the left temporal lobe).

5.4 PLS between cognitive variables and morphological evolution

The correlation maps between the morphological changes and cognitive evolution suggests each variable describes a particular evolution that can also be observed at the anatomical level. In this section we directly model these interactions using PLS.

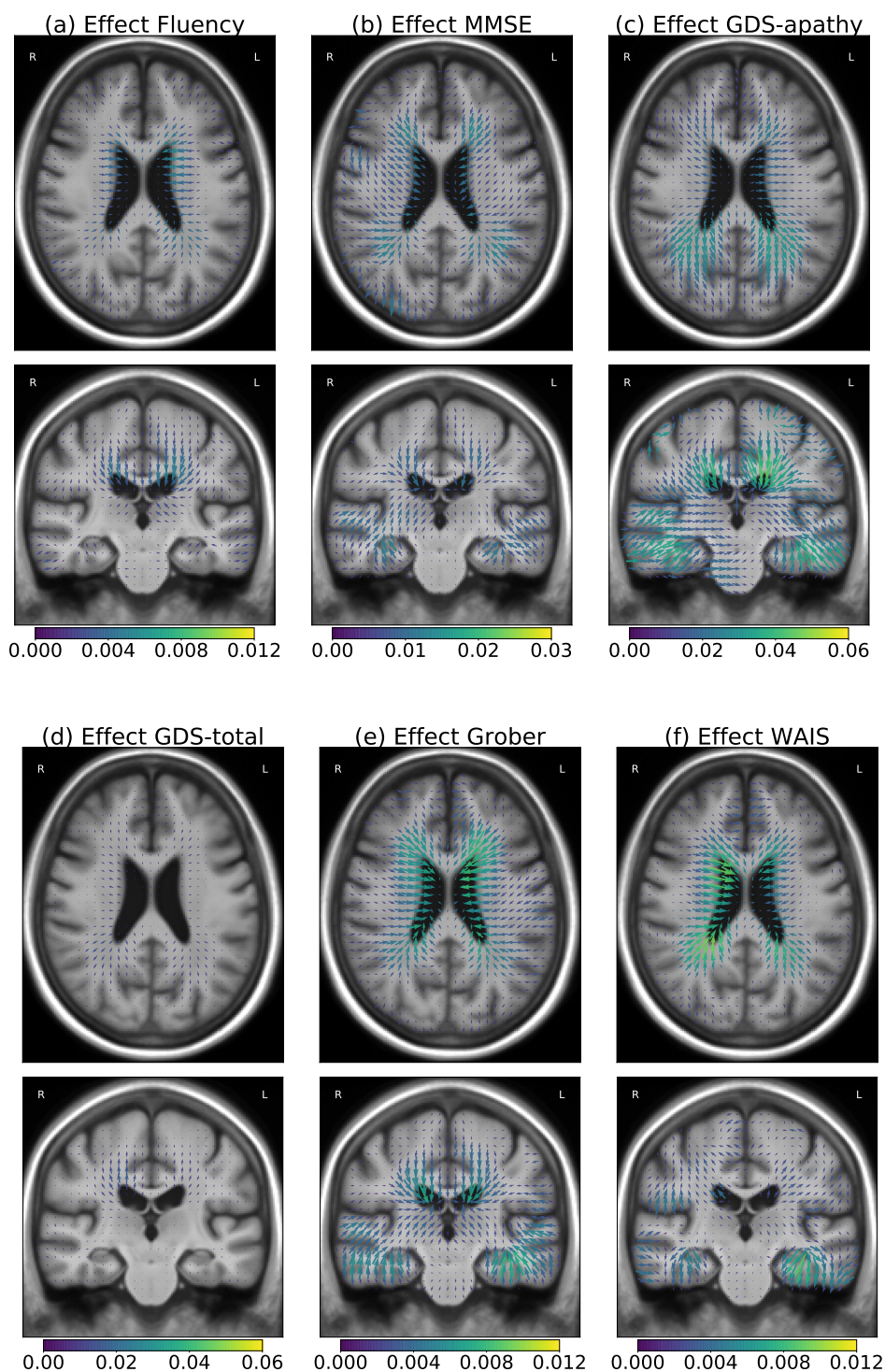


Figure 5.2: SVFs parametrizing the difference in longitudinal evolution associated with several clinical variables. From top left to bottom right: fluency, MMSE, GDS-apathy, GDS, Grober and WAIS tests.

5.4.1 Dimension reduction with PCA

The main objective of PCA here is to reduce the dimensionality of the data for computational reasons without affecting the PLS results: the PLS is performed on the projection of the data computed by the PCA. The variance is mainly concentrated on the first few modes. In particular the data was not centered and the the first mode is really close to the direction of the mean. Even with centering, this axis would encompass a large part of the variability, meaning that the speed of evolution is a crucial information. In the following, we keep the 62 first modes representing 99.9 percent of the variance (see Figure 5.3).

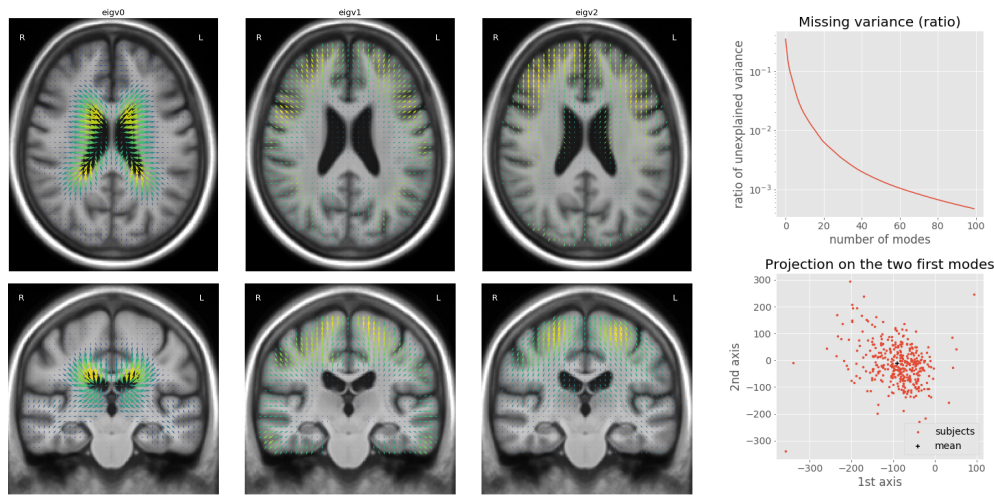


Figure 5.3: Left: First modes of the PCA (normalized). Top right: visualization of the explained variance as a function of the number of modes (log-scale), 3 modes explain 89% of the variance, 62 modes explain 99.9%. Bottom right: Projection of the subjects on the axes defined by the two first modes (and the mean in black).

The first three modes are shown on Figure 5.3. The first mode shows the ventricular expansion and the corresponding deformation in the brain tissues. The second mode could correspond to a global atrophy of the brain but can also be due to misalignment before registration. The third mode shows more deformation in the pre-frontal lobe and in the superior frontal gyrus. Of course these descriptions are simplification (in particular beyond the first mode): no spatial structure is imposed on the PCA and the deformations are spread in the whole brain.

5.4.2 Cross decomposition with PLS

Figure 5.4 summarizes the results of the PLS decomposition. The estimation is repeated 500 times using randomly drawn 2/3rd of the subjects. An additional step to optimize the sign consistency between all the estimations is done at the end. This bootstrapping approach gives an insight about the reliability of the estimation. The results are shown for 4 PLS components.

The boxplots show the distribution of the coefficients for the clinical side. The first mode is associated with a general cognitive decline affecting every variables and the variability between the samples is low. The second mode is more associated with the MDI treatment, the MMSE and most of all the GDS. This can be described as a everyday-life/behavioral factor. The third mode is strongly associated with the MDI treatment, and it seems to distinguish MMSE and Grober from fluency and WAIS, although the coefficients are varying and this observation is not stable. The fourth mode is even less interpretable. The coefficients have large variability across samples and we cannot reliably interpret the results.

These observations are summarized in Figure 5.4(b,c) using the mean and the probability of staying of the same sign for each coefficient. This representation, and especially the sign probability, make it clear that the 4th components is not reliable. This figure also shows the same statistics for the morphological coefficients. In particular we see that the first mode of the PLS is strongly associated with the first mode of the PCA.

Figure 5.5 shows the representation of the 4 morphological modes and the atrophy maps corresponding to these modes computed using the log-Jacobian of the SVFs (some artifacts are visible at the boundaries because the PLS modes were computed using masked versions of the original SVFs).

As expected from the coefficients, the first mode shows some similarities with the mean evolution. It suggests that the main variability is related to the speed of the main evolution due to age, both for the morphology and the cognition. However several differences are noticeable. The global atrophy is really diffuse with less emphasis around the ventricles. A contraction of the ventricles and an expansion pattern in the parenchyma are visible. This is in agreement with the improvement of the cognitive scores in Figure 5.4 : a strong cognitive decline would be associated with a strong atrophy. The SVF and the atrophy map also show an important deformation in the hippocampal areas. The focus of the deformation on these regions, relatively to the mean evolution for example, highlights their importance in the cognitive performance measured by the clinical tests. We could also note that some patterns such as atrophy of parietal lobe and in the hippocampal area and part of the frontal lobe resemble the usual Alzheimer's patterns.

The 2nd and 3rd modes are not obvious to interpret and look a bit noisy. The main characteristics are nevertheless interesting to complement the first mode. The second mode contains atrophy variations with higher spatial frequencies (more sign changes) than the first mode. The combination of these two modes then characterizes the spatial extent of these concentric expansion and contraction patterns. The third mode seems to be oriented along the transverse axis with relatively low volumes changes along the antero-posterior axis. Some asymmetrical deformation patterns are also visible (in particular in the parietal lobes) but are barely reflected by the atrophy.

Finally, the 4th mode is characterized by a strong left/right asymmetry. As this mode was not significantly associated with any cognitive evolution it could be caused by misalignment during the processing of the images but could also indicate

some reconfiguration of the brain between the two hemispheres.

5.5 Discussion

5.5.1 Correlations with cognitive scores

Using multivariate statistical analysis, we are able to spatially correlate morphological changes and cognitive variables. These correlations inform us on the existing functional dependencies in the brain using morphometric features in coherence with the general knowledge of the brain anatomy. More importantly, these observations highlight the relation that are involved in the longitudinal evolutions of elderly subjects. This approach could help understanding how every regions is affected and what are the different realizations of the consequences of aging.

Regarding the MAPT study, doing this analysis with the same data, processing and statistical approach provide insightful examples that can help us interpret the observed MDI effect from the clinical point of view. In particular, we see that the GDS and more specifically the selected items show a pattern of evolution similar the one observed for the multi-domain intervention (Sivera et al., 2019a).

5.5.2 Cross-decomposition of morphological and cognitive evolutions

The PLS leads to a more complex description of the evolutions of the subjects. It highlights several modes of evolutions: the first describing the general cognitive decline, the second focusing more on behavior and depression and the third being associated with the MDI treatment without any strong association with the assessed cognitive variables.

Interestingly, the morphological changes that correspond to these cognitive evolutions do not directly match the correlation observed previously. This suggests numerous dynamics that are not always related. However, the morphometric PLS modes are generally difficult to interpret and a regularization, for example using a sparsity inducing prior, should be included to improve subsequent analyses.

5.5.3 Limitations and perspectives

First, we should note that both the analyses proposed in this work can be sensitive to the pre-processing steps used to estimate the deformations. In particular, the processing pipeline assumes that the head of a subject does not change size. Consequently, the intra-subject alignment, required before the non-linear registration, is rigid in order to limit the introduction of potential artifact. However due to the calibration of the MR scan or local distortion we observed several cases where an affine alignment would have been justified. These small errors can have a large effect, for example regarding the estimation of global PCA modes, as a small rigid misalignment creates a relatively large deformation in term of L_2 norm.

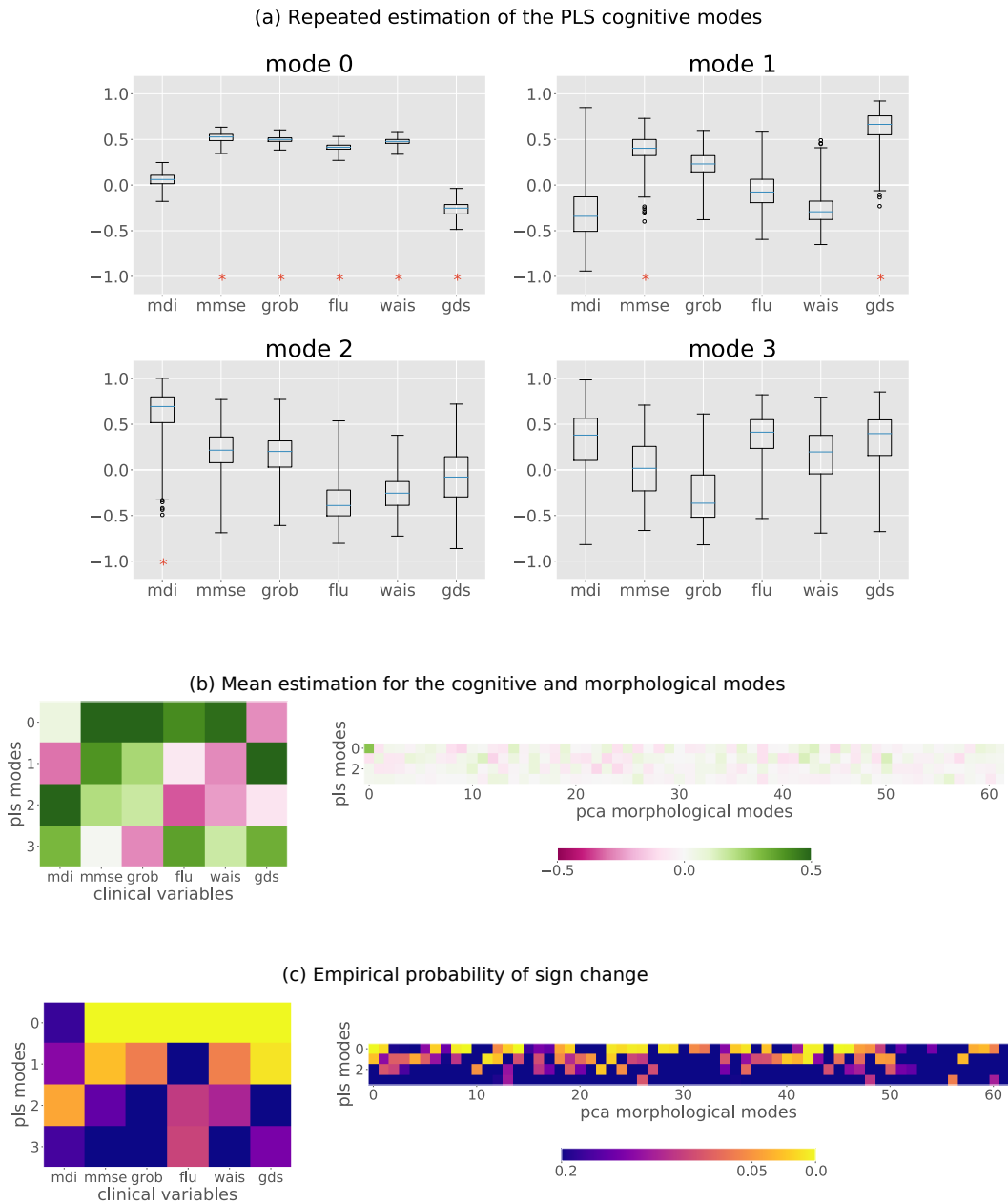


Figure 5.4: PLS coefficients for the clinical variables and the morphological evolutions. The clinical variables used are the MMSE, Grober, fluency, WAIS and GDS cognitive scores and the MDI treatment variables. The morphological variables correspond to the projection of the longitudinal SVF on the 62 first PCA modes. The estimation is repeated 500 times with 2/3rd of the populations randomly drawn in order to estimate its consistency. (a) The boxplots show the distribution of the estimated coefficients for the cognitive modes. The first mode shows general cognitive decline, the second mode corresponds to a specific subtype of cognitive changes, the third one is more associated to the MDI, and the fourth one is not consistent between the different estimations. (b) The mean of the estimation is represented for every coefficient: the clinical side partially reproducing information in the boxplots, the morphological side showing a dominating first coefficient (first PLS mode along the first PCA mode). (c) The sign change probability associated with a coefficient corresponds to the proportion of the estimations that are of opposite sign than the mean. Low values (below 0.05) are mainly found for the first modes, this information is reported in (a) using red stars.

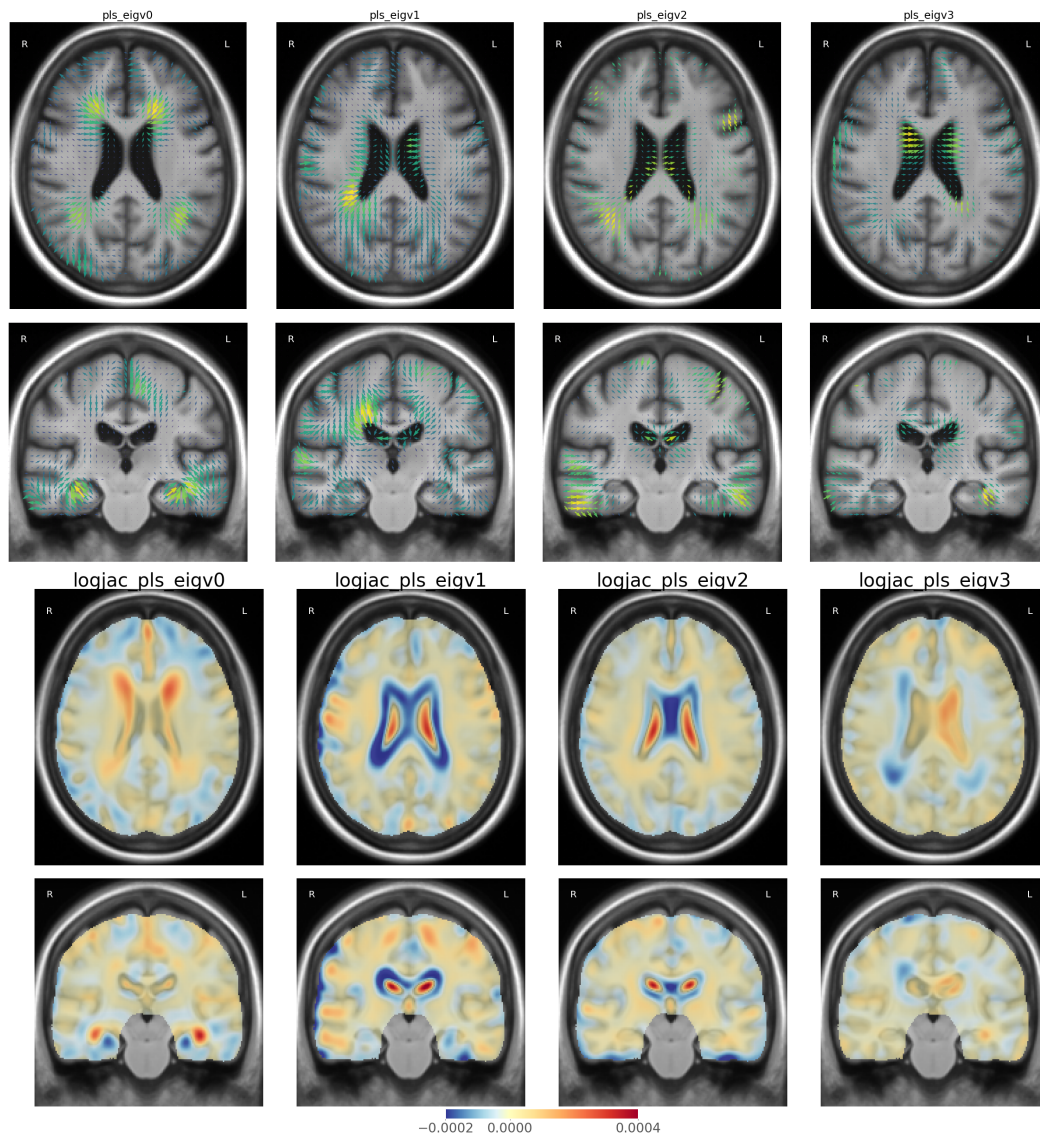


Figure 5.5: Reconstitution of the morphological PLS mode in the SVFs space (top rows) and corresponding atrophy maps (log-Jacobian in bottom rows).

Similarly the exclusion of the cerebellum and the brain stem for the PCA/PLS analysis is arbitrary and only related to inconsistencies of the processing software in the estimation of the deformations in this area. Moreover, this problem highlights the limitations of the decomposition algorithms that rely on global metric such as PCA or PLS.

More generally, the potential noises and biases in the estimation of the deformations are not spatially uniformly distributed and can affect the power of our analyses. Indeed, the observed localization of the differences is impacted by the local effect size. In our case, this classical statistical remark directly affects our interpretation of the results: the localization of the changes is dependent of the amplitude of the changes and the ease to estimate the deformation. Moreover, this work highlights the importance of understanding the nature of non-volumetric changes: the representation of the deformations and their characteristics (using atrophy for example) affects our ability to fully characterize and interpret the brain morphological evolution.

Finally, these analyses should be replicated the analysis on distinct datasets. It would be really interesting to assess the reliability of these results as such structure-cognition associations can be really sensitive to the studied sample and choice of cognitive measures (Raz and Rodrigue, 2006). The MAPT study focuses on a specific population of patients with subjective memory complaints and the importance of the treatment effects is difficult to quantify. In particular, the parietal deformation pattern could be related to the effect of the multi-domain intervention (Sivera et al., 2019a) and its importance could be specific to this dataset.

Conclusion and Perspectives

Brain morphometry can provide useful information for the clinical insight. The observations of the brain and spatially-informed quantitative measurements are crucial for the assessment of the brain condition and the possible development of a neurodegenerative diseases. The objectives are multiple: disease modeling, treatment evaluation in a clinical trial, diagnosis, prognosis or shape extraction for personalisation. In this thesis, we proposed several new approaches for the characterization of the brain morphology of elderly subjects that leverage on the deformation modeling of the morphological changes.

The following section summarizes the main contributions of our work. These contributions towards a more extensive use of brain morphometry and the limitations of the proposed methods give new insights on the considered problems. To this end, we conclude this manuscript with some possible research perspectives of this thesis.

6.1 Summary of the main contributions

Modeling the effects of aging and Alzheimer’s disease on the brain morphology Aging is not a simple risk factor for Alzheimer’s disease. The effect of both factors on the brain is extensively studied and the questions about the role of aging mechanisms in the AD symptoms or the modulating aspect of age on the disease progression (especially for late onset subtype) are also under investigations. In chapter 2, we propose a novel approach that jointly model the effect of both aging and disease in a non-linear deformation space.

For this model, the focus was set on the interpretability of the description. Two scalar markers (the morphological age and the disease score) were used to summarize the progression of these two major processes. We showed that this simple description based on cross-sectional assessments could be used for the characterization of realistic longitudinal evolutions for several clinical conditions. In particular, this aspect of our model was illustrated in a personalized image simulation setting. Despite our approach limitations – the parameter estimation scheme should be optimized and more individual variability could be taken into account – the deformation-based parametric representation coupled with an innovative use of longitudinal and cross-sectional registration could lead to rich insights on the disease progression.

Exploratory morphometry for the MAPT clinical trial The development of a treatment to cure AD or, at least, slow down the apparition of symptoms is one of the main challenge of neurological research. In this context, the contributions of deformation-based morphometry can be invaluable. The MAPT was a large scale longitudinal study designed to assess two potential candidates: an omega-3 supplementation and a multidomain intervention.

Chapter 3 presents the results of our analysis of MAPT morphometric data. In particular, a significant effect of the multidomain intervention on the brain morphology was found using the statistical analysis of the longitudinal deformations. Moreover, we are able to relate this effect to a slower morphological evolution and better performance in cognitive assessments.

Statistical framework for hypothesis testing on deformations Chapter 4 is dedicated to the presentation of complete and general statistical framework that can be used for hypothesis testing in the context of multivariate images and deformation fields. The statistical analysis of medical images is well established since the early 2000s but it is less developed for multivariate images and often relies on strong hypothesis about spatial correlation.

The main contributions of this work is the synthetic presentation of a framework for high-dimensional multivariate images that is easily generalizable and does not require strong ad-hoc hypotheses. This framework enables the use of complex morphometric features while keeping the advantages of the VBM regarding the wide spectrum of possible applications.

We also discussed the use of unorthodox non-parametric statistics and further investigations should be done to precise the optimal choice for the comparison of deformation fields and SVFs.

Relation between morphology and cognition in elderly patients Spatial correlation between morphological features and cognitive function in elderly subjects may shed light on the mechanisms leading to cognitive symptoms in Alzheimer's disease.

In this work, we highlighted specific patterns related to clinical assessments of various cognitive functions (memory, language, reasoning, etc.). The longitudinal evolution of each score is associated with a specific deformation pattern that affect differently the brain regions. We argue that the visualisation of the morphological correlates could help refine the cognitive tests used in the clinical practice.

We also used PLS to highlights several mode of evolution. This exploratory study highlights two modes (one that is more behavior related and one that is associated with MDI treatment specific to the MAPT) in addition to the main trend given by the cognitive decline.

6.2 Publications

This work led to publications and manuscript submissions in peer-reviewed journals.

- *A model of brain morphological changes related to aging and Alzheimer's disease from cross-sectional assessments*, Raphaël Sivera, Hervé Delingette, Marco Lorenzi, Xavier Pennec, and Nicholas Ayache. *NeuroImage*, 2019.
- *Voxel based assessments of treatment effects on brain evolution in the MAPT cohort*, Raphaël Sivera, Nicolas Capet, Valeria Manera, Roxane Fabre, Marco Lorenzi, Hervé Delingette, Xavier Pennec, and Nicholas Ayache. submitted to *Neurobiology of Aging*, 2019.
- *Non parametric statistical approach to the General Linear Model for deformation based morphometry*. Raphaël Sivera, Hervé Delingette, Marco Lorenzi, Xavier Pennec, and Nicholas Ayache. to be submitted, 2019.

6.3 Perspectives

6.3.1 Describing the brain deformations

A recurring question we faced in this thesis was how to synthesize the important information from our SVFs. This issue concerns the modeling of the evolution with the choice of the metric for the projection and the parameter estimation in chapter 2, likewise it occurs in the choice of voxel-wise statistic (see chapter 4) and in the characterization of the changes relatively to cognitive variables in chapter 5 for example.

We believe this question belongs to a even larger problem related to the choice of the parametrization of the deformations. The parametrization of diffeomorphisms by tangent vectors is a convenient approach; the Lie algebra structure allows to easily perform operations between the deformations (composition, inversion, mean, etc) and the space of deformation is large enough to encompass enough variability to precisely describe all the subtle morphological differences. However the parametrization using SVFs is suboptimal regarding the dimensionality and the interpretability, and generally, a similar problem exists for other deformation frameworks such as LDDMM if we use a dense vector field to parametrize the momenta.

The SVFs estimated through registration are generally smooth and represent actual brain deformations or inter-subject matching. The resulting space could be seamlessly described using a parametric space of intrinsic dimensionality much lower than the total dimension of a dense 3D vector fields (usually around 20 millions for 1mm isotropic images of a brain). This redundancy is obvious when performing hypothesis testing. Each voxel test is strongly correlated with its neighbors and is not informative by itself. Such observations push us to look for a description of the SVF at a higher spatial scale. More generally, non-local information may be crucial to understand an evolution or to take advantage of spatial redundancy.

These questions have already been addressed from several points of view. More compact representation of deformations have been proposed for example using control points and kernels or spline representations. These representations can lead to more interpretable local components (McLeod et al., 2013; Gris et al., 2015) or frequency bandwidth constraints (Zhang and Fletcher, 2015). Additional constraints such as sparsity can also be included (Fishbaugh et al., 2013). Some approaches are more data-driven with the use of dictionary, principal geodesic analysis (PGA) or variational auto-encoder to provide local description (of patches) or to estimate the low-dimensional subspace associated with a population of shape (Zhang and Thomas Fletcher, 2015; Krebs et al., 2019). The incorporation of these descriptions in our models could lead to faster processing and possibly better interpretation of the results.

The description of the deformations could also be driven by a bio-mechanical model leading to feasible deformations only. In this thesis we used a brain atrophy simulator (Khanal et al., 2016), but the inverse problem that relate our observation to deeper structural changes (in particular using the anisotropy of the brain tissues) would be really promising. These model-based approaches include knowledge in the modeling and the interpretation but the complexity of the brain and the difficulty to model the ongoing process over several decades will put a particular emphasis on the limitations and the errors of the models. To this end, we should take inspiration from application domains, such as cardiac motion tracking, where the use of models is more prevailing.

6.3.2 Multimodal approach

This work focuses on the morphological evolution, its relation with the two main processes that are aging and AD, and with the cognitive evolution. Possible future developments could include more diverse information. The disease progression is characterized more extensively when combining several biomarkers – including the morphometrical changes – and the relation between the different aspects of neurodegeneration is an interesting research topic.

At an elementary level, the multivariate statistical framework can work with additional channels provided by other imaging modalities. Even when only T1 images are available, it is possible to extract textures and complementary intensity-based information that can be related to the progression of AD (Hett et al., 2017). This could be straight-forwardly implemented but would raise questions about the meaning of the local statistic and the multi-scale interactions (caused by the various image smoothness) on the multiple test comparison.

The problem becomes more complex when we want to use more heterogeneous data sources (Antelmi et al., 2019). The additional variables may have their own properties and the description of their longitudinal evolution, in relation with the geometry of the brain, may require a specific modeling approach. Learning trajectory distributions in this complex geometrical framework is then a complex but crucial problem to understand the neurodegenerative diseases (Durrleman, 2018).

In this context, the main challenge of this work would be to conserve the specificity of deformation data while taking into account other kinds of data. To this end,

6.3.3 Spatio-temporal patterns and disease evolution subtypes

Another interesting question that was not directly addressed in this thesis is the description of the diversity of spatial patterns visible in the longitudinal evolutions of the deformations. We have seen that the deformation are not constant over time; in particular in the context of AD, the progression of the disease is mirrored by specific morphological changes. Evidences of pattern changes with age have also been shown in AD and several dynamics are visible depending on the age of onset (Fiford et al., 2018). The description of these patterns contributes to the understanding of the chronology of the Alzheimer’s disease.

At the same time, the interactions of coincident pathologies, genetic and environment backgrounds with neuropathological differences can lead to a large phenotypic variability in the presentation of the disease. This individual diversity is to combined with the complex longitudinal patterns of evolution making the situation of every patient unique at any time. The description of these patterns contributes to the understanding of the chronology of the disease and can lead to the characterization of several subtypes improving the individual follow-up.

Regarding the morphological aspects specifically, this heterogeneity of evolution has been highlighted in AD. Morphometrics studies using cortical and subcortical measurements (Na et al., 2016; Poulakis et al., 2018; Koval et al., 2017) or VBM (Dong et al., 2016) have lead to the definition of several disease subtypes. For example, these studies often highlight disease progressions that are more medio-temporal focused, some that are more diffuse in the whole brain, and some that are even sparing the hippocampi. These subtypes are especially interesting because they have been shown to be related to the patient evolutions with different rate of evolution and cognitive functions affected.

The longitudinal modeling in chapter 2 or the relation to the cognitive evolution in chapter 5 do not take this heterogeneity into account. However we think that the approaches developed in this thesis could be helpful in this context. The analysis of the complete longitudinal deformations could help precise these observations and show patterns that are not visible yet. We think that the presented methods are well-suited for this objective and that disease heterogeneity should be included in our approaches.

Appendix A: Algorithmic contribution to the measure of the brain evolution

Contents

A.1 Robust-FOV	99
A.1.1 Context	99
A.1.2 Model	100
A.1.3 Implementation	101
A.2 Alignment: quality control	101
A.2.1 Segmentation and skull-stripping	102
A.2.2 Residual affine component in a SVFs	103
A.3 Affine invariant LCC-demons registration	103
A.3.1 Local correlation coefficient	103
A.3.2 Evaluation on controlled experiments	104

A.1 Robust-FOV

A.1.1 Context

In order to get a standardized view of the brain, we want to reorient and reduce the field of view (FOV) to the region of interest (ROI). In our case we want the whole brain while the original image often include the neck and a part of the upper body that do not interest us. The extent of the background can also vary between the images.

The FOV estimation is an important step in the pre-processing of the MRI that has been addressed in multiple ways (Wargo et al., 2013). In this work, we want a fast method that can perform FOV reduction without prior registration or skull removal. We chose to use the FOV reduction software used in brain extraction pipeline of the FSL software (Smith, 2002). The FOV reduction looks at the upper axial slices, starting at the edge of the volume and moving towards the center, and determines whether a slice contains only noise or signal plus noise. It stops when it finds consecutive slices with signal. Using this localization of the top of the head, a

fixed height is then used to cover the brain but cut the jaw and the neck. The needed size to cover every brain is set to 142mm (Mennes et al., 2014). The challenge is to accurately detect the top of the skull (neither too high or too low) to avoid partial brain coverage.

However, the software does not perform as well as expected on our dataset and often cut part of the cerebellum (see Figure A.1). By consequence we propose here a method to distinguish between noise and noise plus signal that we use to adapt the software to our needs.

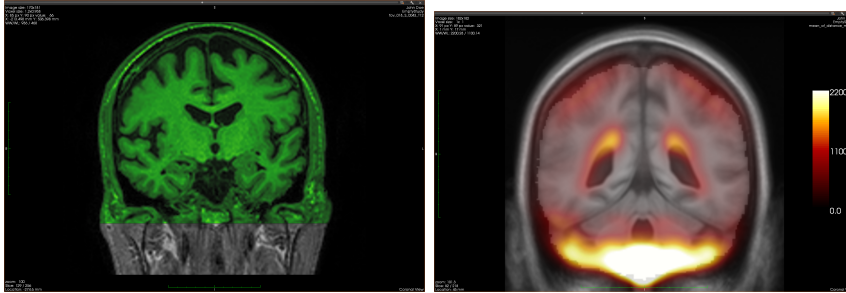


Figure A.1: Left: Example of FOV reduction, in green, on the original image. Right: Total sum of square error of the intra-subject linear model using the vanilla pipeline. The high variability in the cerebellum is caused by FOV reduction that cut the images too high removing part of the cerebellum making the non-linear registration algorithm fail.

A.1.2 Model

In general, the subject only occupies a third to a half of the volume. The rest is background. Trying to describe the actual signal is a complicated task. However the background is much simpler: only air where the variations in intensity are only due to the noise. The description of the noise and/or the denoising in MRI can be a complex problem because of the intensity inhomogeneity and other acquisition biases (Zhuge et al., 2009). However, here, we are only interested in getting a fast and easy background/foreground distinction and the noise in a MRI can generally be modeled by a Rician distribution (Manjon et al., 2008).

The first approach can try to model the intensity distribution with a mixture of a Rayleigh distribution (approximation of the Rician noise for low intensity) and a more complex distribution for the signal plus Rician noise for higher intensity. For example, using a Gaussian distribution, the intensity probability density function, depending on the proportion α of background in the image, would write:

$$f_{\alpha, \sigma_n, \sigma_s, \mu}(x) = (1 - \alpha) \frac{1}{\sigma_n^2} x \exp\left(-\frac{x^2}{2\sigma_n^2}\right) + \alpha \frac{1}{\sigma_s \sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma_s^2}\right)$$

Even if the resulting problem could be solve using EM algorithm, the problem

can be computationally expensive and require strong hypothesis on the intensity distribution of the actual signal.

We choose a second approach that take advantage of the fact that the lower intensity part of the images that constitute the background is large and that the cumulative density function of a Rayleigh distribution is $F(x) = 1 - \exp(-x^2/(2\sigma^2))$ meaning that the quantiles are then equal to:

$$Q(q, \sigma) = \sigma \sqrt{-2 \log(1 - q)}$$

In particular the quantiles are proportional to the noise level given by σ . It is then easy to compare the higher quantiles to the lower quantiles when there is no signal in a slice. On the contrary, any signal will create high intensity aberrations. The last point to consider is the eventual intensity offset that can bias the estimations. Several approaches are then possible using these considerations. Their comparison on the theoretical and practical point of view should be developed in subsequent work. Here we just describe the empirical choice used in this thesis.

A.1.3 Implementation

For every slice, if the head (the foreground) occupies more than an α proportion of the volume, we can write for every the following relation for the quantile in the slice (denoted Q_{sl}) for every x : $Q_{sl}(1 - x) \geq Q_{fg}(1 - x/\alpha)$. The quantile in the foreground are statistically higher than the one for the total image (denoted Q_I) so: $Q_{sl}(1 - x) \geq Q_I(1 - x/\alpha)$. In practice, the idea is to use the 20th and 80th percentiles of the image as the head occupies more than 20% and less than 80% of the volumes, we should get information about the noise and the signal intensity scaling. The test for one slice is:

$$Q_{sl}(0.98) \geq 0.75Q_I(0.20) + 0.25Q_I(0.80)$$

The search is perform from the top and stopped when 3 valid slices are found. A margin is taken into account top and a bounding box of fixed size is then computed as it was proposed in the original software. This method is really empirical and rely on several arbitrary parameters (quantiles choices and weight in the linear combination). To address this problem, we choose to compare the distribution of the estimated quantiles in the slice assigned to the background in order to raise a warning when the thresholding is uncertain. The result can then be checked manually.

This output of the simple statistic used to do the FOV reduction is a really simple aspect but allow for this semi-supervised evaluation of the result that is crucial to guarantee the quality of the results over several thousands of images.

A.2 Alignment: quality control

The affine registration is another important step that can have a dramatic impact on the following steps. The affine inter-subject registration and the rigid intra-subject

are performed fully automatically using the `flirt` software (Jenkinson and Smith, 2001) available in FSL. In this section, we discuss safeguards that can be included in the pipeline to check the quality of the results (and we are not interested in registration algorithms directly).

A.2.1 Segmentation and skull-stripping

Every image is skull-stripped posterior to their alignment in the MNI space. This task can then be used to assess the intra-subject rigid registrations; indeed the brain mask for successive acquisitions of the same subject should be really similar. A similar reasoning can be also done between subject. The anatomies may slightly differ but the mask should overall be superposed. To measure these similarities the Dice score is computed between the mask of the follow-ups and the baseline for every subject and between the baseline images of each subject and the mean of these baseline masks. The results for 1083 images from 300 subjects from the ADNI database are represented in Figure A.2

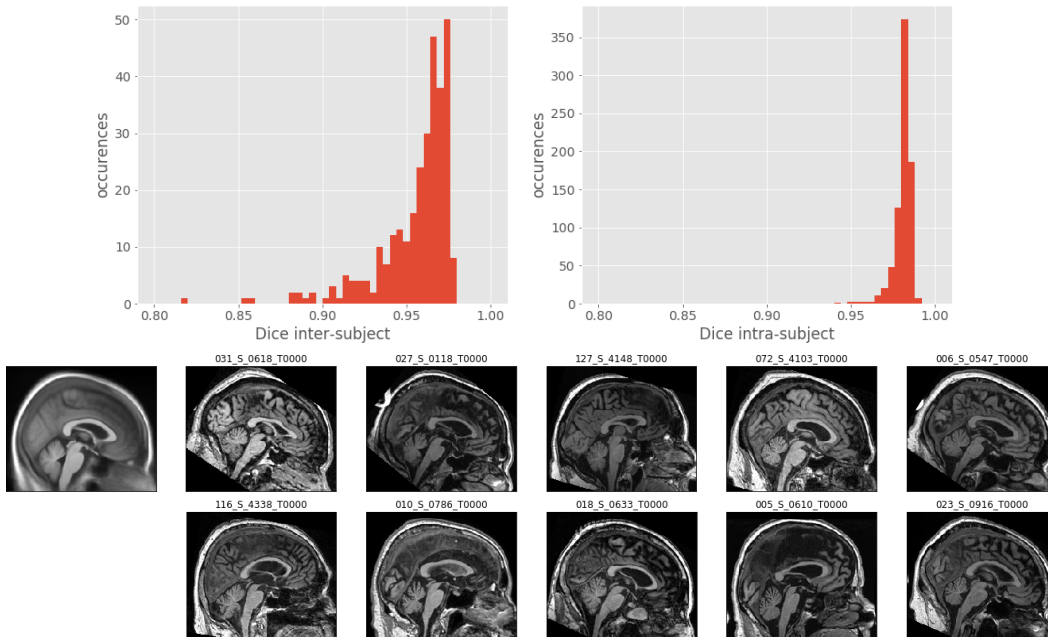


Figure A.2: Top: Dice scores between skull-stripping mask used to assess the results of intra-subject rigid registration (left) and inter-subject affine registration (right). Bottom: Baseline images with a Dice inter-subject below 0.9, the position differences are clearly visible between the images (the first one is the average image)

The images with the lowest threshold can then be registered manually

We can note that these assessments can also help validate the other pre-processing such as the reorientation, the FOV reduction and even the skull-stripping.

A.2.2 Residual affine component in a SVFs

A second registration can help assess the liability of the results. For example, we performed affine registration between intra-subject aligned images to assess for potential scale changes. The extreme cases were manually checked and corrected when possible but most of the time the error is really low and the distortion are not homogeneously associated to a scale change.

Without additional registration, it is also possible to use the result of the non-linear registration. Direct affine transformations can be parametrized by a SVF (given by the logarithm of the matrix of the deformation). Reciprocally, it is possible to estimate the linear deformation that approximate “best” a diffeomorphic deformation using linear regression on SVFs. A large deformation may indicate an error of alignment in the pre-processing.

A.3 Affine invariant LCC-demons registration

A.3.1 Local correlation coefficient

One of the main advantage of the LCC-demons algorithms, introduced by [Lorenzi et al. \(2013\)](#), was the ability to register MRIs in particular of different modality using an affine invariant metric: the local correlation coefficient. The last version of the software, freely available¹, includes a multi-scale scheme and the symmetric use of masks in the registration (see [Hadj-Hamou et al. \(2016\)](#)).

However the metric used between two images F and G is written as a generalized local correlation coefficient (GLCC) as follow (the bar denotes local Gaussian smoothing: $\bar{F} = \mathbf{G}_\sigma * F$):

$$\rho_{GLCC}(F, G) = \int_{\Omega} \frac{\overline{FG}}{\sqrt{\overline{F^2 G^2}}} \quad (\text{A.1})$$

This metric is invariant to linear transformations of the intensities but is not invariant to intensity offsets. This formulation corresponds to the statistical correlation for variables with 0 mean but, in the context of images, there is no reason to assume that the intensities are centered around 0. The expected LCC should use the mean of both F and G , that we denote F_m and G_m , and writes:

$$\rho_{LCC}(F, G) = \rho_{GLCC}(F - F_m, G - G_m)$$

The means can be estimated globally to account for normalization bias or for inter-modality registration. They can also be estimated locally, for example using the same Gaussian kernel i.e. $F_m = \bar{F}$. Obviously, other kernel shapes and sizes are also possible. Our implementation uses two separate scales for the Gaussian smoothing and the correlation in the metric that are defined using the parameters σ_m

¹the LCC registration software can be downloaded here:<https://team.inria.fr/epione/en/software/>.

for the mean and σ_s for the similarity. Using the consistent notation $\bar{F}^{(\sigma)} = \mathbf{G}_\sigma * F$, it writes:

$$\rho_{LCC}(F, G) = \int_{\Omega} \frac{\overline{(F - \bar{F}^{(\sigma_m)})(G - \bar{G}^{(\sigma_m)})}^{(\sigma_s)}}{\sqrt{\overline{(F - \bar{F}^{(\sigma_m)})^2}^{(\sigma_s)} \overline{(G - \bar{G}^{(\sigma_m)})^2}^{(\sigma_s)}}} \quad (\text{A.2})$$

In practice, the mean intensity signal that characterizes the intensity normalization and the slowly varying bias field is typically of lower frequency than the local features used to register images. Consequently, we generally choose $\sigma_s \leq \sigma_m$.

We should however note that the new LCC metric is badly defined over uniform areas. This problem was previously restricted to constant zero areas but the situation is more common now. An epsilon is included to avoid to divide by zero but this bad conditioning can lead to numerical instability. It illustrates how the new metric is less stable than the old one. More generally, a relatively large σ_m is recommended in order to suppress as little signal as possible.

A.3.2 Evaluation on controlled experiments

A.3.2.1 Invariance to intensity changes

The toy example shown in Figure A.3 illustrates how the linearly-invariant metric can lead to incorrect results while the affine-invariant gives a result that is in agreement with the ground-truth deformation. The experiment is simple. We take an image, shift and rescale the intensity ($x \mapsto 2x + 100$) and apply a small translation to the image. The first result is obtained with the linearly invariant metric (previous GLCC), the second with the affine invariant one (proposed LCC).

The default parameters are used. The proposed values for the new affine invariant setting are the following: the local mean is estimated using $\sigma_m = 8mm$, the similarity uses $\sigma_s = 3mm$ (unchanged relatively to the old version), the weight of the image similarity term is increased to take into account the lower values given by the new metric (due to centering) $u = 0.45$. The other parameters (smoothing, regularization, etc.) are kept unchanged.

The registration using the GLCC metric is imperfect: recognizable patterns are visible in the image difference meaning that the structure are not perfectly align. The result is better with the proposed LCC. No differences are visible in the brain (the skull is not registered because of the mask) and the only 'errors' in the SVF are located in the ventricles where the translation may not be observable (due to constant intensity area in the CSF).

A.3.2.2 Accuracy assessments on realistic brain evolutions

The choice of registration metric (LCC or GLCC) and scales is evaluated in a controlled experiment where realistic longitudinal evolutions are simulated from brain images. The simulations are performed using the simulatrophy software ([Khanal](#)

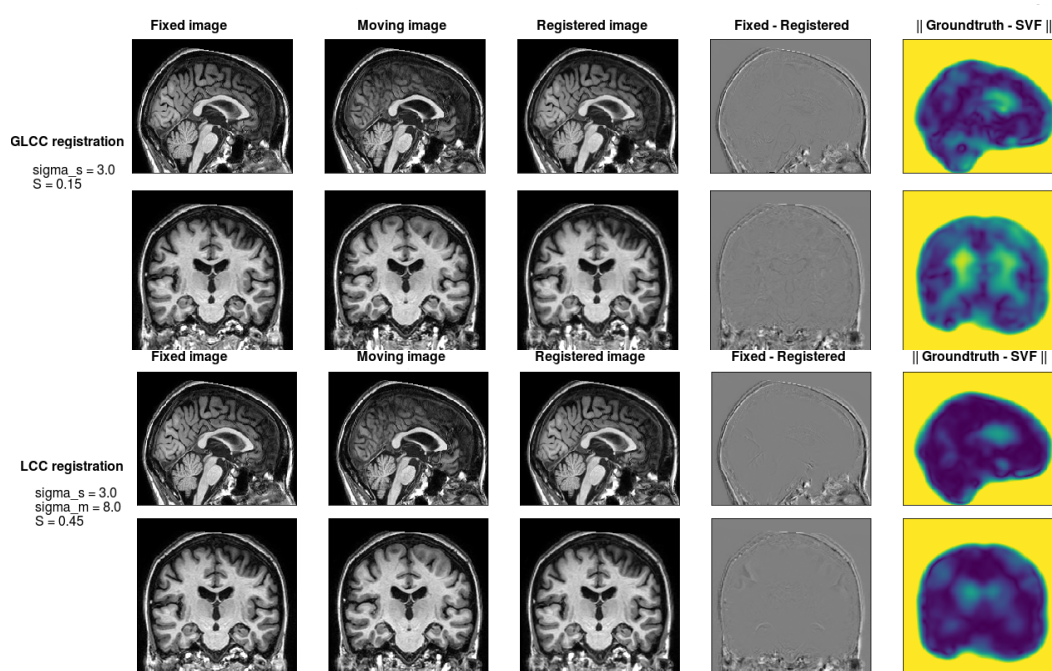


Figure A.3: Synthetic experiment using a translated and intensity-shifted images. The affine invariant metric (proposed LCC) leads to less difference between the registered and the original image and to less 'error' on the SVF relatively to the ground-truth translation relatively to the previous linear invariant metric (so-call GLCC registration).

et al., 2016). The deformations are then applied to the baseline image. No intensity changes are used (at the advantage of the previous metric). The registration is done between the baseline and the simulated image. Besides, an unique scale is used in the case of the LCC ($\sigma_m = \sigma_s$) giving the same parametrization for both algorithm but this choice is certainly sub-optimal and should be discussed in subsequent work.

The quality of the registration is assessed on a segmentation of the brain using the Dice metric. We only consider the white-matter and the cortex globally and the hippocampi. The simulations are done for 43 healthy subjects from ADNI with two atrophy settings (the ones used in Chapter 2 see table 2.2). These segmentation-based scores provide simple (but limited) assessments of the registration accuracy.

The results are presented in Figure A.4. The affine invariant metric does not lead to a improvement of the registration. Conversely, the results are worse and estimating the mean using a too small window even lead to disastrous results. Even if it is possible to get good results with the new metric for a particular set of parameters, it is not possible to use it reliably because there is no mean to fine tune these parameters for every registration.

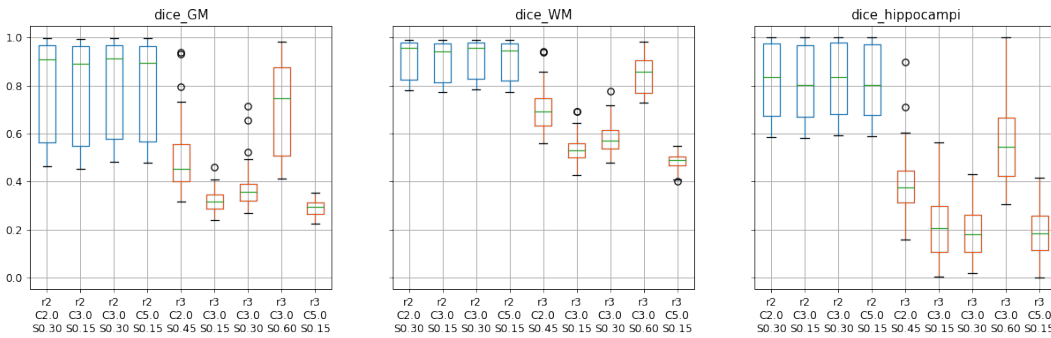


Figure A.4: Evaluation of the registration on simulated deformation for several set of parameters. r: metric type (2=GLCC, 3=LCC), C: scale, S: intensity similarity weight.

A.3.2.3 Limitations and perspectives

The presented experiments are not a validation of our contribution to the LCC-demons algorithm. They aims to justify our use of the old version of the algorithm despite this unexpected limitation of the implementation and to raise potential issues with the registration algorithm that is widely used in this thesis. The evaluation of the algorithm should be conducted in a rigorous setting and the automatic calibration of the scales parameters is an important challenge that should be addressed in future work.

Bibliography

- F. Agosta, M. Pievani, S. Sala, C. Geroldi, S. Galluzzi, G. B. Frisoni, and M. Filippi. White Matter Damage in Alzheimer Disease and Its Relationship to Gray Matter Atrophy. *Radiology*, 258(3):853–863, 2011. (Cited on page 57.)
- Alzheimer’s Association and others. 2017 Alzheimer’s disease facts and figures. *Alzheimer’s & Dementia*, 13(4):325–373, 2017. (Cited on pages 1 and 13.)
- S. Andrieu, P.-J. Ousset, N. Coley, M. Ouzid, H. Mathiex-Fortunet, and B. Velas. GuidAge study: a 5-year double blind, randomised trial of EGb 761 for the prevention of Alzheimer’s disease in elderly subjects with memory complaints. i. rationale, design and baseline data. *Current Alzheimer Research*, 5(4):406–415, 2008. (Cited on page 44.)
- S. Andrieu, S. Guyonnet, N. Coley, C. Cantet, M. Bonnefoy, S. Bordes, L. Bories, M.-N. Cufi, T. Dantoine, J.-F. Dartigues, and others. Effect of long-term omega 3 polyunsaturated fatty acid supplementation with or without multidomain intervention on cognitive function in elderly adults with memory complaints (MAPT): a randomised, placebo-controlled trial. *The Lancet Neurology*, 16(5):377–389, 2017. (Cited on pages 45 and 49.)
- L. Antelmi, N. Ayache, P. Robert, and M. Lorenzi. Sparse Multi-Channel Variational Autoencoder for the Joint Analysis of Heterogeneous Data. In K. Chaudhuri and R. Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 302–311, Long Beach, California, USA, 2019. PMLR. (Cited on page 96.)
- V. Arsigny, O. Commowick, X. Pennec, and N. Ayache. A log-euclidean framework for statistics on diffeomorphisms. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2006*, pages 924–931, 2006. (Cited on pages 4, 16 and 46.)
- J. Ashburner. SPM: a history. *Neuroimage*, 62(2):791–800, 2012. (Cited on page 62.)
- J. Ashburner and K. J. Friston. Voxel-Based Morphometry—The Methods. *NeuroImage*, 11(6):805–821, 2000. (Cited on pages 4 and 45.)
- J. Ashburner and K. J. Friston. Unified segmentation. *NeuroImage*, 26(3):839–851, 2005. (Cited on page 3.)
- J. Ashburner, C. Hutton, R. Frackowiak, I. Johnsrude, C. Price, and K. Friston. Identifying global anatomical differences: Deformation-based morphometry. *Human brain mapping*, 6(5-6):348–357, 1998. (Cited on pages 4, 45 and 62.)

- M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes. Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, 2005. (Cited on page 5.)
- M. Bilgel, J. L. Prince, D. F. Wong, S. M. Resnick, and B. M. Jernak. A multivariate nonlinear mixed effects model for longitudinal image analysis: Application to amyloid imaging. *NeuroImage*, 134:658–670, 2016. (Cited on page 14.)
- N. A. Bishop, T. Lu, and B. A. Yankner. Neural mechanisms of ageing and cognitive decline. *Nature*, 464(7288):529–535, 2010. (Cited on page 80.)
- A. Bjørnebekk, L. T. Westlye, K. B. Walhovd, and A. M. Fjell. Everyday memory: Self-perception and structural brain correlates in a healthy elderly population. *Journal of the International Neuropsychological Society*, 16(6):1115–1126, 2010. (Cited on page 80.)
- D. J. Blezek and J. V. Miller. Atlas stratification. *Medical Image Analysis*, 11(5):443–457, 2007. (Cited on page 38.)
- M. Bossa, M. Hernandez, and S. Olmos. Contributions to 3d diffeomorphic atlas estimation: application to brain images. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 667–674, 2007. (Cited on page 17.)
- A. Bowring, T. E. Nichols, and C. Maumet. Same data-different software-different results? analytic variability of group fmri results. In *OHBM 2018-24th Annual Meeting of the Organization for Human Brain Mapping*, pages 1–3, 2018. (Cited on page 58.)
- E. E. Bron, M. Smits, W. M. van der Flier, H. Vrenken, F. Barkhof, P. Scheltens, J. M. Papma, R. M. Steketee, C. Méndez Orellana, R. Meijboom, M. Pinto, J. R. Meireles, C. Garrett, A. J. Bastos-Leite, A. Abdulkadir, O. Ronneberger, N. Amoroso, R. Bellotti, D. Cárdenas-Peña, A. M. Álvarez Meza, C. V. Dolph, K. M. Iftexharuddin, S. F. Eskildsen, P. Coupé, V. S. Fonov, K. Franke, C. Gaser, C. Ledig, R. Guerrero, T. Tong, K. R. Gray, E. Moradi, J. Tohka, A. Routier, S. Durrleman, A. Sarica, G. Di Fatta, F. Sensi, A. Chincarini, G. M. Smith, Z. V. Stoyanov, L. Sørensen, M. Nielsen, S. Tangaro, P. Inglese, C. Wachinger, M. Reuter, J. C. van Swieten, W. J. Niessen, and S. Klein. Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CADDementia challenge. *NeuroImage*, 111:562–579, 2015. (Cited on page 45.)
- D. Cardebat, B. Doyon, M. Puel, P. Goulet, and Y. Joanette. Formal and semantic lexical evocation in normal subjects. Performance and dynamics of production as a function of sex, age and educational level. *Acta Neurologica Belgica*, 90(4):207–217, 1990. (Cited on page 49.)

- O. Carmichael, D. G. McLaren, D. Tommet, D. Mungas, and R. N. Jones. Coevolution of brain structures in amnesic mild cognitive impairment. *NeuroImage*, 66: 449 – 456, 2013. (Cited on page 22.)
- O. T. Carmichael, H. A. Aizenstein, S. W. Davis, J. T. Becker, P. M. Thompson, C. C. Meltzer, and Y. Liu. Atlas-based hippocampus segmentation in Alzheimer’s disease and mild cognitive impairment. *NeuroImage*, 27(4):979–990, 2005. (Cited on page 3.)
- D. M. Cash, C. Frost, L. O. Ithme, D. Ünay, M. Kandemir, J. Fripp, O. Salvado, P. Bourgeat, M. Reuter, B. Fischl, M. Lorenzi, G. B. Frisoni, X. Pennec, R. K. Pierson, J. L. Gunter, M. L. Senjem, C. R. Jack, N. Guizard, V. S. Fonov, D. L. Collins, M. Modat, M. J. Cardoso, K. K. Leung, H. Wang, S. R. Das, P. A. Yushkevich, I. B. Malone, N. C. Fox, J. M. Schott, and S. Ourselin. Assessing atrophy measurement techniques in dementia: Results from the MIRIAD atrophy challenge. *NeuroImage*, 123:149–164, 2015. (Cited on page 23.)
- G. E. Christensen, R. D. Rabbitt, and M. I. Miller. 3d brain mapping using a deformable neuroanatomy. *Physics in medicine and biology*, 39(3):609, 1994. (Cited on page 16.)
- J. H. Cole, S. J. Ritchie, M. E. Bastin, M. C. Valdés Hernández, S. Muñoz Maniega, N. Royle, J. Corley, A. Pattie, S. E. Harris, Q. Zhang, N. R. Wray, P. Redmond, R. E. Marioni, J. M. Starr, S. R. Cox, J. M. Wardlaw, D. J. Sharp, and I. J. Deary. Brain age predicts mortality. *Molecular Psychiatry*, 2017. (Cited on page 13.)
- A. Collie, P. Maruff, D. G. Darby, and M. McSTEPHEN. The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test–retest intervals. *Journal of the International Neuropsychological Society*, 9(3):419–428, 2003. (Cited on page 80.)
- O. Colliot, G. Chételat, M. Chupin, B. Desgranges, B. Magnin, H. Benali, B. Dubois, L. Garnero, F. Eustache, and S. Lehéricy. Discrimination between Alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus. *Radiology*, 248(1):194–201, 2008. (Cited on pages 4 and 45.)
- P. Coupé. *Patch-based MRI Analysis: From voxel to knowledge*. HDR Thesis, 2018. (Cited on page 4.)
- P. Coupé, J. V. Manjón, V. Fonov, J. Pruessner, M. Robles, and D. L. Collins. Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. *NeuroImage*, 54(2):940–954, 2011. (Cited on page 3.)
- R. W. Cox, G. Chen, D. R. Glen, R. C. Reynolds, and P. A. Taylor. FMRI Clustering in AFNI: False-Positive Rates Redux. *Brain Connectivity*, 7(3):152–171, 2017. (Cited on pages 62 and 78.)

- C. Cury, M. Lorenzi, D. Cash, J. M. Nicholas, A. Routier, J. Rohrer, S. Ourselin, S. Durrleman, and M. Modat. Spatio-temporal shape analysis of cross-sectional data for detection of early changes in neurodegenerative disease. In *International Workshop on Spectral and Shape Analysis in Medical Imaging*, pages 63–75. Springer, 2016. (Cited on pages 4 and 14.)
- P. A. Dacks, D. W. Shineman, and H. M. Fillit. Current evidence for the clinical use of long-chain polyunsaturated N-3 fatty acids to prevent age-related cognitive decline and Alzheimer’s disease. *The journal of nutrition, health & aging*, 17(3): 240–251, 2013. (Cited on page 44.)
- C. Davatzikos, F. Xu, Y. An, Y. Fan, and S. M. Resnick. Longitudinal progression of Alzheimer’s-like patterns of atrophy in normal older adults: the SPARE-AD index. *Brain*, 132(8):2026–2035, 2009. (Cited on page 13.)
- C. DeCarli, J. Massaro, D. Harvey, J. Hald, M. Tullberg, R. Au, A. Beiser, R. D’Agostino, and P. A. Wolf. Measures of brain morphology and infarction in the framingham heart study: establishing what is normal. *Neurobiology of aging*, 26(4):491–510, 2005. (Cited on page 13.)
- S. Deprez, F. Amant, R. Yigit, K. Porke, J. Verhoeven, J. V. d. Stock, A. Smeets, M.-R. Christiaens, A. Leemans, W. V. Hecke, J. Vandenberghe, M. Vandembulcke, and S. Sunaert. Chemotherapy-induced structural changes in cerebral white matter and its correlation with impaired cognitive functioning in breast cancer patients. *Human Brain Mapping*, 32(3):480–493, 2011. (Cited on page 80.)
- D. P. Devanand, G. Pradhaban, X. Liu, A. Khandji, S. De Santi, S. Segal, H. Rusinek, G. H. Pelton, L. S. Honig, R. Mayeux, Y. Stern, M. H. Tabert, and M. J. de Leon. Hippocampal and entorhinal atrophy in mild cognitive impairment: Prediction of Alzheimer disease. *Neurology*, 68(11):828–836, 2007. (Cited on pages 4 and 45.)
- A. Dong, J. B. Toledo, N. Honnorat, J. Doshi, E. Varol, A. Sotiras, D. Wolk, J. Q. Trojanowski, C. Davatzikos, and for the Alzheimer’s Disease Neuroimaging Initiative. Heterogeneity of neuroanatomical patterns in prodromal Alzheimer’s disease: links to cognition, progression and biomarkers. *Brain*, page aww319, 2016. (Cited on page 97.)
- M. C. Donohue, H. Jacqmin-Gadda, M. Le Goff, R. G. Thomas, R. Raman, A. C. Gamst, L. A. Beckett, C. R. Jack, M. W. Weiner, J.-F. Dartigues, and P. S. Aisen. Estimating long-term multivariate progression from short-term data. *Alzheimer’s & Dementia*, 10(5):S400–S410, 2014. (Cited on page 14.)
- K. L. Double, G. M. Halliday, J. J. Krill, J. A. Harasty, K. Cullen, W. S. Brooks, H. Creasey, and G. A. Broe. Topography of brain atrophy during normal aging and alzheimer’s disease. *Neurobiology of Aging*, 17(4):513 – 521, 1996. (Cited on page 31.)

- S. Durrleman. *Geometrical Approaches in Statistical Learning for the Construction of Digital Models of the Human Brain*. HDR Thesis, 2018. (Cited on page 96.)
- A. Eklund, T. E. Nichols, and H. Knutsson. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, 113(28):7900–7905, 2016. (Cited on page 62.)
- C. M. Fiford, G. R. Ridgway, D. M. Cash, M. Modat, J. Nicholas, E. N. Manning, I. B. Malone, G. J. Biessels, S. Ourselin, O. T. Carmichael, M. J. Cardoso, and J. Barnes. Patterns of progressive atrophy vary with age in Alzheimer’s disease patients. *Neurobiology of Aging*, 63:22–32, 2018. (Cited on page 97.)
- B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A. M. Dale. Whole Brain Segmentation. *Neuron*, 33(3):341–355, 2002. (Cited on pages 3 and 22.)
- J. Fishbaugh, M. Prastawa, G. Gerig, and S. Durrleman. Geodesic image regression with a sparse parameterization of diffeomorphisms. In *International Conference on Geometric Science of Information*, pages 95–102. Springer, 2013. (Cited on page 96.)
- A. M. Fjell, K. B. Walhovd, C. Fennema-Notestine, L. K. McEvoy, D. J. Hagler, D. Holland, K. Blennow, J. B. Brewer, A. M. Dale, and t. A. D. N. Initiative. Brain Atrophy in Healthy Aging Is Related to CSF Levels of Ab1-42. *Cerebral Cortex*, 20(9):2069–2079, 2010. (Cited on page 22.)
- M. F. Folstein, S. E. Folstein, and P. R. McHugh. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3):189–198, 1975. (Cited on page 49.)
- H. M. Fonteijn, M. Modat, M. J. Clarkson, J. Barnes, M. Lehmann, N. Z. Hobbs, R. I. Scahill, S. J. Tabrizi, S. Ourselin, N. C. Fox, and D. C. Alexander. An event-based model for disease progression and its application in familial Alzheimer’s disease and Huntington’s disease. *NeuroImage*, 60(3):1880–1889, 2012. (Cited on page 14.)
- K. Franke and C. Gaser. Longitudinal Changes in Individual *BrainAGE* in Healthy Aging, Mild Cognitive Impairment, and Alzheimer’s Disease. *GeroPsych*, 25(4): 235–245, 2012. (Cited on page 14.)
- K. Franke, G. Ziegler, S. Klöppel, C. Gaser, A. D. N. Initiative, and others. Estimating the age of healthy subjects from T 1-weighted MRI scans using kernel methods: Exploring the influence of various parameters. *Neuroimage*, 50(3):883–892, 2010. (Cited on page 13.)

- K. J. Friston, editor. *Statistical parametric mapping: the analysis of functional brain images*. Elsevier/Academic Press, Amsterdam ; Boston, 1st ed edition, 2007. (Cited on page 68.)
- Y. Fujikoshi. Likelihood Ratio Tests in Multivariate Linear Model. In V. N. Katsikis, editor, *Applied Linear Algebra in Action*. InTech, 2016. (Cited on page 65.)
- D. Galasko, D. A. Bennett, M. Sano, D. Marson, J. Kaye, S. D. Edland, and Alzheimer’s Disease Cooperative Study. ADCS Prevention Instrument Project: assessment of instrumental activities of daily living for community-dwelling elderly individuals in dementia prevention clinical trials. *Alzheimer Disease and Associated Disorders*, 20(4 Suppl 3):S152–169, 2006. (Cited on page 49.)
- C. D. Good, I. S. Johnsruide, J. Ashburner, R. N. Henson, K. J. Friston, and R. S. Frackowiak. A Voxel-Based Morphometric Study of Ageing in 465 Normal Adult Human Brains. *NeuroImage*, 14(1):21–36, 2001. (Cited on pages 12 and 80.)
- B. Gris, S. Durrleman, and A. Trouvé. A sub-Riemannian modular approach for diffeomorphic deformations. In *Geometric Science of Information*, pages 39–47. Springer, 2015. (Cited on page 96.)
- E. Grober, H. Buschke, H. Crystal, S. Bang, and R. Dresner. Screening for dementia by memory testing. *Neurology*, 38(6):900–903, 1988. (Cited on page 49.)
- A. Guimond, J. Meunier, and J.-P. Thirion. Average brain models: A convergence study. *Computer vision and image understanding*, 77(2):192–210, 2000. (Cited on page 20.)
- J. M. Guralnik, L. Ferrucci, C. F. Pieper, S. G. Leveille, K. S. Markides, G. V. Ostir, S. Studenski, L. F. Berkman, and R. B. Wallace. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, 55(4):M221–231, 2000. (Cited on page 49.)
- M. Hadj-Hamou. *Beyond volumetry in longitudinal deformation-based morphometry: application to sexual dimorphism during adolescence*. PhD thesis, Université Côte d’Azur, 2016. (Cited on pages 23 and 26.)
- M. Hadj-Hamou, M. Lorenzi, N. Ayache, and X. Pennec. Longitudinal analysis of image time series with diffeomorphic deformations: a computational framework based on stationary velocity fields. *Frontiers in neuroscience*, 10, 2016. (Cited on pages 5, 19, 38, 48, 63, 81 and 103.)
- K. Hett, V.-T. Ta, J. V. Manjón, P. Coupé, A. D. N. Initiative, and others. Adaptive fusion of texture-based grading: application to Alzheimer’s disease detection. In *International Workshop on Patch-based Techniques in Medical Imaging*, pages 82–89. Springer, 2017. (Cited on page 96.)

- W. Huizinga, D. Poot, M. Vernooij, G. Roshchupkin, E. Bron, M. Ikram, D. Rueckert, W. Niessen, and S. Klein. A spatio-temporal reference model of the aging brain. *NeuroImage*, 169:11–22, 2018. (Cited on page 13.)
- C. Hutton, B. Draganski, J. Ashburner, and N. Weiskopf. A comparison between voxel-based cortical thickness and voxel-based morphometry in normal aging. *NeuroImage*, 48(2):371–380, 2009. (Cited on pages 4, 13 and 45.)
- M. D. Ikonomic, W. E. Klunk, E. E. Abrahamson, C. A. Mathis, J. C. Price, N. D. Tsopelas, B. J. Lopresti, S. Ziolk, W. Bi, W. R. Paljug, M. L. Debnath, C. E. Hope, B. A. Isanski, R. L. Hamilton, and S. T. DeKosky. Post-mortem correlates of in vivo PiB-PET amyloid imaging in a typical case of Alzheimer’s disease. *Brain*, 131(6):1630–1645, 2008. (Cited on page 3.)
- C. R. Jack, T. M. Therneau, H. J. Wiste, S. D. Weigand, D. S. Knopman, V. J. Lowe, M. M. Mielke, P. Vemuri, R. O. Roberts, M. M. Machulda, M. L. Senjem, J. L. Gunter, W. A. Rocca, and R. C. Petersen. Transition rates between amyloid and neurodegeneration biomarker states and to dementia: a population-based, longitudinal cohort study. *The Lancet Neurology*, 15(1):56–64, 2016. (Cited on page 45.)
- M. Jenkinson and S. Smith. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5(2):143–156, 2001. (Cited on page 102.)
- N. M. Kemppainen, S. Aalto, I. A. Wilson, K. Nagren, S. Helin, A. Bruck, V. Oikonen, M. Kailajarvi, M. Scheinin, M. Viitanen, R. Parkkola, and J. O. Rinne. Voxel-based analysis of PET amyloid ligand [11c]PIB uptake in Alzheimer disease. *Neurology*, 67(9):1575–1580, 2006. (Cited on page 4.)
- B. Khanal, M. Lorenzi, N. Ayache, and X. Pennec. A biophysical model of brain deformation to simulate and analyze longitudinal MRIs of patients with Alzheimer’s disease. *NeuroImage*, pages 35–52, 2016. (Cited on pages 21, 22, 96 and 104.)
- M. Kivipelto, F. Mangialasche, and T. Ngandu. Lifestyle interventions to prevent cognitive impairment, dementia and Alzheimer disease. *Nature Reviews Neurology*, 14(11):653–666, 2018. (Cited on pages 44 and 58.)
- S. Klöppel, A. Abdulkadir, C. R. Jack, N. Koutsouleris, J. Mourão-Miranda, and P. Vemuri. Diagnostic neuroimaging across diseases. *Neuroimage*, 61(2):457–463, 2012. (Cited on page 13.)
- I. Koval, J.-B. Schiratti, A. Routier, M. Bacci, O. Colliot, S. Allasonnière, S. Durrleman, and A. D. N. Initiative. Statistical learning of spatiotemporal patterns from longitudinal manifold-valued networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 451–459. Springer, 2017. (Cited on pages 4, 14 and 97.)

- J. Krebs, H. Delingette, B. Mailhe, N. Ayache, and T. Mansi. Learning a Probabilistic Model for Diffeomorphic Registration. *IEEE Transactions on Medical Imaging*, pages 1–1, 2019. (Cited on page 96.)
- R. Levy, Working Party of the International Psychogeriatric Association in collaboration with the World Health Organization, et al. Aging-Associated Cognitive Decline. *International Psychogeriatrics*, 6(1):63–68, 1994. (Cited on page 80.)
- J. A. Lieberman. Antipsychotic Drug Effects on Brain Morphology in First-Episode Psychosis. *Archives of General Psychiatry*, 62(4):361, 2005. (Cited on page 45.)
- Y. Liu, K. Wang, C. Yu, Y. He, Y. Zhou, M. Liang, L. Wang, and T. Jiang. Regional homogeneity, functional connectivity and imaging markers of Alzheimer’s disease: A review of resting-state fMRI studies. *Neuropsychologia*, 46(6):1648–1656, 2008. (Cited on page 4.)
- X. Long, W. Liao, C. Jiang, D. Liang, B. Qiu, and L. Zhang. Healthy Aging. *Academic Radiology*, 19(7):785–793, 2012. (Cited on pages 12 and 80.)
- M. Lorenzi and X. Pennec. Geodesics, parallel transport & one-parameter subgroups for diffeomorphic image registration. *International journal of computer vision*, 105(2):111–127, 2013. (Cited on page 18.)
- M. Lorenzi and X. Pennec. Efficient Parallel Transport of Deformations in Time Series of Images: From Schild’s to Pole Ladder. *Journal of Mathematical Imaging and Vision*, 50(1-2):5–17, 2014. (Cited on pages 4 and 46.)
- M. Lorenzi, N. Ayache, G. B. Frisoni, X. Pennec, A. D. N. I. (ADNI, and others. LCC-Demons: a robust and accurate symmetric diffeomorphic registration algorithm. *NeuroImage*, 81:470–483, 2013. (Cited on pages 5, 16, 20, 48, 68 and 103.)
- M. Lorenzi, X. Pennec, G. B. Frisoni, N. Ayache, A. D. N. Initiative, and others. Disentangling normal aging from Alzheimer’s disease in structural magnetic resonance images. *Neurobiology of aging*, 36:S42–S52, 2015. (Cited on pages 14, 26 and 28.)
- M. Lorenzi, M. Filippone, G. B. Frisoni, D. C. Alexander, and S. Ourselin. Probabilistic disease progression modeling to characterize diagnostic uncertainty: Application to staging and prediction in Alzheimer’s disease. *NeuroImage*, 2017. (Cited on pages 14 and 45.)
- R. Luerding, T. Weigand, U. Bogdahn, and T. Schmidt-Wilcke. Working memory performance is correlated with local brain morphology in the medial frontal and anterior cingulate cortex in fibromyalgia patients: structural correlates of pain–cognition interaction. *Brain*, 131(12):3222–3231, 2008. (Cited on page 80.)
- J. K. Mai, M. Majtanik, and G. Paxinos. *Atlas of the human brain*. Academic Press, 2015. (Cited on page 3.)

- J. Manjon, J. Carbonellcaballero, J. Lull, G. Garciamarti, L. Martibonmati, and M. Robles. MRI denoising using Non-Local Means. *Medical Image Analysis*, 12(4):514–523, 2008. (Cited on page 100.)
- J. V. Manjón and P. Coupé. volBrain: An Online MRI Brain Volumetry System. *Frontiers in Neuroinformatics*, 10, 2016. (Cited on page 3.)
- A. Marchewka, F. Kherif, G. Krueger, A. Grabowska, R. Frackowiak, B. Draganski, and T. A. D. N. Initiative. Influence of magnetic field strength and image registration strategy on voxel-based morphometry in a study of Alzheimer’s disease. *Human Brain Mapping*, 35(5):1865–1874, 2014. (Cited on page 38.)
- E. Maris and R. Oostenveld. Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1):177–190, 2007. (Cited on page 63.)
- C. R. McDonald, L. K. McEvoy, L. Gharapetian, C. Fennema-Notestine, D. J. Hagler, D. Holland, A. Koyama, J. B. Brewer, A. M. Dale, and For the Alzheimer’s Disease Neuroimaging Initiative. Regional rates of neocortical atrophy from normal aging to early Alzheimer disease. *Neurology*, 73(6):457–465, 2009. (Cited on pages 45 and 57.)
- K. Mcleod, C. Seiler, M. Serresant, and X. Pennec. Spatio-temporal dimension reduction of cardiac motion for group-wise analysis and statistical testing. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 501–508. Springer, 2013. (Cited on page 96.)
- D. McNair and R. Kahn. Self-assessment of cognitive deficits. *Assessment in geriatric psychopharmacology*, 137:143, 1983. (Cited on page 49.)
- Z. A. Medvedev. An attempt at a rational classification of theories of ageing. *Biological Reviews*, 65(3):375–398, 1990. (Cited on page 12.)
- M. Mennes, M. Jenkinson, R. Valabregue, J. K. Buitelaar, C. Beckmann, and S. Smith. Optimizing full-brain coverage in human brain MRI through population distributions of brain size. *NeuroImage*, 98:513–520, 2014. (Cited on page 100.)
- P. Muralidharan, J. Fishbaugh, E. Y. Kim, H. J. Johnson, J. S. Paulsen, G. Gerig, and P. T. Fletcher. Bayesian covariate selection in mixed-effects models for longitudinal shape analysis. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 656–659, Prague, Czech Republic, 2016. IEEE. (Cited on page 39.)
- H. K. Na, D. R. Kang, S. Kim, S. W. Seo, K. M. Heilman, Y. Noh, and D. L. Na. Malignant progression in parietal-dominant atrophy subtype of Alzheimer’s disease occurs independent of onset age. *Neurobiology of Aging*, 47:149–156, 2016. (Cited on page 97.)
- National Institute of Aging. NIA information portal on Alzheimer’s disease, 2019. (Cited on page 1.)

- T. Ngandu, J. Lehtisalo, A. Solomon, E. Levälähti, S. Ahtiluoto, R. Antikainen, L. Bäckman, T. Hänninen, A. Jula, T. Laatikainen, J. Lindström, F. Mangialasche, T. Pajananen, S. Pajala, M. Peltonen, R. Rauramaa, A. Stigsdotter-Neely, T. Strandberg, J. Tuomilehto, H. Soininen, and M. Kivipelto. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. *The Lancet*, 385(9984):2255–2263, 2015. (Cited on page 44.)
- T. E. Nichols and A. P. Holmes. Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, 15(1):1–25, 2002. (Cited on pages 48, 62, 65, 66 and 77.)
- T. Ohnishi, H. Matsuda, T. Tabira, T. Asada, and M. Uno. Changes in brain morphology in Alzheimer disease and normal aging: is Alzheimer disease an exaggerated aging process? *American Journal of Neuroradiology*, 22(9):1680–1685, 2001. (Cited on page 13.)
- D. C. Park and P. Reuter-Lorenz. The Adaptive Brain: Aging and Neurocognitive Scaffolding. *Annual Review of Psychology*, 60(1):173–196, 2009. (Cited on pages 12 and 80.)
- F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011. (Cited on pages 31 and 82.)
- J. B. Pereira, N. Ibarretxe-Bilbao, M.-J. Marti, Y. Compta, C. Junqué, N. Bargallo, and E. Tolosa. Assessment of cortical degeneration in patients with Parkinson’s disease by voxel-based morphometry, cortical folding, and cortical thickness. *Human Brain Mapping*, 33(11):2521–2534, 2012. (Cited on page 3.)
- L. Pini, M. Pievani, M. Bocchetta, D. Altomare, P. Bosco, E. Cavedo, S. Galluzzi, M. Marizzoni, and G. B. Frisoni. Brain atrophy in Alzheimer’s Disease and aging. *Ageing Research Reviews*, 30:25–48, 2016. (Cited on pages 2 and 22.)
- K. Poulakis, J. B. Pereira, P. Mecocci, B. Vellas, M. Tsolaki, I. Kłoszewska, H. Soininen, S. Lovestone, A. Simmons, L.-O. Wahlund, and E. Westman. Heterogeneous patterns of brain atrophy in Alzheimer’s disease. *Neurobiology of Aging*, 65:98–108, 2018. (Cited on page 97.)
- A. Qiu, L. Younes, L. Wang, J. T. Ratnanather, S. K. Gillepsie, G. Kaplan, J. Csernansky, and M. I. Miller. Combining anatomical manifold information via diffeomorphic metric mappings for studying cortical thinning of the cingulate gyrus in schizophrenia. *NeuroImage*, 37(3):821–833, 2007. (Cited on pages 4 and 45.)

- N. Raz and K. M. Rodrigue. Differential aging of the brain: Patterns, cognitive correlates and modifiers. *Neuroscience & Biobehavioral Reviews*, 30(6):730–748, 2006. (Cited on pages 80 and 91.)
- R. M. Reitan. Validity of the Trail Making Test as an Indicator of Organic Brain Damage. *Perceptual and Motor Skills*, 8(3):271–276, 1958. (Cited on page 49.)
- M. Reuter, N. J. Schmansky, H. D. Rosas, and B. Fischl. Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, 61(4):1402–1418, 2012. (Cited on pages 38 and 46.)
- S. L. Risacher, L. Shen, J. D. West, S. Kim, B. C. McDonald, L. A. Beckett, D. J. Harvey, C. R. Jack, M. W. Weiner, and A. J. Saykin. Longitudinal MRI atrophy biomarkers: Relationship to conversion in the ADNI cohort. *Neurobiology of Aging*, 31(8):1401–1418, 2010. (Cited on pages 4 and 45.)
- K. M. Rodrigue and N. Raz. Shrinkage of the entorhinal cortex over five years predicts memory performance in healthy adults. *Journal of Neuroscience*, 24(4):956–963, 2004. (Cited on page 12.)
- M.-M. Rohé, M. Sermesant, and X. Pennec. Barycentric Subspace Analysis: a new Symmetric Group-wise Paradigm for Cardiac Motion Tracking. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 300–307. Springer, 2016. (Cited on page 38.)
- A. C. Rosen, M. W. Prull, J. D. E. Gabrieli, T. Stoub, R. O’Hara, L. Friedman, J. A. Yesavage, and L. deToledo Morrell. Differential Associations Between Entorhinal and Hippocampal Volumes and Memory Performance in Older Adults. *Behavioral Neuroscience*, 117(6):1150–1160, 2003. (Cited on pages 12 and 80.)
- D. S. Roy, A. Arons, T. I. Mitchell, M. Pignatelli, T. J. Ryan, and S. Tonegawa. Memory retrieval by activating engram cells in mouse models of early Alzheimer’s disease. *Nature*, 531(7595):508–512, 2016. (Cited on page 3.)
- M. Sabuncu, S. Balci, M. Shenton, and P. Golland. Image-Driven Population Analysis Through Mixture Modeling. *IEEE Transactions on Medical Imaging*, 28(9):1473–1487, 2009. (Cited on page 38.)
- M. R. Sabuncu, R. S. Desikan, J. Sepulcre, B. T. T. Yeo, H. Liu, N. J. Schmansky, M. Reuter, M. W. Weiner, R. L. Buckner, R. A. Sperling, and B. Fischl. The Dynamics of Cortical and Hippocampal Atrophy in Alzheimer Disease. *ARCH NEUROL*, 68(8):9, 2011. (Cited on page 45.)
- V. D. Santos, P. A. Thomann, T. Wüstenberg, U. Seidl, M. Essig, and J. Schröder. Morphological Cerebral Correlates of CERAD Test Performance in Mild Cognitive Impairment and Alzheimer’s Disease. *Journal of Alzheimer’s Disease*, 23(3):411–420, 2011. (Cited on page 80.)

- A. Saykin, H. Wishart, L. Rabin, R. Santulli, L. Flashman, J. West, T. McHugh, and A. Mamourian. Older adults with cognitive complaints show brain atrophy similar to that of amnesic MCI. *Neurology*, 67(5):834–842, 2006. (Cited on page 45.)
- C. Scarpazza, S. Tognin, S. Frisciata, G. Sartori, and A. Mechelli. False positive rates in Voxel-based Morphometry studies of the human brain: Should we be worried? *Neuroscience & Biobehavioral Reviews*, 52:49–55, 2015. (Cited on page 78.)
- J.-B. Schiratti, S. Allasonnière, O. Colliot, and S. Durrleman. A Bayesian mixed-effects model to learn trajectories of changes from repeated manifold-valued observations. *The Journal of Machine Learning Research*, 18(1):4840–4872, 2017. (Cited on page 14.)
- T. Schmidt-Wilcke, S. Poljansky, S. Hierlmeier, J. Hausner, and B. Ibach. Memory performance correlates with gray matter density in the ento-/perirhinal cortex and posterior hippocampus in patients with mild cognitive impairment and healthy controls — A voxel based morphometry study. *NeuroImage*, 47(4):1914–1920, 2009. (Cited on page 45.)
- D. Schmitter, A. Roche, B. Maréchal, D. Ribes, A. Abdulkadir, M. Bach-Cuadra, A. Daducci, C. Granziera, S. Klöppel, P. Maeder, R. Meuli, and G. Krueger. An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer’s disease. *NeuroImage: Clinical*, 7:7–17, 2015. (Cited on page 13.)
- A. Schwartzman, R. F. Dougherty, and J. E. Taylor. Cross-subject comparison of principal diffusion direction maps. *Magnetic Resonance in Medicine*, 53(6):1423–1431, 2005. (Cited on page 62.)
- M. Silver, G. Montana, and T. E. Nichols. False positives in neuroimaging genetics using voxel-based morphometry data. *NeuroImage*, 54(2):992–1000, 2011. (Cited on page 78.)
- N. Singh, J. Hinkle, S. Joshi, and P. T. Fletcher. A hierarchical geodesic model for diffeomorphic longitudinal shape analysis. In *International Conference on Information Processing in Medical Imaging*, pages 560–571. Springer, 2013. (Cited on page 37.)
- R. Sivera, N. Capet, V. Manera, R. Fabre, M. Lorenzi, H. Delingette, X. Pennec, N. Ayache, and P. Robert. Voxel based assessments of treatment effects on brain evolution in the MAPT cohort. 2019a. (Cited on pages 67, 78, 80, 88 and 91.)
- R. Sivera, H. Delingette, M. Lorenzi, X. Pennec, and N. Ayache. A model of brain morphological changes related to aging and Alzheimer’s disease from cross-sectional assessments. *NeuroImage*, 2019b. (Cited on page 45.)

- J. D. Sluimer, W. M. van der Flier, G. B. Karas, N. C. Fox, P. Scheltens, F. Barkhof, and H. Vrenken. Whole-brain atrophy rate and cognitive decline: longitudinal MR study of memory clinic patients. *Radiology*, 248(2):590–598, 2008. (Cited on page 45.)
- S. M. Smith. Fast robust automated brain extraction. *Human Brain Mapping*, 17(3):143–155, 2002. (Cited on page 99.)
- S. M. Smith, M. Jenkinson, H. Johansen-Berg, D. Rueckert, T. E. Nichols, C. E. Mackay, K. E. Watkins, O. Ciccarelli, M. Z. Cader, P. M. Matthews, and T. E. Behrens. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*, 31(4):1487–1505, 2006. (Cited on page 62.)
- E. R. Sowell, B. S. Peterson, P. M. Thompson, S. E. Welcome, A. L. Henkenius, and A. W. Toga. Mapping cortical change across the human life span. *Nature neuroscience*, 6(3):309–315, 2003. (Cited on page 13.)
- J. Talairach. Co-planar stereotaxic atlas of the human brain-3-dimensional proportional system. *An approach to cerebral imaging*, 1988. (Cited on page 3.)
- S. J. Teipel, C. Born, M. Ewers, A. L. Bokde, M. F. Reiser, H.-J. Möller, and H. Hampel. Multivariate deformation-based analysis of brain atrophy to predict Alzheimer’s disease in mild cognitive impairment. *NeuroImage*, 38(1):13–24, 2007. (Cited on page 62.)
- N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15(1):273–289, 2002. (Cited on page 31.)
- F. van der Lijn, T. den Heijer, M. M. Breteler, and W. J. Niessen. Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts. *NeuroImage*, 43(4):708–720, 2008. (Cited on page 3.)
- M. van Oijen, F. Jan de Jong, A. Hofman, P. J. Koudstaal, and M. Breteler. Subjective memory complaints, education, and risk of Alzheimer’s disease. *Alzheimer’s & Dementia*, 3(2):92 – 97, 2007. (Cited on page 44.)
- E. F. van Velsen, M. W. Vernooij, H. A. Vrooman, A. van der Lugt, M. M. Breteler, A. Hofman, W. J. Niessen, and M. A. Ikram. Brain cortical thickness in the general elderly population: the Rotterdam Scan Study. *Neuroscience letters*, 550:189–194, 2013. (Cited on page 13.)
- B. Vellas, I. Carrie, S. Gillette-Guyonnet, J. Touchon, T. Dantoine, M. N. Cuffi, S. Bordes, Y. Gasnier, P. Robert, L. Bories, O. Rouaud, K. Sudres, M. Bonnefoy, A. Pesce, C. Dufouil, S. Lehericy, J. F. Mangin, P. Payoux, D. Adel, P. Legrand, D. Catheline, M. Zaim, L. Molinier, N. Costa, J. Delrieu, T. Voisin,

- C. Faisant, F. Lala, F. Nourhashémi, Y. Rolland, G. A. V. Kan, C. Dupuy, P. Cestac, S. Belleville, S. Willis, M. Cesari, M. W. Weiner, P. J. Ousset, and S. Andrieu. MAPT study: A multidomain approach for preventing Alzheimer's disease: design and baseline data. *The journal of prevention of Alzheimer's disease*, page 23, 2015. (Cited on pages 45 and 63.)
- L. Wang, F. Beg, T. Ratnanather, C. Ceritoglu, L. Younes, J. C. Morris, J. G. Csernansky, and M. I. Miller. Large Deformation Diffeomorphism and Momentum Based Hippocampal Shape Discrimination in Dementia of the Alzheimer type. *IEEE Transactions on Medical Imaging*, 26(4):462–470, 2007. (Cited on page 16.)
- C. J. Wargo, J. Moore, and J. C. Gore. A comparison and evaluation of reduced-FOV methods for multi-slice 7t human imaging. *Magnetic Resonance Imaging*, 31(8):1349–1359, 2013. (Cited on page 99.)
- D. Wechsler. *Manual for the Wechsler adult intelligence scale - revised*. Psychological Corp., 1981. (Cited on page 49.)
- J. Wegelin. A survey of Partial Least Squares (PLS) methods, with emphasis on the two-block case. Technical Report 371, University of Washington, 2000. (Cited on page 82.)
- J. L. Whitwell, S. A. Przybelski, S. D. Weigand, D. S. Knopman, B. F. Boeve, R. C. Petersen, and C. R. Jack Jr. 3d maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease. *Brain*, 130(7):1777–1786, 2007. (Cited on page 46.)
- A. M. Winkler, G. R. Ridgway, M. A. Webster, S. M. Smith, and T. E. Nichols. Permutation inference for the general linear model. *NeuroImage*, 92:381–397, 2014. (Cited on page 62.)
- C.-W. Woo, A. Krishnan, and T. D. Wager. Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *NeuroImage*, 91:412–419, 2014. (Cited on page 72.)
- J. A. Yesavage, T. L. Brink, T. L. Rose, O. Lum, V. Huang, M. Adey, and V. O. Leirer. Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of Psychiatric Research*, 17(1):37–49, 1982. (Cited on page 49.)
- J. Zhang, Q. Peng, Q. Li, N. Jahanshad, Z. Hou, M. Jiang, N. Masuda, D. R. Langbehn, M. I. Miller, S. Mori, C. A. Ross, and W. Duan. Longitudinal characterization of brain atrophy of a Huntington's disease mouse model by automated morphological analyses of magnetic resonance images. *NeuroImage*, 49(3):2340–2351, 2010. (Cited on page 4.)

-
- M. Zhang and P. T. Fletcher. Finite-dimensional Lie algebras for fast diffeomorphic image registration. In *International Conference on Information Processing in Medical Imaging*, pages 249–260. Springer, 2015. (Cited on page 96.)
- M. Zhang and P. Thomas Fletcher. Bayesian principal geodesic analysis for estimating intrinsic diffeomorphic image variability. *Medical Image Analysis*, 25(1): 37–44, 2015. (Cited on page 96.)
- Y. Zhuge, J. K. Udupa, J. Liu, and P. K. Saha. Image background inhomogeneity correction in MRI via intensity standardization. *Computerized Medical Imaging and Graphics*, 33(1):7–16, 2009. (Cited on page 100.)