# A model of brain morphological changes caused by aging and Alzheimer's disease for cross-sectional assessments

Raphaël Sivera, Hervé Delingette, Marco Lorenzi, Xavier Pennec, Nicholas Ayache

# A model of brain morphological changes caused by aging and Alzheimer's disease for cross-sectional assessments

Raphaël Sivera*[1], Hervé Delingette[1], Marco Lorenzi[1], Xavier Pennec[1], and Nicholas Ayache[1]

for the Alzheimer's Disease Neuroimaging Initiative†

[1]Epione Research Project, Université Côte d'Azur, Inria Sophia Antipolis, France.

December 7, 2018

## Abstract

In this study we propose a deformation-based framework to jointly model the influence of aging and Alzheimer's disease (AD) on the brain morphological evolution. Our approach combines a spatio-temporal description of both processes into a generative model. A reference morphology is deformed along specific trajectories to match subject specific morphologies. It is used to define two imaging progression markers: 1) a *morphological age* and 2) a *disease score*. These markers can be computed locally in any brain region.

The approach is evaluated on brain structural magnetic resonnance images (MRI) from the ADNI database. The generative model is first estimated on a control population, then, for each subject, the markers are computed for each acquisition. The longitudinal evolution of these markers is then studied in relation with the clinical diagnosis of the subjects and used to generate possible morphological evolution.

In the model, the morphological changes associated with normal aging are mainly found around the ventricles, while the Alzheimer's disease specific changes are more located in the temporal lobe and the hippocampal area. The statistical analysis of these markers highlights differences between clinical conditions even though the inter-subject variability is quiet high. In this context, the model can be used to generate plausible morphological trajectories associated with the disease.

Our method gives two interpretable scalar imaging biomarkers assessing the effects of aging and disease on brain morphology at the individual and population level. Our proposed progressions

*Corresponding author at: Epione Research Project, INRIA Sophia-Antipolis, 2004, route des Lucioles, 06902 Sophia-Antipolis, France, raphael.sivera@inria.fr

†Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: `http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.p`

markers confirm an acceleration of apparent aging for Alzheimer's subjects and can help discriminate clinical conditions even in prodromal stages. More generally, the joint modeling of normal and pathological evolutions shows promising results to describe age-related brain diseases over long time scales.

Keywords: normal aging, Alzheimer's disease, registration based framework, spatio-temporal model, cross-sectional progression markers, morphological age, disease score, imaging biomarkers.

# 1 Introduction

Age-related diseases are a growing public health concern with the aging of the population. 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 [1], leading to the definition of surrogate age variables based on the quantification of the biological changes.

## 1.1 Modeling brain morphological aging

In this paper 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 [2] and it was also shown to be directly correlated with structural changes such as brain atrophy and decrease of white matter integrity [3, 4].

The normal brain morphology has been studied from the development stage to the most advanced ages in image-based studies. 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 led to an initial understanding of the normal brain shape evolution across life span [5]. These descriptions have been used to regress models characterizing brain aging in order to highlight differences between subjects or brain areas [6].

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* [7] 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 [8]. More generally, these surrogate brain age estimates have been associated with an increase of risk factors for several age-related disorders such as cardio-vascular diseases [9, 10]. In a longitudinal setting, a brain age measurement could be used to compare the evolution of several clinical conditions (see Figure 1).

## 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, the most common form of dementia [11]. These diseases come with their own specific apparent brain morphological changes [12]. Computer-aided diagnosis techniques using neuroimaging features have shown

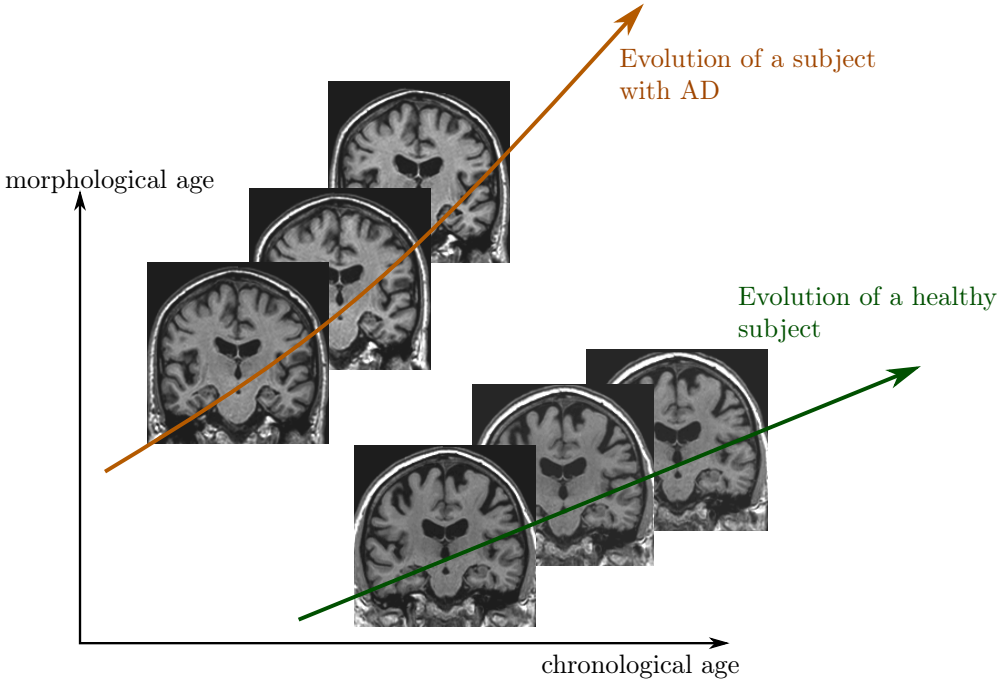



Figure 1: Schematic representation of two evolutions relatively to an hypothetic *morphological age* reflecting the structural status of the brain relatively to the aging process.

promising results to classify and to predict clinical evolutions [13, 14]. They have been used to lead to precise and robust diagnostic measures, as compared to the use of clinical criteria alone [15].

It is also interesting to describe how the pathology impacts the brain morphology over time. 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 are either event-based, and rely on a discretisation of the evolution [16], or inspired by random effect modeling of continuous trajectories (temporal evolutions). They use a population model to describe the global evolution and put in relation the individual trajectories that could only be observed a limited number of times: only once or, at best, a few times in a relatively short time-windows.

They have been used to model the progression of biomarkers [17] but also directly brain shape [18] or spatio-temporal patterns in brain images [19, 20]. This random-effect modeling also allow the use of a complementary set of variables in addition of the morphology (for example markers from amyloid imaging [21]) to improve the staging and to better align the subjects time-wise before analyzing individual neuroimaging data. These models produce good representations of the disease progression and can combine a variety of available biomarkers for patient monitoring [22].

### 1.3 Toward a joint model of brain aging and disease progression

The morphological aging of the brain and the 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 [23] or to put aside the aging part in order to focus on the disease specific changes. Lorenzi et al. [24] proposed to model the normal aging evolution to disentangle the relative contributions of aging from a reminder, not explicitly modeled, that is more related to 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. [24] 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.

In section 2, the generative model used to represent the brain morphology is introduced while section 3 describes how the model parameters are estimated and how an inverse problem is solved to compute the morphological age and the disease score of a subject. Experimental results are then presented in section 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 follow the evolution of elderly patients. Finally, the interest and limits of this approach and its perspectives are discuss in section 5.

## 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 two 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 2.1, we expose the main ideas behind the definition of our reference parametric space. Then in section 2.2, we explain how the morphological evolutions are modeled in our framework. Finally in section 2.3, we show how we use the deformation framework to model individual morphologies relatively to this reference.

### 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 $\phi_k$ and an intensity noise $\epsilon_k$ is added in the subject space accounting for local intensity variability. Therefore the images are modeled as follows:

$$I_k = T_0 \circ \phi_k + \epsilon_k$$

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 $\Phi$ parametrized by two variables that measure the progress of each process: the *morphological age* $\lambda_{MA}$ and the *disease score* $\lambda_{DS}$. The two variables $\lambda_{MA}$ and $\lambda_{DS}$ 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 $\lambda_{MA}$ years of normal aging and $\lambda_{DS}$ 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 diseased patient of age $t$ with a disease onset of $t'$ years. If one is able to associate an age $\lambda_0$ to the image $T_0$, and assumes that this is the image of a healthy subject, then it is natural to enforce $\Phi(\lambda_0, 0) = \mathrm{Id}$.

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

$$\mathcal{T} = \{T_0 \circ \Phi(\lambda_{MA}, \lambda_{DS}) \text{ for } \lambda_{MA}, \lambda_{DS} \in \mathbb{R}\} \tag{1}$$

A first model of the subject images is then:

$$I_k = T_0 \circ \Phi(\lambda^k_{MA}, \lambda^k_{DS}) \circ \phi_k + \epsilon_k$$

$\lambda^k_{MA}$ and $\lambda^k_{DS}$ are the subject morphological age and disease score while $\phi_k$ encodes differences specific to the subject morphology. Of course multiple options are available to combine the longitudinal deformation $\Phi$ with the subject specific changes $\phi^k$. Here we write this operation as a right-composition but this choice will be discussed in section 2.2 and a similar but symetrical 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 of the 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 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 [25].

In this paper, we use the stationary velocity field (SVF) framework [26] for its ability to describe complex and realistic diffeomorphic (smooth and invertible) brain deformations in a straightforward manner [27]. 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 $\phi$ we integrate trajectories along the vector field $v$ for a unit of time.

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

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

By writing $\Phi(\lambda_{MA}, \lambda_{DS}) = \text{Exp}(\lambda_{MA}\text{v}_A + \lambda_{DS}\text{v}_D)$, we propose a linear model in the SVF space (*i.e.* the space of the parameter of the deformations) parametrized by two SVFs $\text{v}_A$ and $\text{v}_D$. In particular, the two template trajectories are then separately parametrized: $\text{v}_A$ controls the normal aging template trajectory and $\text{v}_D$ the disease specific template trajectory.

Besides, for each subject, the processes are meant to be intertwined and it 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 [28] 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 $\Phi$ modeling the effects of the aging and the disease on a reference morphology $T_0$ is parametrized by two SVFs: $\text{v}_A$ and $\text{v}_D$. This ideal model generates a surface $\mathcal{T}$ of possible images describing the evolution of the template morphology:

$$\mathcal{T} = \{T_0 \circ \text{Exp}(\lambda_{MA}\text{v}_A + \lambda_{DS}\text{v}_D) \text{ for } \lambda_{MA}, \lambda_{DS} \in \mathbb{R}\} \tag{2}$$

## 2.3 Individual morphological variability and generative model

An individual image is modeled as follows:

$$I_k = T_0 \circ \text{Exp}((\lambda_{MA}^k - \lambda_0)\text{v}_A + \lambda_{DS}^k\text{v}_D) \circ \phi_k + \epsilon_k \tag{3}$$

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

$$\text{Exp}((\lambda_{MA}^k - \lambda_0)\text{v}_A + \lambda_{DS}^k\text{v}_D) \circ \phi_k = \text{Exp}((\lambda_{MA}^k - \lambda_0)\text{v}_A + \lambda_{DS}^k\text{v}_D + w_r^k)$$

In this formula, $\text{Exp}(w_r^k)$ is approximatively equal to $\phi_k$ 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 $\text{v}_A$ and $\text{v}_D$.

As we can see in Figure 2, the model parametrized by $\text{v}_A$ and $\text{v}_D$ allows us to characterize the subject morphology with two scalar variables, the morphological age $\lambda_{MA}^k$ and the disease score $\lambda_{DS}^k$, 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:

$$w^k = (\lambda_{MA}^k - \lambda_0)\text{v}_A + \lambda_{DS}^k\text{v}_D + w_r^k \tag{4}$$

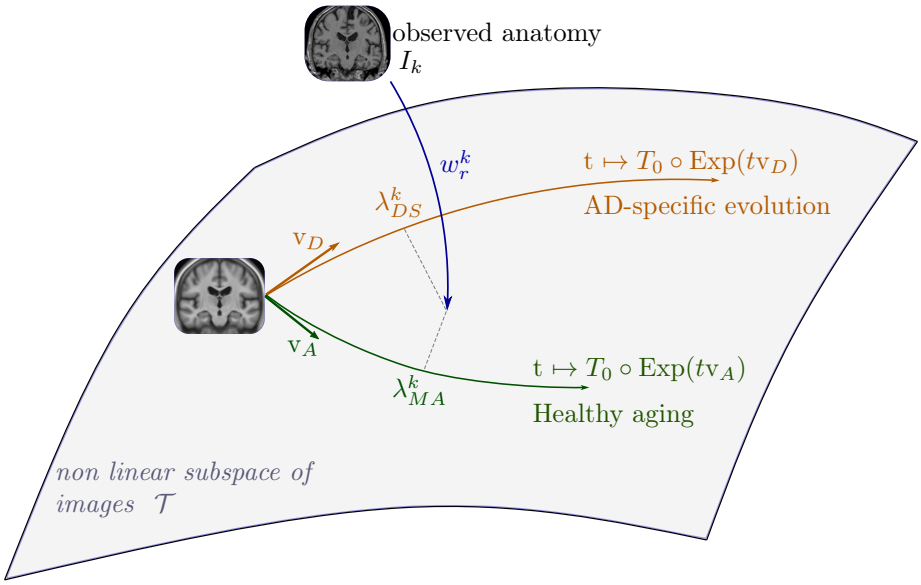



Figure 2: Our two-trajectory model. The template image $T_0$, the normal aging template trajectory parametrized by $\mathrm{v}_A$ and the disease specific template trajectory parametrized by $\mathrm{v}_D$ define a subspace of possible morphologies of reference. An individual morphology is characterized by a morphological age $\lambda_{MA}^k$, a disease score $\lambda_{DS}^k$, and an SVF $w_r^k$ modeling the subject-specific part. Each image can be projected on the template subspace $\mathcal{T}$ using a decomposition of the deformation between the image and the template $T_0$.

# 3 Estimation of the model parameters

The model parameters are of two kinds: the population parameters ($T_0$, $\mathrm{v}_A$ and $\mathrm{v}_D$) and the subjects parameters ($\lambda_{MA}^{k,i}$, $\lambda_{DS}^{k,i}$ and $w_r^{k,i}$). Moreover, the model includes an intensity noise and a deformation noise for every image, a regularization term for both template trajectories and subject specific deformations and every longitudinal evolution is described relatively to the subject-specific morphology. 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 [29]. In practice, it allows us to work only with intra-subject deformation to estimate the model population parameters. It simplifies the optimisation 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 subjects in the training set.

3. Intra-subject deformations are relatively small and smooth. Consequently, the registration

regularisation has less impact on the estimated deformation. It 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 ($T_0$, $\mathrm{v}_A$ and $\mathrm{v}_D$) are estimated using only intra-subject longitudinal evolutions.. Finally, the subject parameters are estimated for each subject.

### 3.1 Estimation of the template trajectories $\mathrm{v}_A$ and $\mathrm{v}_D$ 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. [30]. Images are preprocessed, rigidly aligned to the MNI-152 template and then longitudinally registered. Intra-subject deformations between follow-up images $I_{k,i}$ and the baseline image $I_{k,0}$ are computed using non-linear registration. The resulting intra-subject model in the subject space is estimated using an 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 $s^k_{MA}$), a disease progression rate (noted $s^k_{DS}$) and an orthogonal component (noted $v^k_r$):

$$v^k = s^k_{MA}\mathrm{v}_A + s^k_{DS}\mathrm{v}_D + v^k_r$$

The estimation is done on two groups of subjects: $\mathcal{G}_h$ with the healthy subjects and $\mathcal{G}_{ad}$ with patients diagnosed with AD. We assume that each healthy subject of $\mathcal{G}_h$ is aging at normal speed $s^k_{MA} = 1, \forall k \in \mathcal{G}_h$ and does not have any evolution toward the disease $s^k_{DS} = 0, \forall k \in \mathcal{G}_h$. Similarly, each patient of $\mathcal{G}_{ad}$ has a normal morphological aging rate $s^k_{MA} = 1, \forall k \in \mathcal{G}_{ad}$ and a constant unit disease progression rate $s^k_{DS} = 1, \forall k \in \mathcal{G}_{ad}$. Finally the subject specific components are assumed to be centered, uncorrelated and to have a fixed variance. The maximum likelihood problem writes:

$$\min_{\mathrm{v}_A, \mathrm{v}_D} \sum_{k \in \mathcal{G}_h} \|v^k - \mathrm{v}_A\|^2 + \sum_{k \in \mathcal{G}_{ad}} \|v^k - \mathrm{v}_A - \mathrm{v}_D\|^2 \tag{5}$$

The solution of the optimization problem is explicit:

$$\hat{\mathrm{v}}_A = \frac{1}{|\mathcal{G}_h|} \sum_{k \in \mathcal{G}_h} v^k \tag{6}$$

$$\hat{\mathrm{v}}_D = \left( \frac{1}{|\mathcal{G}_{ad}|} \sum_{k \in \mathcal{G}_{ad}} v^k \right) - \hat{\mathrm{v}}_A \tag{7}$$

We should however note that $\|\hat{\mathrm{v}}_A\|$ (resp. $\|\hat{\mathrm{v}}_D\|$) is a biased estimator of $\|\mathrm{v}_A\|$ (resp. $\|\mathrm{v}_D\|$). We detail the bias estimation in the Appendix 8.2.

### 3.2 Estimation of the template morphology $T_0$

The population specific template morphology is computed using the algorithm proposed by Guimond et al. [31] by alternating the registration of subject images to the template and the recomputation of the template intensity. However, in our approach, the subject image should not be registered to $T_0$ directly but to its projection on the template space.

To tackle this problem, we propose an iterative procedure where we iteratively 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 [27]):

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

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

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

**Algorithm 1:** Iterative registration algorithm

As 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 optimisation. The projection coordinates are kept from one iteration to the next and the images are registered to their estimated projections in the template space. The deformation update $u$ is then computed and finally the new atlas image $T$ and the estimated projections are updated (see Algorithm 2).

**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**
  $w_r^k = \mathrm{registration}(T \circ \mathrm{Exp}(w_t^k), I_k)$ ;
  $u = \mathrm{mean}(w_r^k)$ ;
  $T = \mathrm{mean}(I_k \circ \mathrm{Exp}(-w_r^k + u))$ ;
  $w_t^k = w_t^k + \mathrm{proj}_{\mathcal{T}}(w_r^k - u)$ ;
**until** *convergence*;

**Algorithm 2:** Iterative template space estimation algorithm

It is even possible to go further and to update the SVFs parametrizing $\mathcal{T}$ at each iteration. In this work, we do it and transport the subject intra-subject models to the template space and update the template trajectories at each iteration. $T$ is initialized using the MNI-152 template and the

convergence is manually assessed comparing the template for successive iterations. At convergence, we then have the template image $T_0$ and both template trajectories $\mathrm{v}_A$ and $\mathrm{v}_D$.

### 3.3 Estimation of the subject parameters

If the population parameters are set, 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 decomposed linearly $w^k = (\lambda^k_{MA} - \lambda_0)\mathrm{v}_A + \lambda^k_{DS}\mathrm{v}_D + w^k_r$ by solving the following linear system:

$$w^k \cdot \mathrm{v}_A = \|\mathrm{v}_A\|^2(\lambda^k_{MA} - \lambda_0) + \mathrm{v}_D \cdot \mathrm{v}_A\lambda^k_{DS} \tag{8}$$

$$w^k \cdot \mathrm{v}_D = \mathrm{v}_D \cdot \mathrm{v}_A(\lambda^k_{MA} - \lambda_0) + \|\mathrm{v}_D\|^2\lambda^k_{DS} \tag{9}$$

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

## 4 Results

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

#### 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 diseased patients. Atrophy of the aging brain and the effect of AD have been extensively studied [33] and the atrophy measurements may vary depending on the methodology and the population studied. In this work, we choose to prescribe piecewise-constant atrophy map with constant value in brain areas delimited by the segmentation software *FreeSurfer*. The atrophy value of a region is sampled around a fixed population mean with an additive Gaussian noise of relative standard deviation of 5% for every subject. The healthy population is designed to have a slight atrophy in the whole brain while the diseased 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. [34] for healthy aging and in Carmichael et al. [35] with an additional scaling for the pathological evolution. We detail the exact regional specifications in table 2 in the appendix (see 8.1).

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 the pathological and the healthy settings. The deformation is then extrapolated 5 times before being applied to the original image. We then have two matched populations of 40 pairs of images.

#### 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 (see Figure 3). The atrophy ground truth is computed using the estimated template anatomy as the anatomical variability is not explicitly model.

The estimated atrophy patterns are smoothed versions of the simulated ground truth. This effect was already observed [32]. First of all, the registration algorithm is unaware of the underlying simulation model and is unable to localize the atrophy in homogeneous areas. Moreover, the spatial regularization of the registration and the parametrisation using SVFs are also smoothing the estimated atrophy patterns. It is particularly visible in small (hippocampus) or thin regions (cortex). This results in a consistent bias when the atrophy measurements are integrated over the regions (see Figure 15 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 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 more reproducible [36, 37]. We want to see how these limitations of the deformation based morphometry are translated into the proposed age and disease progression markers.

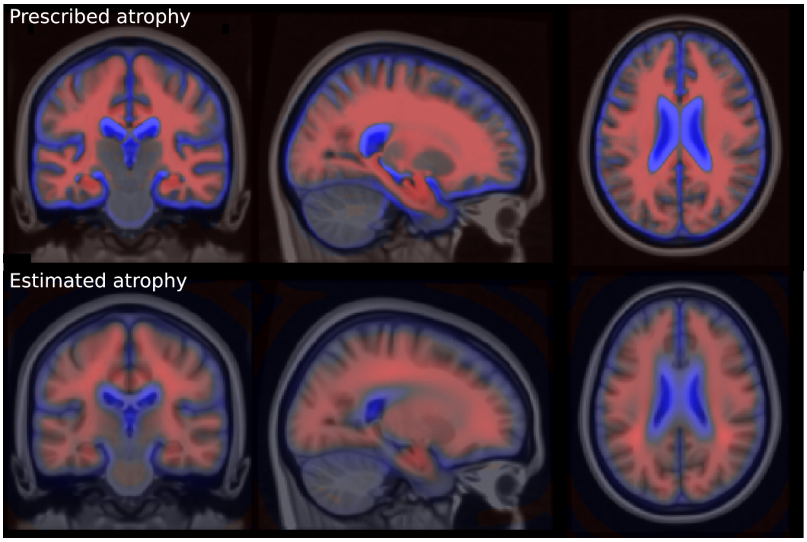
#### 4.1.3 Imaging biomarkers estimation

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 can compare the simulated differences at t=5 or the evolution of the cross-sectional assessments for each subject (see Figure 4).

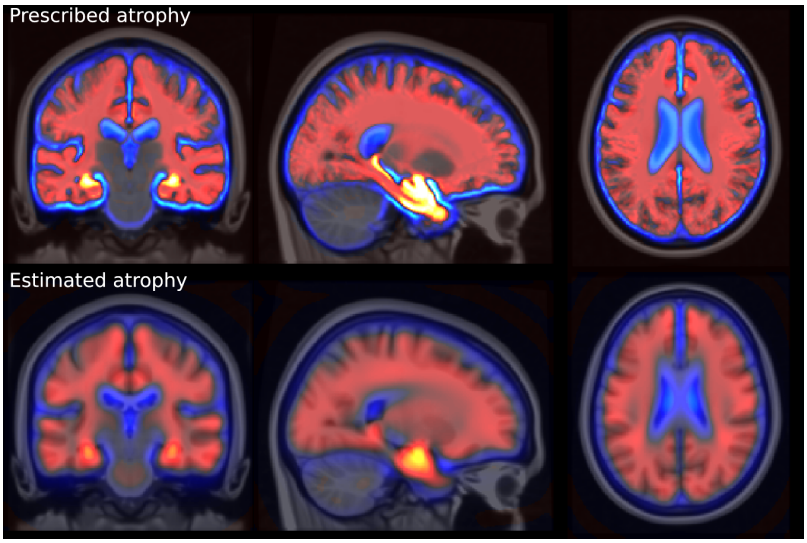
In this experiment, the initial anatomical variability is really important, indeed the standard deviation at t=0 is equal to 48 years for MA(0) and 24 years for DS(0). At t=5, a difference is visible for the disease score but it is still diluted in the inter-subject variability. It is also possible to 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. Here, the statistical analysis gives that, for a significativity level of 0.05, the disease score would be significantly discriminant after 13 years of evolution. It is done for synthetic data but it highlights the slowness of the disease and the interest in modeling to be able to extrapolate the evolutions.

Looking at the evolution of these cross-sectional biomarkers, we see that the measure is relatively stable despite the relatively large inter-subject variability. It seems that doing the measure at two different time points gives a good estimate of the longitudinal evolution. Even if the simulation images was not done independently of our generative model, it is able to explained most of the simulated changes. In practice the estimations are a bit 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 the mean of the estimation is equal to 3.92 for the healthy group and to 3.57 for the diseased group. But more importantly, the standard deviation is really 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 $\Delta DS$, the mean is close to 0 for the healthy group and close to 5 for the diseased group. 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 $\Delta DS$.

In a clinical context where longitudinal data is not always available, it is important to get as




(a) Healthy evolution




(b) Pathological evolution


relative volume change (year$^{-1}$)

Figure 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 estimation is smoother but even local pattern such as in the one in hippocampal area are captured and globally the simulated and estimated maps are qualitatively similar.

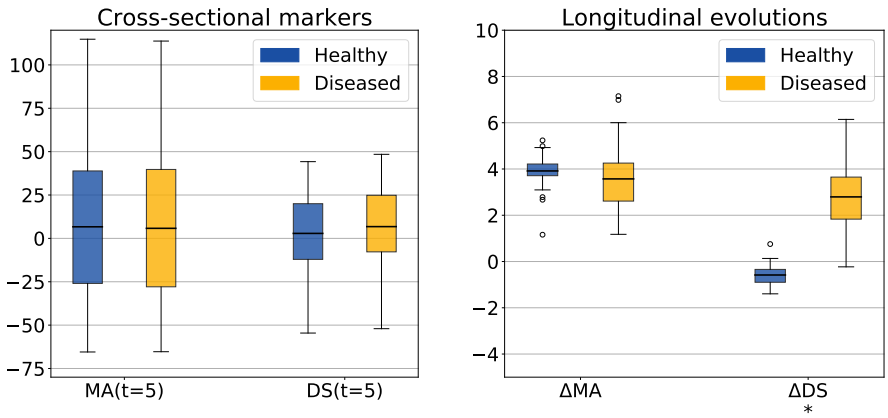



Figure 4: Evolution of the imaging biomarkers estimated on simulated data. MA=morphological age, DS=disease score. The longitudinal evolutions $\Delta$MA and $\Delta$DS are the differences between the two cross-sectional assessments i.e. $\Delta$MA = MA(5) - MA(0). The star indicates that the difference between the healthy and the diseased subjects is significant (p-value $< 0.05$) for an unpaired t-test. By construction there is no difference between the two populations at baseline. The changes are generally underestimated when using the difference between the two cross-sectional measurements but the longitudinal evolutions show the stability of the estimation despite a strong variability inter-subject.

much information as possible from every cross-sectional measurement. We see here that the markers may not be the best to separate the two populations but the cross-sectional measurement gives a relatively stable assessment and their evolutions are strongly associated with the clinical diagnosis.

## 4.2 Experiments with real data

Longitudinal T1 sequences were obtained from the Alzheimer's Disease National Initiative (ADNI). 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 normal diagnosis and the MCI stable group are also sub-classified using the beta-amyloid 1-42 biomarker that has been shown to be related to prodromal condition. 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+, MCIc (MCI converter to AD) and AD (diagnosed with Alzheimer's disease starting from the beginning). The table 1 sums up the demographic description of the population. We can notice that the group sizes are a bit unbalanced as we divided the cognitively normal group (272 subjects in total) in three but otherwise the groups are well matched in age with slightly uneven gender distributions.

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

| 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 1: Socio-demographic and clinical information of the study cohort. Standard deviations are shown in brackets.

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).

#### 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 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.

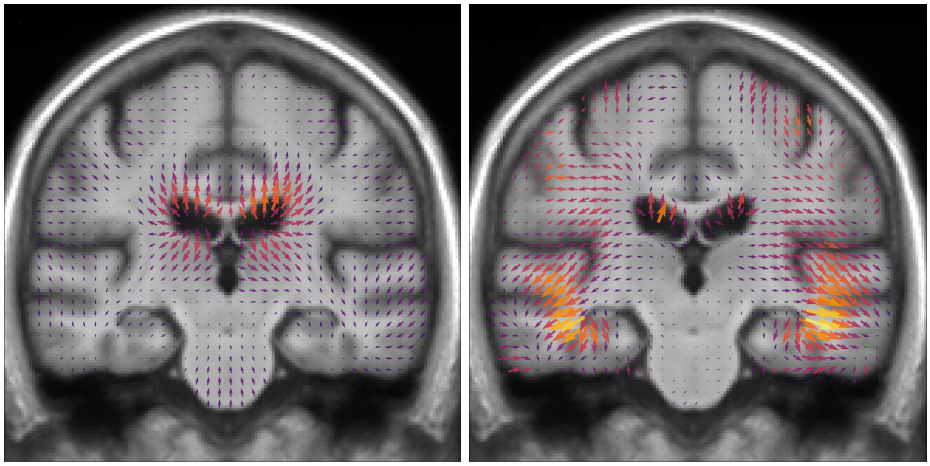
Atrophy patterns can be measured by the divergence of the velocity field. We see in Figure 6 that they are similar to what was already observed in the past for healthy and the diseased subjects [36]. The precise localization of the atrophy is always difficult with a morphometrical approach but we see for the healthy group a well spread and relatively small atrophy while the additional disease specific atrophy is particularly strong in the temporal area. It also seems that the remaining atrophy in the rest of the brain, is mainly located near the cortical surface.

The generative model can be used to get images in the reference template plane to visualize the modeled morphologies. Figure 7 shows the evolution of the template morphology in the two main directions. It shows the evolution over 20 years. The difference of intensity overlay is used to shows the changes of 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 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.

#### 4.2.2 Intra-subject variability of our progression markers

In Lorenzi et al. [24], 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. And if the subject is healthy, we can expect a linear increase with a slope of 1. We should also see the increase of disease score for the patients while this marker stays stable and close to 0 for the healthy subjects. These hypotheses are evaluated using data from the ADNI datasets.

The morphological age and the disease score are computed for each subject at each time point. Figure 8 shows the evolution of these cross-sectional measurements. First, at the population level,



(a) Normal aging template trajectory (b) Disease specific template trajectory

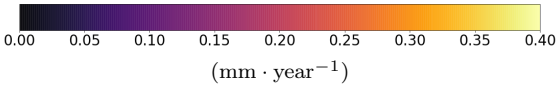


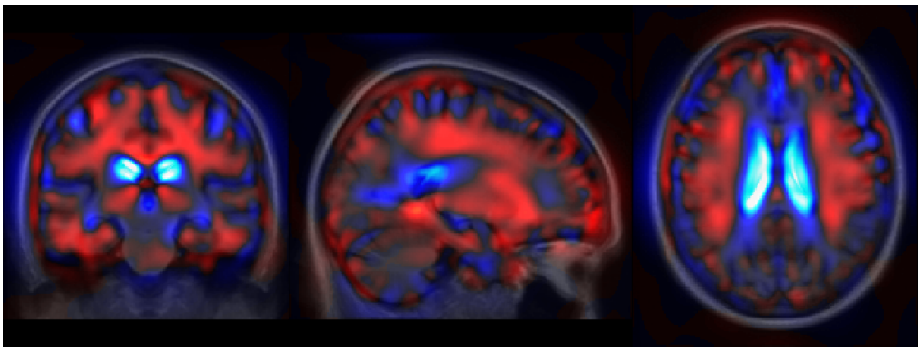

Figure 5: Template image and SVFs parametrizing the two template trajectories SVFs. Left: normal aging trajectory $v_A$ showing a ventricular expansion related to a global atrophy. Right: disease specific trajectory $v_D$ showing specific patterns, especially in the temporal lobes around the hippocampi areas. The color encodes the amplitude of the velocity at each position.

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 also steadily increases for each subject. Finally, we can also note that the AD subjects look older, age faster and have a higher disease score than the healthy ones.

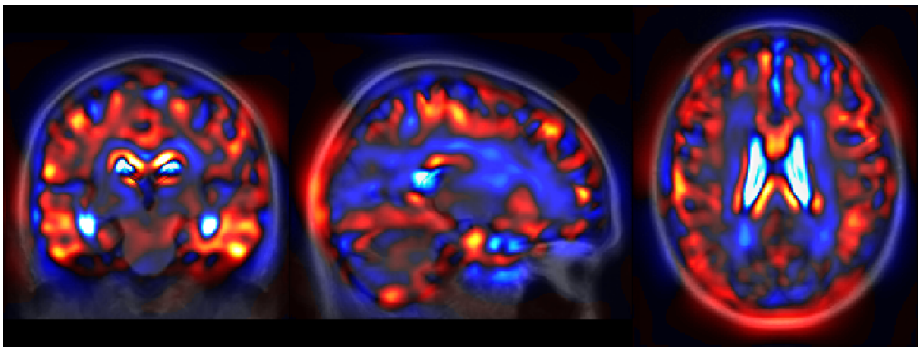
A linear random effects model can help us to quantify these observations. The model is fitted, for both morphological age and disease score, with fixed effects on age and sex and a random intercept and slope for each subject. The focus is set on the analysis of the regression for the CN and the AD groups.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

The model is first fitted to the morphological age measures in the CN group. It gives a coefficient of $0.26 \pm 0.11$ for the fixed effect of age while the mean subject slope is $0.10 \pm 0.02$. 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 significatively) coefficient for male subjects ($1.81 \pm 1.2$) 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 $0.12 \pm 0.1$ for the fixed effect of age and $0.12 \pm 0.01$ for the mean individual slope.

For the AD group, the linear model is also well adapted (p-values inferior to $10^{-6}$ for the



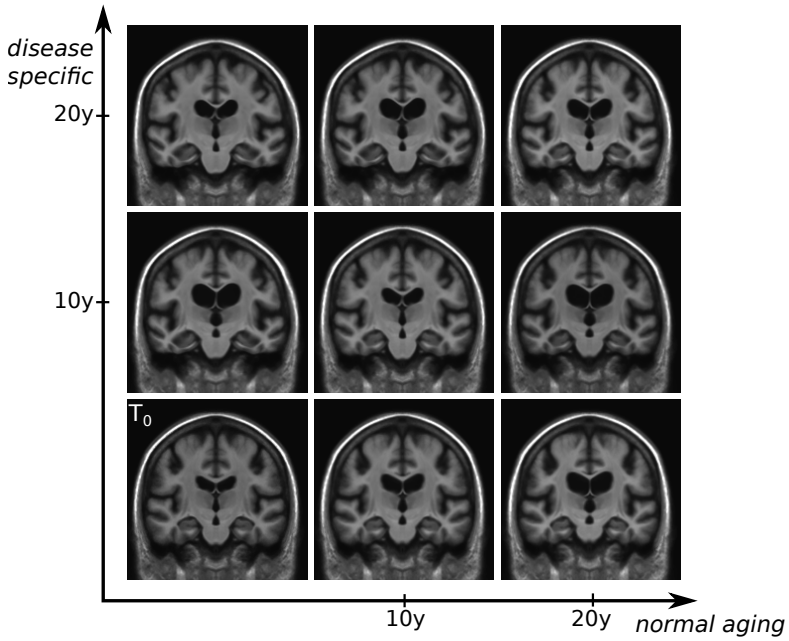
(a) Atrophy along the normal aging template trajectory



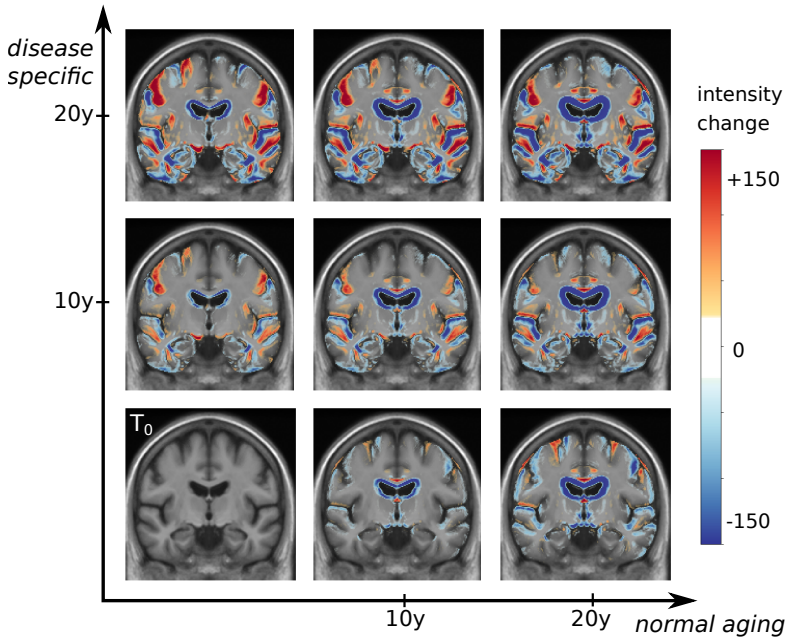
(b) Atrophy along the disease specific template trajectory


relative volume change ($\text{year}^{-1}$)

Figure 6: Atrophy measured by the divergence of the SVFs parametrizing the two template trajectories.




(a) Images in the atlas plane $\mathcal{T}$




(b) Images in the atlas plane $\mathcal{T}$ with intensity differences

Figure 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 these figures, 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.

likelihood ratio test). The main remark is probably that the intra-subject slopes are in average more important than for healthy subjects (around $0.52 \pm 0.06$ for the morphological age and $0.71 \pm 0.05$ for the disease score) while the fixed effect related to age of $0.23 \pm 0.07$ (for the morphological age) and $0.14 \pm 0.04$ (for the disease score) are more similar to the one observed previously.

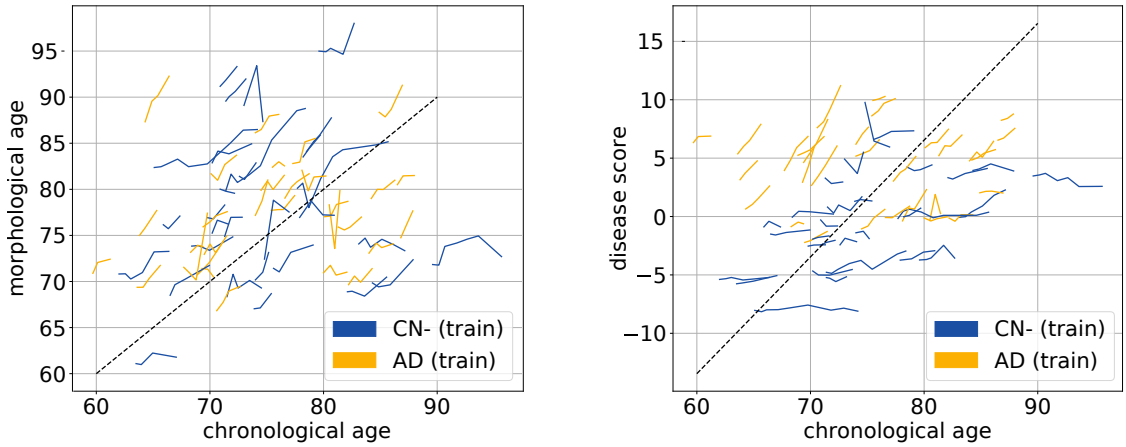



Figure 8: Evolution of cross-sectional markers for every subject of the two training sets. [Left] chronological age, the dashed line corresponds to the expected evolution of healthy subjects i.e. the morphological age is equal to the chronological age. [Right] disease score, the dashed line is the expected pathological evolution.

#### 4.2.3 Cross-sectional discriminating power

We saw that the two markers are able to represent realistic evolutions of healthy and diseased subjects. We want to study their relation to the disease progression. We start with a simple discriminant analysis using only the first image available for each subject.Figure 9 shows the distribution of the estimations for each group. We see a gradual increase of both markers towards more advanced 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. Both markers show the same trend, but marginal pattern can be distinguished. For instance, the difference between the MCI stable and the MCI converters is stronger for the morphological age while the disease score differentiate more between the MCIs- and the MCIs+. 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 classification using the scikit-learn python library. The error penalty weights are adjusted between the two classes to balance the trade-off between true positive rate (correct classification of AD subjects) and false positive rate (incorrect classification of healthy ones). The mean classification accuracy using a 10-folds cross-validation scheme is equal to 0.59. Of course we do not reach the performance of state-of-the-art dedicated algorithm (the same classifier using the normalised

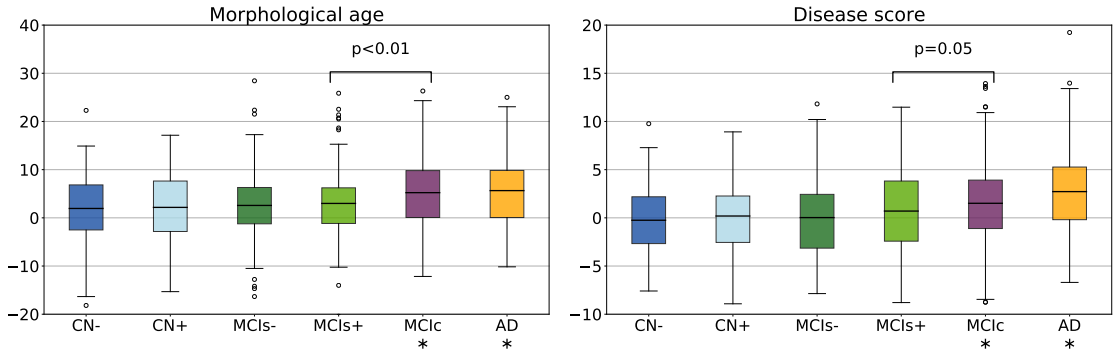



Figure 9: Box-plot of the group-wise markers estimated for the clinical groups. Stars below the name of the group indicate a significant difference to the CN- group for a t-test at the level 0.05. Both markers gradually increase towards more advanced disease states.

volume of the hippocampi has a 0.67 accuracy) but it allows us to see how both markers are associated with the diagnosis. Moreover this discriminative approach could be extended by using information in a subset of targeted areas.

Here, the linear decision function is equivalent to the projection on the SVF $\mathrm{v}_A - 0.003\mathrm{v}_D$, so the differences between MCIs and MCIc subjects is, in our model, only associated with the aging trajectory. The disease specific changes do not seem to have an impact before the conversion. We can compare this to the similar experiment between CN and AD where the linear classifier corresponds to the projection on $\mathrm{v}_A + 0.49\mathrm{v}_D$ so approximately $(v_{bn} + v_{ad})/2$, i.e. the mean evolution of the whole population.

#### 4.2.4 Regional analysis of the progression

In the context of Alzheimer's disease, most of the morphological changes are known to be located in the temporal lobe. Using the AAL atlas [38], 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. Results are shown in Figure 10.

The region is not adapted 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.

#### 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 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. It is a bit more complex for the disease score. The evolution is almost

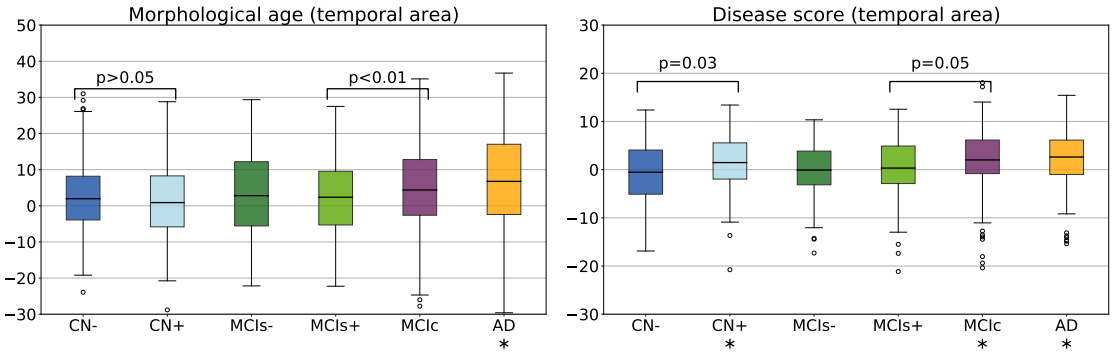



Figure 10: Box-plot of the markers estimated in the temporal lobes. Stars below the name of the group indicate a significant difference to the CN- reference for a t-test at the level 0.05. The area known to be related to the AD makes the disease score estimation less sensitive to the overall noise.

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 converter subjects, but also between MCIs- and MCIs+ (or more generally between subjects with negative amyloid or positive amyloid marker). It means that we are able to capture the global progression of the disease. The changes are stronger for diagnosed patients but similar patterns of evolution are observed in the early stage of the disease. It is a good sign to indicate that the marker may be able to identify precursor sign of the disease, however it is difficult to draw definitive conclusions because of the difference is small relatively to the variability and the statistical tests are not corrected for multiple comparisons. A significant difference is also observable between the CN- and CN+ group for the temporal disease score slope.

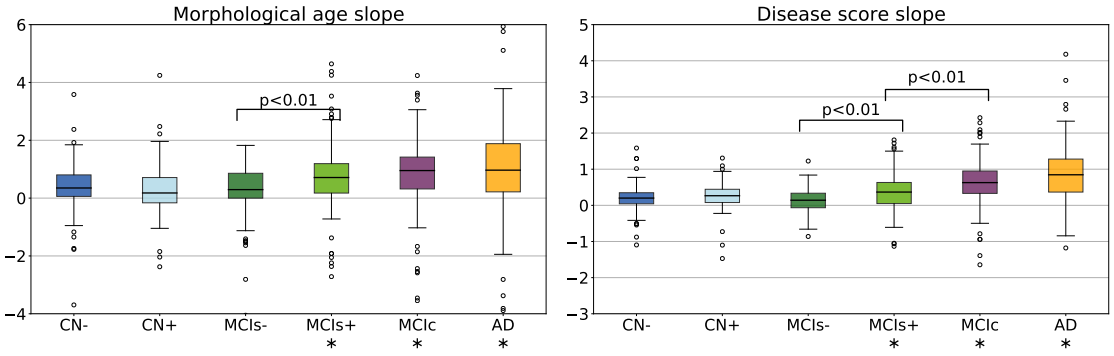



Figure 11: Box-plot of the rate of evolution of the markers computed using individual linear regressions. Stars indicate a significant difference to the CN- reference for a t-test at the level 0.05. Top row shows the results for the whole brain while the bottom row shows the result for the temporal lobe only. A gradation is visible from the healthy to the diseased 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 divergence 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 observer previously and be in part explained by the estimation procedure: the norm of the SVF parametrizing the template trajectory and by consequence the normalization are known to be biased (see 8.2).

## 4.3 Generating diagnosis driven morphological evolution

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

### 4.3.1 Modeling the group evolution

To predict a subject evolution it is natural to look at subjects that are similar. This similarity will be based on the morphological age and the disease score and will be restricted to subjects of the same clinical condition.

In practice, each subject is associated with a couple (position, speed) in the parameter plane, then, for each group, a vector field is regressed using a kernel ridge regression to estimate the local speed at every point. It uses a RBF kernel with a spatial scale set to 10 years (for both the morphological age and the disease score axis) in order 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 12 for the CN (i.e CN-, CN+ together), MCIs, MCIc and AD groups. The figure is centered on the data and the extrapolation can be really unreliable.

Differences in amplitude, i.e. speed of evolution, and orientation are clearly visible between the groups and echo the results about the linear regression coefficients 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 a 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. It 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 a bit stronger but overall really similar to the CN one. The MCI converters however show a stronger and more vertical evolution. The angle of the average vector is equal to 38° while it was equal 30° (resp. 31°) for the MCIs (resp. CN) group. We should also note that subjects with high morphological age and low disease score (bottom right) seem to follow a different, more horizontal evolution. It means 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 (angle equal to 46° for the average). A main evolution is visible from bottom left to top right with a sightly 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. The difference between the main evolution these subjects in the bottom right gives a possible stratification of the disease in two sub-categories.

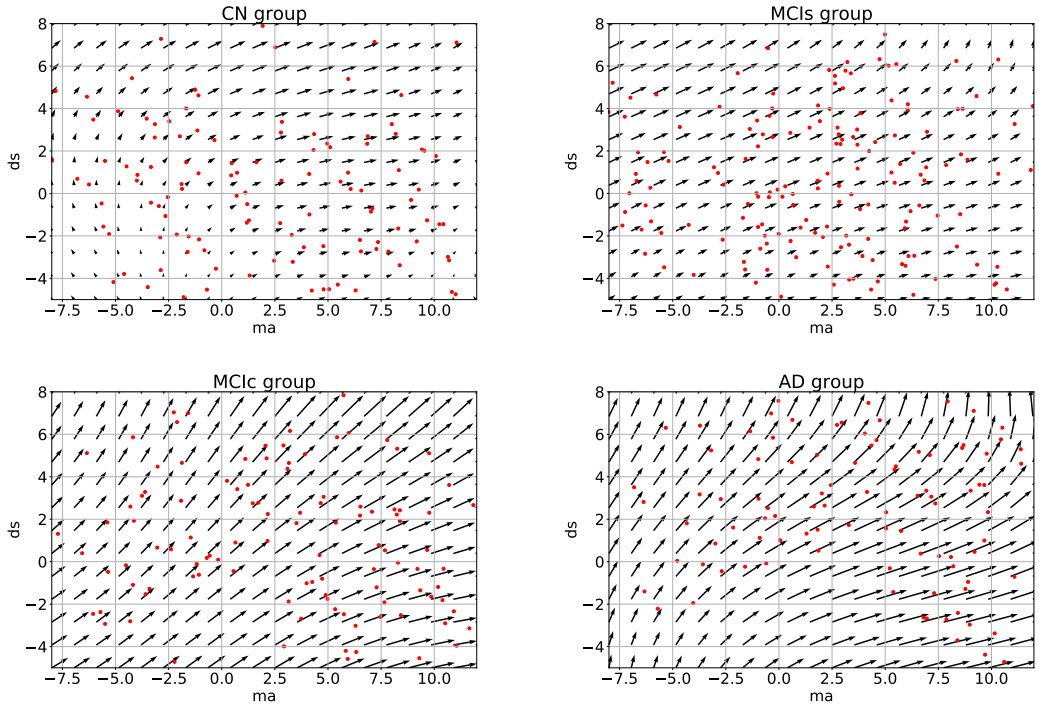



Figure 12: Results of the kernel ridge regression for the markers' evolution for the CN, MCIs, MCIc and AD groups. It shows the regressed vector field with data points shown in red. Amplitude and orientation variability is visible between the groups (with stronger and more vertical evolution for AD than for CN) but also within each group giving non-linear morphological evolutions.

### 4.3.2 Generating a subject specific evolution

The regression model can then be used to extrapolate the evolution of one subject. We choose here to model the evolution of MCIc subject in order to capture the changes happening around the time when the disease is diagnosed.

For example, Figure 13 shows several evolutions 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.

The observed evolution is shown by the dashed black line, the dots correspond to acquisition times available for this subject. The measurements can be a bit erratic between successive acquisitions but a global trend is visible. The evolution is globally between the healthy and the pathological trajectories. Using only the data from this subject, we need several years to get a stable assessment

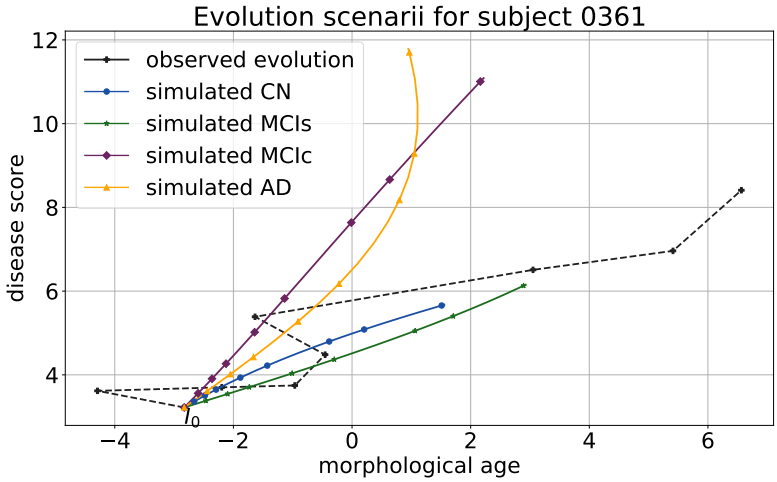



Figure 13: Real and generated marker evolutions for the subject 0361. The observed evolution appear to be between the CN and the AD evolution and close to the MCIc one.

of the morphological evolution while the models regressed from the population are able to simulate realistic changes in the medium to long term.

We see that the observation is more similar to the MCIs or the MCIc model. It highlights that using additional clinical condition improves the model as our cross-sectional imaging biomarkers do not fully capture the status of the subject relatively to the disease and are not designed to replace a diagnosis or any clinical information.

These markers changes can be directly translated in brain images to visualize the morphological evolution. Figure 14 shows the results for the end point of the trajectories. Results are shown for the CN and AD models only. We have seen that the estimation of the markers was biased. For both the synthetic dataset and the training samples the measured changes only correspond to 80% of the expected value. In order to get images that correspond to the desired changes we then multiply the deformations by 1.25. Images are generated by deforming the baseline image using the simulated deformation transported in the subject space.

The cortical atrophy in particular in the temporal lobe is clearly visible in the real image and to a lesser extent in the AD simulation. The expansion of the ventricles is also visible in every image, but again even in the AD simulation the volume change is inferior to the observation. It seems that we only partially capture the changes related to AD. Other aspects of the evolution are hard to quantify and often poorly documented. For example the evolution of the shape of the ventricles 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 of lowering of the temporal lobe and a local rotation in the image for the real case and the AD simulation. Overall, the morphological changes simulated looks realistic, even if they do not matches perfectly the observed changes because the model is generic and really simple. It gives a range of possible evolution that can be further exploited to see what is expected or not.

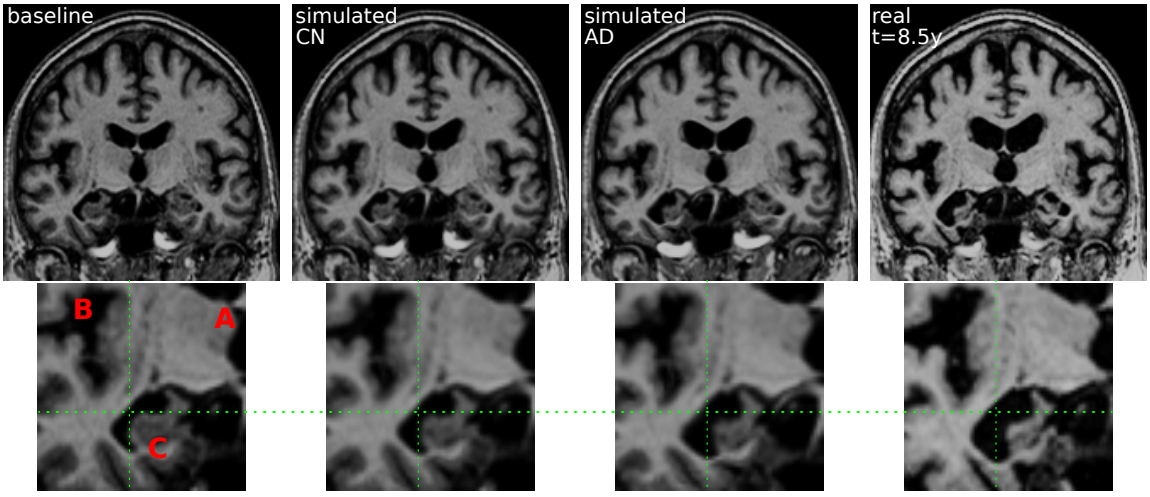



Figure 14: Evolution of the subject 0361 over more 8.5 years. The subject was diagnosed MCI (for 5y) then AD. From left to right: (1) real image at baseline, (2) simulated image from the baseline image using the healthy (CN) evolution, (3) using the pathological (AD) evolution, (4) real image at t=8.5y. The second row zooms in the most interesting areas between the ventricle (A), the lateral sulcus (B) and the hippocampus area (C). 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.

# 5 Contributions and limitations

In this work, we proposed to describe the evolution of elderly patients using two trajectories that model the morphological evolution of the brain for healthy or pathological cases. 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 which is not seen as a simple accelerated aging process or the divergence from the normal evolution. However we are not able to fully describe the morphological changes that can occur while aging and the description is not precise enough to model accurately 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 change the approach.

## 5.1 Approximations and technical limitations

One of the limitations of our approach is the difficulty to accurately estimate the markers as it can be seen in the longitudinal intra-subject evolutions where some time points may look like outliers. A change of calibration of the MR scanner or a bad pause correction can affect the registration results and then the linear decomposition of the SVF is not adapted to this kind of variability. Using the model to guide the registration could help to combine the constraints imposed on the deformations by the registration and by the model.

Then, in this work, we used an orthogonal projection using the $L2$ scalar product in the SVF

space to define the subject specific deformation. It provides an intuitive interpretation and a simple decomposition but this choice is arbitrary and we saw that, for example, using another region of interest could totally change this constraint. It would be interesting to match the subject with the closest morphology in the model for a metric more adapted to deformations. Another solution could try to decorrelate the two information (for example using an ICA).

We saw, using synthetic and longitudinal data, that the we were under-estimating the speed of evolution. This bias is partially explained by the estimation procedure that does not take into account the uncertainty and noise in the population parameters of the model. It is possible to estimate the magnitude of this bias but it is not easy to correct it for two reasons: first it is hard to quantify the bias caused by the use of the same subjects for the template anatomy and the healthy template trajectory, 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.

## 5.2 Possible extensions of the description

Probably the main limitation of our approach is the inter-subject variability of the markers. It is related to the inter-subject anatomical variability that is not easy to model. Better markers should be less sensitive to these subject differences and more specific to time related changes. An alternative would be to somehow align the different evolutions before comparing them.

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 or we could do something similar to what was proposed by Rohé et al. [39] using a barycentric subspace prior to the registration.

The use of a segmentation to compute the progression markers in multiple regions would be one 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 as the spatial analysis of brain deformation remains a research topic.

The modeling of the disease is also really simple and one dimensional. And, in practice, the subject specific field $w_{ss}^k$ is actually encoding changes that are related to the disease and its variability but are not currently modeled. To enhance the description, it would be interesting be able to really integrate the clinical condition and maybe several sub-condition for the diseased patients in the model. For example we saw that the MCIc subjects evolve similarly to the AD subjects. It suggest that a more advanced stage of the disease corresponds to a later stage but with no changes of the ongoing process. It 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.

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

# 6 Conclusion and perspectives

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 disease progression with 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, even if they are by construction strongly dependent on the original anatomy of the brain, they are also consistent for intra-subject longitudinal analyses and 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 condition.

Further work should be done in order to improve the model parameters estimation to make the markers less anatomically dependent. It would also be interesting to take advantage of the regional description to produce a spatially varying disease progression model for each subject.

# 7 Acknowledgment

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.;Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.;Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

# References

[1] Zhores A Medvedev. An attempt at a rational classification of theories of ageing. *Biological Reviews*, 65(3):375–398, 1990.

[2] Denise C Park and Patricia Reuter-Lorenz. The adaptive brain: aging and neurocognitive scaffolding. *Annual review of psychology*, 60:173–196, 2009.

[3] Allyson C Rosen, Matthew W Prull, John DE Gabrieli, Travis Stoub, Ruth O'hara, Leah Friedman, Jerome A Yesavage, and Leyla deToledo Morrell. Differential associations between entorhinal and hippocampal volumes and memory performance in older adults. *Behavioral neuroscience*, 117(6):1150, 2003.

[4] Karen M Rodrigue and Naftali Raz. Shrinkage of the entorhinal cortex over five years predicts memory performance in healthy adults. *Journal of Neuroscience*, 24(4):956–963, 2004.

[5] Xiaojing Long, Weiqi Liao, Chunxiang Jiang, Dong Liang, Bensheng Qiu, and Lijuan Zhang. Healthy aging: an automatic analysis of global and regional morphological alterations of human brain. *Academic radiology*, 19(7):785–793, 2012.

[6] Elizabeth R Sowell, Bradley S Peterson, Paul M Thompson, Suzanne E Welcome, Amy L Henkenius, and Arthur W Toga. Mapping cortical change across the human life span. *Nature neuroscience*, 6(3):309–315, 2003.

[7] 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*, April 2017. ISSN 1359-4184, 1476-5578. doi: 10.1038/mp.2017.62. URL `http://www.nature.com/doifinder/10.1038/mp.2017.62`.

[8] Katja Franke, Gabriel Ziegler, Stefan Klöppel, Christian Gaser, Alzheimer's Disease Neuroimaging Initiative, et al. 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.

[9] Charles DeCarli, Joseph Massaro, Danielle Harvey, John Hald, Mats Tullberg, Rhoda Au, Alexa Beiser, Ralph D'Agostino, and Philip 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.

[10] Evert FS van Velsen, Meike W Vernooij, Henri A Vrooman, Aad van der Lugt, Monique MB Breteler, Albert Hofman, Wiro J Niessen, and M Arfan Ikram. Brain cortical thickness in the general elderly population: the rotterdam scan study. *Neuroscience letters*, 550:189–194, 2013.

[11] Alzheimer's Association and others. 2017 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 13(4):325–373, 2017. URL `http://www.sciencedirect.com/science/article/pii/S1552526017300511`.

[12] Takashi Ohnishi, Hiroshi Matsuda, Takeshi Tabira, Takashi Asada, and Masatake 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.

[13] Christos Davatzikos, Feng Xu, Yang An, Yong Fan, and Susan 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.

[14] Daniel Schmitter, Alexis Roche, Bénédicte Maréchal, Delphine Ribes, Ahmed Abdulkadir, Meritxell Bach-Cuadra, Alessandro Daducci, Cristina Granziera, Stefan Klöppel, Philippe Maeder,

Reto Meuli, and Gunnar Krueger. An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer's disease. *NeuroImage: Clinical*, 7:7–17, 2015. ISSN 22131582. doi: 10.1016/j.nicl.2014.11.001. URL `http://linkinghub.elsevier.com/retrieve/pii/S221315821400165X`.

[15] Stefan Klöppel, Ahmed Abdulkadir, Clifford R Jack, Nikolaos Koutsouleris, Janaina Mourão-Miranda, and Prashanthi Vemuri. Diagnostic neuroimaging across diseases. *Neuroimage*, 61 (2):457–463, 2012.

[16] Hubert M. Fonteijn, Marc Modat, Matthew J. Clarkson, Josephine Barnes, Manja Lehmann, Nicola Z. Hobbs, Rachael I. Scahill, Sarah J. Tabrizi, Sebastien Ourselin, Nick C. Fox, and Daniel 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, April 2012. ISSN 10538119. doi: 10.1016/j.neuroimage.2012.01.062. URL `http://linkinghub.elsevier.com/retrieve/pii/S1053811912000791`.

[17] Michael C. Donohue, Hélène Jacqmin-Gadda, Mélanie Le Goff, Ronald G. Thomas, Rema Raman, Anthony C. Gamst, Laurel A. Beckett, Clifford R. Jack, Michael W. Weiner, Jean-François Dartigues, and Paul S. Aisen. Estimating long-term multivariate progression from short-term data. *Alzheimer's & Dementia*, 10(5):S400–S410, October 2014. ISSN 15525260. doi: 10.1016/j.jalz.2013.10.003. URL `http://linkinghub.elsevier.com/retrieve/pii/S1552526013028732`.

[18] Claire Cury, Marco Lorenzi, David Cash, Jennifer M. Nicholas, Alexandre Routier, Jonathan Rohrer, Sebastien Ourselin, Stanley Durrleman, and Marc 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.

[19] Igor Koval, J-B Schiratti, Alexandre Routier, Michael Bacci, Olivier Colliot, Stéphanie Allassonnière, Stanley Durrleman, Alzheimer's Disease Neuroimaging Initiative, et al. 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.

[20] Jean-Baptiste Schiratti, Stéphanie Allassonniere, Olivier Colliot, and Stanley Durrleman. A Bayesian mixed-effects model to learn trajectories of changes from repeated manifold-valued observations. 2016. URL `https://hal.archives-ouvertes.fr/hal-01540367/`.

[21] Murat Bilgel, Jerry L. Prince, Dean F. Wong, Susan M. Resnick, and Bruno M. Jedynak. A multivariate nonlinear mixed effects model for longitudinal image analysis: Application to amyloid imaging. *NeuroImage*, 134:658–670, July 2016. ISSN 10538119. doi: 10.1016/j.neuroimage.2016.04.001. URL `http://linkinghub.elsevier.com/retrieve/pii/S1053811916300349`.

[22] Marco Lorenzi, Maurizio Filippone, Giovanni B Frisoni, Daniel C Alexander, Sébastien Ourselin, Alzheimer's Disease Neuroimaging Initiative, et al. Probabilistic disease progression modeling to characterize diagnostic uncertainty: application to staging and prediction in alzheimer's disease. *NeuroImage*, 2017.

[23] Katja Franke and Christian Gaser. Longitudinal Changes in Individual *BrainAGE* in Healthy Aging, Mild Cognitive Impairment, and Alzheimer's Disease. *GeroPsych*, 25(4):235–245, January 2012. ISSN 1662-9647, 1662-971X. doi: 10.1024/1662-9647/a000074. URL `http://econtent.hogrefe.com/doi/abs/10.1024/1662-9647/a000074`.

[24] Marco Lorenzi, Xavier Pennec, Giovanni B. Frisoni, Nicholas Ayache, Alzheimer's Disease Neuroimaging Initiative, and others. Disentangling normal aging from Alzheimer's disease in structural magnetic resonance images. *Neurobiology of aging*, 36:S42–S52, 2015. URL `http://www.sciencedirect.com/science/article/pii/S0197458014005594`.

[25] Gary E Christensen, Richard D Rabbitt, and Michael I Miller. 3d brain mapping using a deformable neuroanatomy. *Physics in medicine and biology*, 39(3):609, 1994.

[26] Vincent Arsigny, Olivier Commowick, Xavier Pennec, and Nicholas Ayache. A log-euclidean framework for statistics on diffeomorphisms. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2006*, pages 924–931, 2006.

[27] Marco Lorenzi, Nicholas Ayache, Giovanni B. Frisoni, and Xavier Pennec. LCC-Demons: a robust and accurate symmetric diffeomorphic registration algorithm. *NeuroImage*, 81:470–483, 2013. URL `http://www.sciencedirect.com/science/article/pii/S1053811913004825`.

[28] Matias Bossa, Monica Hernandez, and Salvador Olmos. Contributions to 3d diffeomorphic atlas estimation: application to brain images. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2007*, pages 667–674, 2007.

[29] Marco Lorenzi and Xavier Pennec. Geodesics, parallel transport & one-parameter subgroups for diffeomorphic image registration. *International journal of computer vision*, 105(2):111–127, 2013. URL `http://link.springer.com/article/10.1007/s11263-012-0598-4`.

[30] Mehdi Hadj-Hamou, Marco Lorenzi, Nicholas Ayache, and Xavier Pennec. Longitudinal analysis of image time series with diffeomorphic deformations: a computational framework based on stationary velocity fields. *Frontiers in neuroscience*, 10, 2016. URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4891339/`.

[31] Alexandre Guimond, Jean Meunier, and Jean-Philippe Thirion. Average brain models: A convergence study. *Computer vision and image understanding*, 77(2):192–210, 2000.

[32] Bishesh Khanal, Marco Lorenzi, Nicholas Ayache, and Xavier Pennec. A biophysical model of brain deformation to simulate and analyze longitudinal MRIs of patients with Alzheimer's disease. *NeuroImage*, pages 35–52, July 2016. doi: 10.1016/j.neuroimage.2016.03.061. URL `https://hal.inria.fr/hal-01305755`.

[33] Lorenzo Pini, Michela Pievani, Martina Bocchetta, Daniele Altomare, Paolo Bosco, Enrica Cavedo, Samantha Galluzzi, Moira Marizzoni, and Giovanni B. Frisoni. Brain atrophy in alzheimer's disease and aging. *Ageing Research Reviews*, 30:25 – 48, 2016. ISSN 1568-1637. doi: https://doi.org/10.1016/j.arr.2016.01.002. URL `http://www.sciencedirect.com/science/article/pii/S1568163716300022`. Brain Imaging and Aging.

[34] Anders M. Fjell, Kristine B. Walhovd, Christine Fennema-Notestine, Linda K. McEvoy, Donald J. Hagler, Dominic Holland, Kaj Blennow, James B. Brewer, Anders M. Dale, and the

Alzheimer's Disease Neuroimaging Initiative. Brain atrophy in healthy aging is related to csf levels of ab1-42. *Cerebral Cortex*, 20(9):2069–2079, 2010. doi: 10.1093/cercor/bhp279. URL http://dx.doi.org/10.1093/cercor/bhp279.

[35] Owen Carmichael, Donald G. McLaren, Douglas Tommet, Dan Mungas, and Richard N. Jones. Coevolution of brain structures in amnestic mild cognitive impairment. *NeuroImage*, 66:449 – 456, 2013. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2012.10.029. URL http://www.sciencedirect.com/science/article/pii/S1053811912010324.

[36] Mehdi Hadj-Hamou. *Beyond volumetry in longitudinal deformation-based morphometry: application to sexual dimorphism during adolescence*. PhD thesis, Université Côte d'Azur, 2016. URL https://tel.archives-ouvertes.fr/tel-01416569/.

[37] David M. Cash, Chris Frost, Leonardo O. Iheme, Devrim Ünay, Melek Kandemir, Jurgen Fripp, Olivier Salvado, Pierrick Bourgeat, Martin Reuter, Bruce Fischl, Marco Lorenzi, Giovanni B. Frisoni, Xavier Pennec, Ronald K. Pierson, Jeffrey L. Gunter, Matthew L. Senjem, Clifford R. Jack, Nicolas Guizard, Vladimir S. Fonov, D. Louis Collins, Marc Modat, M. Jorge Cardoso, Kelvin K. Leung, Hongzhi Wang, Sandhitsu R. Das, Paul A. Yushkevich, Ian B. Malone, Nick C. Fox, Jonathan M. Schott, and Sebastien Ourselin. Assessing atrophy measurement techniques in dementia: Results from the MIRIAD atrophy challenge. *NeuroImage*, 123:149–164, December 2015. doi: 10.1016/j.neuroimage.2015.07.087.

[38] Nathalie Tzourio-Mazoyer, Brigitte Landeau, Dimitri Papathanassiou, Fabrice Crivello, Olivier Etard, Nicolas Delcroix, Bernard Mazoyer, and Marc 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.

[39] Marc-Michel Rohé, Maxime Sermesant, and Xavier 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.

# 8 Appendix

## 8.1 Validation on a synthetic datasets

The regional values of atrophy set are given in Table 2. For each subject, the values are sampled around these means with a 5% standard deviation.

The comparison with the values obtained after simulation or estimated through registration are shown in Figure 15. 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.

## 8.2 Bias on the estimated template trajectories

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

| 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 |
| enthorinal cortex | 6.5 | 0.7 |
| temporal poles | 6.2 | 0.6 |
| other areas | 0.0 | 0.0 |

Table 2: Specified mean regional atrophy for the healthy and the pathological evolutions. It gives simple but realistic atrophy patterns. It should be noted that these are the values of the divergence of the SVF in the area and not the volume changes. In order to compute the volume changes, the integration must follow the the boundaring shifting over time.

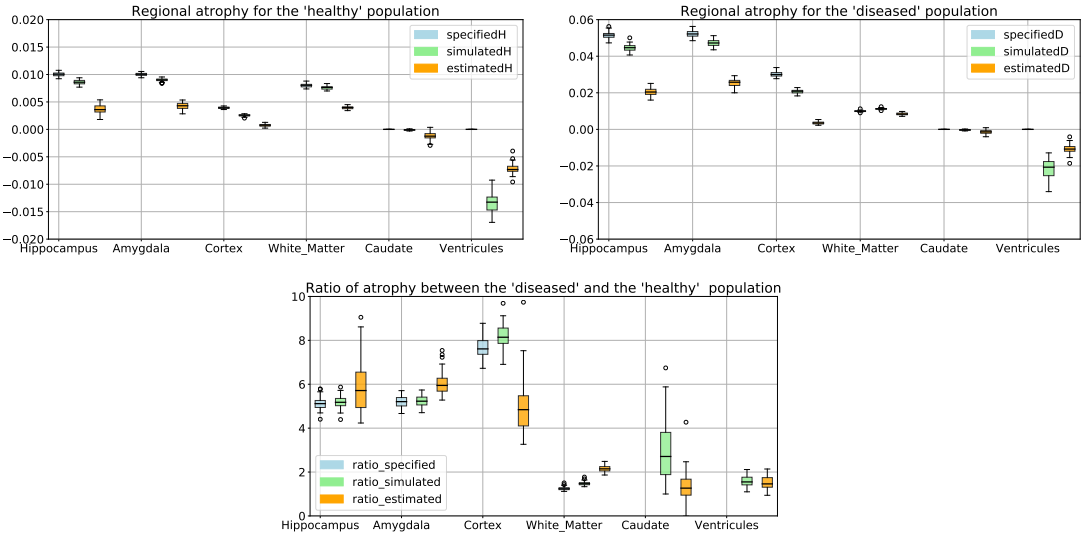



Figure 15: Comparison of prescribed, simulated and estimated atrophy values for several brain regions. The difference between prescription and simulation can be explain by numerical approximation in the biophysical model and the 5 years extrapolation of the deformation. The estimation is biased due to spatial regularization but the relative changes seem to be similar between the two populations.

For the normal aging trajectory we have:

$$
\begin{aligned}
E(\|\hat{\mathrm{v}}_A\|^2) &= \frac{1}{|\mathcal{G}_h|^2} E(< \sum_{k\in\mathcal{G}_h} v^k | \sum_{k\in\mathcal{G}_h} v^k >) \\
&= \frac{1}{|\mathcal{G}_h|^2} \sum_{i,j\in\mathcal{G}_h} E(< s^i_{MA}\mathrm{v}_A + s^i_{DS}\mathrm{v}_D + v^i_r | s^j_{MA}\mathrm{v}_A + s^j_{DS}\mathrm{v}_D + v^j_r >) \\
&= \frac{1}{|\mathcal{G}_h|^2} \sum_{i,j\in\mathcal{G}_h} E(s^i_{MA}s^j_{MA})\|\mathrm{v}_A\|^2 + E(s^i_{DS}s^j_{DS})\|\mathrm{v}_D\|^2 + E(s^i_{MA}s^j_{DS} + s^i_{DS}s^j_{MA}) < \mathrm{v}_A|\mathrm{v}_D > + E(< v^i_r | v^j_r >)
\end{aligned}
$$

Assuming that the subjects are independant and identically distributed:

$$E(\|\hat{\mathrm{v}}_A\|^2) = \|\mathrm{v}_A\|^2 + \frac{1}{|\mathcal{G}_h|^2}\sum_{i\in\mathcal{G}_h}\mathrm{Var}(s^i_{MA})\|\mathrm{v}_A\|^2 + \mathrm{Var}(s^i_{DS})\|\mathrm{v}_D\|^2 + 2E(s^i_{MA}s^i_{DS}) < \mathrm{v}_A|\mathrm{v}_D > + E(\|v^i_r\|^2)$$
$$= \left(1 + \frac{\mathrm{Var}(s^h_{MA})}{|\mathcal{G}_h|}\right)\|\mathrm{v}_A\|^2 + \frac{\mathrm{Var}(s^h_{DS})}{|\mathcal{G}_h|}\|\mathrm{v}_D\|^2 + \frac{E(\|v^h_r\|^2)}{|\mathcal{G}_h|}$$

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

$$E(\|\hat{\mathrm{v}}_D\|^2) = E\left(\|\frac{1}{|\mathcal{G}_{ad}|}\sum_{k\in\mathcal{G}_{ad}} v^k - \hat{\mathrm{v}}_A\|^2\right)$$
$$= E\left(\|\frac{1}{|\mathcal{G}_{ad}|}\sum_{k\in\mathcal{G}_{ad}} v^k\|^2\right) + E(\|\hat{\mathrm{v}}_A\|^2) - 2E\left(< \frac{1}{|\mathcal{G}_{ad}|}\sum_{k\in\mathcal{G}_{ad}} v^k|\hat{\mathrm{v}}_A >\right)$$

We develop the expression of $\|\sum_{k\in\mathcal{G}_{ad}} v^k\|^2$ and $< \sum_{k\in\mathcal{G}_{ad}} v^k|\sum_{k\in\mathcal{G}_{bn}} v^k >$. It leads us to:

$$E(\|\hat{\mathrm{v}}_D\|^2) = \left(1 + \frac{\mathrm{Var}(s^{ad}_{MA})}{|\mathcal{G}_{ad}|}\right)\|\mathrm{v}_A\|^2 + \left(1 + \frac{\mathrm{Var}(s^{ad}_{DS})}{|\mathcal{G}_{ad}|}\right)\|\mathrm{v}_D\|^2 + 2 < \mathrm{v}_A|\mathrm{v}_D > + \frac{E(\|v^{ad}_r\|^2)}{|\mathcal{G}_{ad}|}$$
$$+ E(\|\hat{\mathrm{v}}_A\|^2) - 2(\|\mathrm{v}_A\|^2 + < \mathrm{v}_A|\mathrm{v}_D >)$$
$$= \left(\frac{\mathrm{Var}(s^{ad}_{MA})}{|\mathcal{G}_{ad}|} + \frac{\mathrm{Var}(s^h_{MA})}{|\mathcal{G}_h|}\right)\|\mathrm{v}_A\|^2 + \left(1 + \frac{\mathrm{Var}(s^{ad}_{DS})}{|\mathcal{G}_{ad}|} + \frac{\mathrm{Var}(s^h_{DS})}{|\mathcal{G}_h|}\right)\|\mathrm{v}_D\|^2 + \frac{E(\|v^{ad}_r\|^2)}{|\mathcal{G}_{ad}|} + \frac{E(\|v^h_r\|^2)}{|\mathcal{G}_h|}$$

The variances are estimated in the train population. We should however note that we use the subjects from the $\mathcal{G}_h$ to estimate the template anatomy $T_0$ and by consequence we are certainly underestimating the bias coming from the intra-subject morphological variability for the this population. In practice, using our training dataset we estimated that $\|\hat{\mathrm{v}}_A\|^2 \approx 1.65\|\mathrm{v}_A\|^2$ and $\|\hat{\mathrm{v}}_D\|^2 \approx 1.21\|\mathrm{v}_D\|^2$.