A non-conservative Lagrangian framework for statistical fluid registration - SAFIRA

Caroline C. Brun, Natasha Leporé, Xavier Pennec, Yi-Yu Chou, Agatha D. Lee, Greig de Zubicaray, Katie McMahon, Margaret J. Wright, James C. Gee, Paul M. Thompson

***Abstract*—In this paper, we used a non-conservative Lagrangian mechanics approach to formulate a new statistical algorithm for fluid registration of 3D brain images. This algorithm is named SAFIRA, acronym for Statistically-Assisted Fluid Image Registration Algorithm. A non-statistical version of this algorithm was implemented [9], where the deformation was regularized by penalizing deviations from a zero rate of strain. In [9], the terms regularizing the deformation included the covariance of the deformation matrices (Σ) and the vector fields (q). Here we used a Lagrangian framework to re-formulate this algorithm, showing that the regularizing terms essentially allow non-conservative work to occur during the flow. Given 3D brain images from a group of subjects, vector fields and their corresponding deformation matrices are computed in a first round of registrations using the non-statistical implementation. Covariance matrices for both the deformation matrices and the vector fields are then obtained and incorporated (separately or jointly) in the non-conservative terms, creating four versions of SAFIRA. We evaluated and compared our algorithms' performance on 92 3D brain scans from healthy monozygotic and dizygotic twins; 2D validations are also shown for corpus callosum shapes delineated at midline in the same subjects. After preliminary tests to demonstrate each method, we compared their detection power using tensor-based morphometry (TBM), a technique to analyze local volumetric differences in brain structure. We compared the accuracy of each algorithm variant using various statistical metrics derived from the images and deformation fields. All these tests were also run with a traditional fluid method, which has been quite widely used in TBM studies. The versions incorporating vector-based empirical statistics on brain variation were consistently more accurate than their counterparts, when used for automated volumetric quantification in new brain images. This suggests the advantages of this approach for large-scale neuroimaging studies.**

***Index Terms*—Registration, statistics, non-conservative Lagrangian, fluid mechanics, Riemannian framework, tensors, twins**

C. Brun is with the Laboratory of Neuro Imaging, UCLA, School of Medicine, Los Angeles, CA 90095, USA, and the Penn Image Computing and Science Laboratory, University of Pennsylvania, Philadelphia, PA, USA

N. Leporé is with the Children's Hospital, University of South California, Los Angeles, CA 90027, USA.

Y.-Y. Chou, A. Lee, P. Thompson are with the Laboratory of Neuro Imaging, UCLA, School of Medicine, Los Angeles, CA 90095, USA. (see: http://www.loni.ucla.edu).

X. Pennec is part of the Asclepios Research Project, INRIA, 06902 Sophia-Antipolis Cedex, France. (see: http://www-sop.inria.fr/members/Xavier.Pennec/).

G. de Zubicaray, K. McMahon are with the Centre for Magnetic Resonance, University of Queensland, Brisbane, Australia. (see: http://www.cmr.uq.edu.au/).

M. Wright is with the Genetic Epidemiology Lab, QIMR, Brisbane, Australia. (see: http://genepi.qimr.edu.au/).

J. Gee is with the Penn Image Computing and Science Laboratory, University of Pennsylvania, Philadelphia, PA, USA. (see: http://www.picsl.upenn.edu/).

Manuscript received March 10, 2009

I. Introduction

NONLINEAR registration is an image analysis procedure that warps one image or volume onto another, matching the two images using biological or geometrical features present in both images. It is widely used in medical applications, most commonly to align images from multiple subjects or modalities to a common coordinate space, prior to voxel-based statistical analysis [27]. Registration may also be used for computational morphometry, as the applied deformations can be analyzed statistically to study regional volume or shape differences in conditions such as Alzheimer's disease [40], HIV/AIDS [12], [48], blindness [51], neurogenetic disorders such as Williams syndrome [13], or during childhood development [71]. Statistical analysis of image deformations has also been used to study genetic influences on brain structure [10].

Most image registration algorithms select a similarity term (also called a fidelity term or cost function) that compares information in the two images, and a regularizing term (or statistical prior) that prevents tears, shears or holes from appearing in the resulting registered image, or encourages deformations with certain properties [38], [54]. Brain structure varies widely across subjects, and across the human lifespan [67], and some registration methods have been developed to encode information on the natural variability in brain structure [24], [26], [61], [70]. Cortical surface variation is especially complex, and some researchers have used registrations of curves and surface landmarks to build models of cortical variation based on diffeomorphic currents [21], or $6D$ covariance tensors that related vector-valued deformations between pairs of points on the cortex [26].

Other researchers have taken a more formal approach to modeling brain variation using deformable template methods. Grenander's pattern theory [34] considers the variation of a deformable anatomical template to be modeled using a stochastic PDE of the form $Lu = e$, where e is vector-valued noise, u is the $3D$ deformation vector field aligning the template anatomy onto new subjects' scans, and L is a self-adjoint differential operator (such as the Laplacian or Cauchy-Navier elasticity operator) regularizing the deformation. Since L is an infinite-dimensional operator with known eigenfunctions, many approaches have aimed to develop a spectral model of anatomical variation by projecting the template deformations onto the basis of eigenfunctions, assembling statistics of the basis coefficients, and using them to perform statistical

inference regarding group anatomical differences, or as a prior to control the registration of new images [28]–[30]. More recently, these continuum-mechanical registration approaches have been extended to methods that guarantee diffeomorphic warps (smooth mappings with smooth inverses) and develop metrics on the flow fields [1], [3].

Even so, very few widely used registration methods incorporate empirical information on population variability in brain structure; the use of empirical statistics has been advocated many times, but none of the 14 nonlinear registration methods evaluated in [45] uses empirical information on brain variation during the registration process.

In registration, the cost function is commonly taken as a distance between common anatomical or stereotactic landmarks explicitly defined in the images (such as $3D$ points, curves and surfaces) or intensity-based measures over the whole image such as the squared-intensity difference (L^2-norm), cross-correlation or more complex metrics derived from information theory, such as the Jensen-Rényi divergence [13], [63]. The regularizer compensates for the effects of the data fidelity term and enforces desirable properties in the deformation, such as smoothness, invertibility and inverse-consistency [11], [17], [77].

One method to account for these properties is to add the variational derivative, or gradient, of the similarity criterion as a distributed force field (also known as a body force) in the mechanical equations that govern elastic [6] (through Hooke's Law) or viscous fluid motions (via the Navier-Stokes equation) [8], [16], [33], [49]. Other non-physical regularization models such as Gaussian filtering have been implemented, as these tend to be more efficient than the filters that are needed to implement continuum-mechanical operators [75], [76].

Alternatively, some algorithms focus on formally guaranteeing specific axiomatic properties of the deformation, enforcing for instance diffeomorphic trajectories for the mapping [37], [55]. All these registration processes have different levels of efficiency and precision depending on the accuracy and speed requirements [45].

Very efficient implementations are crucial when real-time registration is required such as in surgical procedures, but longer runtimes may be acceptable for alignment of functional and structural images in research studies of large populations. For instance, registration methods are commonly used in neuroimaging for population studies of brain structure, and for computational morphometry. These studies typically require precise registration to map the influence of disease, genetics or normal development throughout the brain.

In the registration methods mentioned so far, a field of $2D$ or $3D$ displacement vectors is usually computed based on univariate data (one value per voxel) or from pre-defined landmarks. Consequently, as the information that can be consistently identified is limited, a realistic model is needed to interpolate the deformation to the rest of the brain. As noted by other proponents [29], if such empirical information on brain variation was added in the registration process, the statistics of the modeled differences in brain structure would better match those that truly occur, possibly improving registration accuracy, stability, and convergence. Some approaches have used *ad hoc* methods to incorporate regionally-adaptive statistics into the registration. For instance, in [19], the authors nonlinearly rescaled the statistics of the strain tensors so that they could be used in a Demons-like registration algorithm. In [52], [68], non-stationary Gaussian filtering was used to take into account tissue type information during the registration. Early studies by Gee and colleagues also suggested that principal component analysis of intra-subject registration fields could be used to develop an empirical model of brain shape, for use as a Bayesian prior to constrain registrations [29]. Parallel work on active shape models by Cootes, Taylor and others suggested that deformable segmentation methods, for extracting and tracking structures, benefited from prior empirical knowledge of the covariance structure of the deformations [20]. Miller and Grenander also pioneered a spectral method in which the deformations were projected onto the eigenfunctions of the operator governing the deformation, and statistical analysis of the resulting coefficients could be used to infer anatomical abnormalities [34].

In this paper, we focused on fluid registration as it overcomes several known limitations of some continuum-mechanical elastic models, which are derived under small deformation assumptions and the resulting mappings may not be invertible if large image transformations are needed [15]. Fluid transformations remain diffeomorphic even for large deformations. We introduce a new Lagrangian approach where different types of (vector and tensor) statistics on the expected deformations can be seamlessly taken into account in the registration.

The use of a prior to describe the anatomical variations in a population relies on some assumptions. Clearly, it assumes that the brains being analyzed are representative of the population used to create the prior; a prior encoding the variations typically observed in normal subjects may not be optimal for registration of data from subjects with gross lesions (such as tumors) or with anatomical distortions (such as atrophy). Even so, the methods here generalize to creating an atlas or prior from mixed populations of patients and controls. Likewise, a prior model of anatomical variation based on adult brains may not be optimal for analyzing data from children or infants, but the methods proposed here could be used to generate a prior from pediatric data. In other words, our approach assumes that a dataset is available that is representative of the scans that will be analyzed. In terms of formal assumptions for building a prior, there is no model or physical law determining a priori how one brain deforms to match another brain. This is partly because there is no physical or biological process that transforms one brain onto another, unless the images are taken from the same subject over time. In the case of intra-subject registration, some authors have proposed generative priors based on the supposed biological causes of the growth process [35], [36]; [58]. But in the case of registration across subjects, the prior encodes empirical statistics on the positions of points that match across subjects, and their covariances. There has been substantial work on understanding the patterns of anatomical variance and covariances (i.e., spatial correlations) in human populations. The variability of anatomy is not uniform and has preferred directions [69] [26], so this

motivates a locally varying prior rather than one based on a stationary (spatially homogeneous) differential operator such as the Laplacian. Also, the pattern of deformation within a population does not follow an identical and independent law on all deformation parameters (in our case, the vector fields or their deformation tensors), and the covariance tensors of corresponding points across a population are not spherical [26]. Consequently, the standard regularization, which treats deformations in each direction equivalently, is not optimal. For instance, one can imagine that the deformation rate should not be constrained so much in highly variable areas (e.g., temporo-parietal areas of the cortical) compared to more anatomically invariant areas, such as the central sulcus. Our prior is just one term in a constrained optimization problem. As such, does not enforce deformations learned from previous datasets, but promotes them. As these deformations are more likely, the chance of proceeding more rapidly and with greater likelihood to an accurate local minimum is higher. If a deformation deviates considerably from those in the learning set, our method would change the trajectory in the transformation space towards the optimum, which might lead to a different result than that obtained via standard regularization. In that case, the question would become: which of the two results is more plausible? By building the prior using a representative dataset, our adapted metric should lead in general to a mapping that is as plausible, or more so, than obtained from a non-statistical prior. Finally, when the algorithm is used to find groupwise differences, the prior should be built from a representative sample of subjects from both groups, to avoid bias.

Practically, one of the main ideas is to rephrase the image registration minimization problem as the evolution of a non-conservative dynamic system (i.e., a system subjected to conservative forces, which derive from a potential and to non-conservative forces, the most common of which is the dissipative force). In this way, the kinetic energy gained by the gradient descent at each step is transferred to become part of the non-conservative forces, depending on the expected deformations. In mechanics, a force is conservative if the work it does in moving a particle between two points is independent of the path taken. Conservative forces can be written as the gradient of a potential while non-conservative forces cannot. In physical systems, non-conservative forces include friction, drag, and other contact forces. We rewrote the Isotropic Riemannian Fluid registration algorithm developed in [9] using the Lagrangian formalism, where the regularizing terms are formally derived from a *Non-conservative Lagrangian*, which is a function that describes the energy of a dynamical system. This concept, introduced by the Italian mathematician Joseph Louis Lagrange, can be used to re-formulate problems in classical mechanics by solving for the trajectories of systems of particles in terms of conservation laws for momentum and energy, and action functionals.

These terms consist of the Riemannian term and the dissipative term and can be modified to include statistical information on the covariances for the deformation tensors (DTs) and the displacement vector fields (VF), respectively. This leads to four different versions of our algorithm, depending on which types of statistical constraints are added: DTs and VFs statistics both at the same time, or just one of these, or neither (the non-statistical version).

As a result, a mechanically meaningful framework is created that allows the incorporation of biological information in the registration process. This new Lagrangian formulation can help in understanding the constraints and forces that the mechanical system experiences (here, the $3D$ volume). To a reader familiar with the concepts of classical and Lagrangian mechanics, the statistical constraints on the deforming system can be represented using dissipative (non-conservative) forces that are not naturally represented by a classic Newtonian approach. In the evolving mechanical system, the action of the similarity term corresponds to a conservative force that drives the system towards the desired minimum. The Riemannian and dissipative terms correspond to non-conservative forces that favor empirically more likely deformations along the trajectory towards that minimum.

One particular application we use to test our algorithm - for which it was primarily designed - is Tensor-Based Morphometry (TBM). TBM is an image analysis method that has gained popularity and has been used successfully to detect morphometric differences associated with HIV/AIDS [48], Williams sydrome [13], Fragile X syndrome [46], schizophrenia [31], and normal brain development [41]. It consists of a registration of all the subjects' MRI images to an atlas followed by a statistical analysis, which aims to find the profile of volume differences between populations (or similarities in the case of a population of twin subjects). Our registration algorithm is well-suited for TBM as in the most general method, statistics of the Jacobian matrices $J = (Id + \nabla q)$ are computed from the displacement fields q [48]. P-value maps are generated to show local volumetric (or surface in $2D$) differences, via the statistical analysis of J, or its determinant, which measures relative volumetric gain or loss. Local shape differences may also be studied using a multivariate statistical analysis of the DT $\Sigma = J^T J = (Id+\nabla q)^T (Id+\nabla q)$. Σ retains not only the local volume differences but also the directional characteristics of differences [48]. As such, there is a consistency between the subsequent statistical analysis of tensors arising from the registration, and the empirical information used to estimate the tensors in the first place.

Here, we start by introducing the Lagrangian, L, and show how its expression varies depending on the type of forces that a given system is subjected to (conservative or non-conservative). The Hamiltonian is also computed, and reflects the energy of the system (Section II). Section III links the Lagrangian theory to our statistical fluid equation. The Lagrangian framework contains an appropriate structure to provide a clear mechanical explanation of SAFIRA. In Section IV, a TBM analysis resulting in heritability measures is performed on both the corpus callosum ($2D$) and cerebrum ($3D$) from a structural brain MRI dataset (from identical and fraternal twin pairs). We also used the $LPBA40$ image dataset to test segmentation and volume quantification accuracy. Tests were run with the different versions of SAFIRA, and compared with a traditional fluid registration algorithm [49]. The latter method was chosen as a basis for comparison as it has been the registration method of choice in several published TBM

analyses. These tests of registration accuracy and statistical power in anatomical studies are used to validate our method. Sections V and VI present and discuss the results. For TBM, in both 2D and 3D, the biological findings were consistent for all algorithms, but the method incorporating vector-based statistics gave more powerful results (greater effect sizes). Based on the tests of volume labeling accuracy (and registration error), the use of vector-based statistics greatly improved the registration accuracy, compared to using tensor-based statistics or no statistics at all.

II. The General Lagrangian Framework

A. Background

As stated in section I, a number of existing methods regard the $3D$ brain image volume as embedded in a deformable continuum-mechanical system: each voxel is seen as a particle of this deforming system. In such cases, its dynamic behavior may be studied using Newtonian mechanics, as in [16], where the displacement of each particle (i.e., the voxels) is constrained by the Navier-Stokes equation for viscous fluid systems. Here, we modify the traditional fluid equation to allow the integration of statistical data computed from the dataset. A consequence of this modification is that the Newton formulation no longer explains the role of the similarity and regularizing terms mechanically speaking. On the contrary, the generalized framework determined by the non-conservative Lagrangian structure is flexible enough to contain the two types of statistical information (obtained from the dataset) proposed here while remaining mechanically meaningful. In fact, SAFIRA is a modified version of a standard fluid equation, whose terms do not all have a Newtonian mechanical equivalent.

B. A generalized Lagrangian

The Lagrangian L can express the dynamic behavior of a system that is subject to conservative forces [32].

$$L = T - V$$

In the conservative case, $T(\dot{q})$ and $V(q)$ represent the kinetic energy and the potential energy of the system, respectively. q is the displacement, and $\dot{q}$ is the velocity of the system. One way to define the Lagrangian L is to examine its integral, called the action, $\mathcal{S}$. $\mathcal{S}$ is defined as $\mathcal{S} = \int_{t_0}^{t_1} L dt$, and the paths followed by a mechanical system between the times t_0 and t_1 are the ones that minimize the action.

In cases where the system is also subjected to non-conservative forces, the definition of $\mathcal{S}$ changes as the work produced by the non-conservative forces, W, must be added to $\mathcal{S}$:

$$\mathcal{S} = \int_{t_0}^{t_1} L + W \; dt \qquad \text{with} \qquad \delta W = \vec{F}.\delta\vec{r}$$

δW is the work created by the non-conservative forces $\vec{F}$ during the virtual displacement $\delta\vec{r}$. $\delta\vec{r}$ is a variation associated with the possible body position $\vec{r}$, and not with the actual solution $\vec{r}(t)$. $\vec{r}$ is thus chosen such that the force $\vec{F}$ remains constant during the displacement $\delta\vec{r}$ and depends only on q. In fact, by definition a virtual displacement is a displacement that occurs during an infinitesimal time t and that agrees with the constraints of the system. It can also be seen as the difference between two permissible but unequal displacements taken over the same time interval of time (t to $t + dt$) (see Figure 1). It is always orthogonal to the forces constraining it (an example of constraint force may be illustrated by a bead that is forced to slide on a ring) [32], [60]. To find the path followed by the dynamic system (i.e., the path that minimizes the action), $\delta\mathcal{S}$ can be derived as

$$\delta\mathcal{S} = \int_{t_0}^{t_1} \delta L + \delta W \; dt = 0$$

Hence,

$$\left(\frac{\partial L}{\partial q}\right) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) + \vec{F}\left(\frac{\partial \vec{r}}{\partial q}\right) = 0; \tag{1}$$

(see Appendix A and [74] for further explanation). This dynamic equation defines the movement of a non-conservative system at each time t.

C. A detour through Hamiltonian mechanics: conservation of energy?

Using a Lagrangian structure makes each term easier to interpret; it also facilitates the computation of the Hamiltonian H, a quantity that, among other things, characterizes the energy conservation of the system with time. While the energy of conservative systems is maintained with t, this is no longer the case when non-conservative forces are added. Within the context of registration, summarizing the transfers between the different types of energies is of considerable interest, as it gives a clear understanding of the energy minimization scheme.

The Hamiltonian H represents the energy of the conservative system and may be derived from the conservative Lagrangian L as

$$H = p\dot{q} - L \qquad \text{with} \qquad p = \left(\frac{\partial L}{\partial q}\right)_{q\dot{q}} \tag{2}$$

where p is the momentum of the system and q is the displacement. The variation of H w.r.t. time indicates if the energy is conserved, and if not, where it is transferred to. For NC systems, we obtain

$$\frac{dH}{dt} = \vec{F}.\left(\frac{\partial \vec{r}}{\partial q}\right)\dot{q} \tag{3}$$

(see Appendix B for the complete derivation). The energy H of the NC system is not constant with time: this equation shows that the kinetic and potential energies lost by the system are transferred to the NC terms.

III. SAFIRA: An adaptive Fluid Registration Algorithm

A. Previous work

1) The Riemannian Elastic energy: In [57], Pennec introduced a new elastic registration algorithm, which replaces the commonly used Euclidean metric for the elastic regularizer by

a more suitable Riemannian one. More specifically, he defined a statistical regularizer using a Mahalanobis distance on the space of DTs Σ ($\Sigma = J^T J = (Id + \nabla q)^T (Id + \nabla q)$, with J and q the Jacobian matrix and the displacement, respectively. The Σ's are considered as random variables in the Riemannian space of DTs). The Σ's are matrices defined at each voxel that characterize the distortion and volume change of each voxel from the registration. As they are symmetric positive-definite matrices and form a convex half-cone in the vector space of 3x3 matrices, standard Euclidean operations are extrinsic to the manifold that they form; hence the need for a Riemannian (intrinsic) metric in the regularizer. In [57], this metric was determined through the Log-Euclidean framework [2] as we describe below; this formalism allows simple computations to be performed intrinsically on the manifold.

In standard elastic registration, each voxel is considered as a particle whose movement is controlled by Hooke's law and follows an equation that is derived from the Saint-Venant Kirchoff elastic energy:

$$Reg_{SVKE}(\vec{q}) = \int \frac{\mu}{4} Tr((\Sigma - Id)^2) + \frac{\lambda}{8} Tr(\Sigma - Id)^2$$

where λ and μ are the Lamé coefficients. When $Reg_{SVKE}(\vec{q})$ is redefined in a Riemannian framework, the elastic energy becomes [57]:

$$\begin{aligned} Reg_{RE}(\vec{q}) &= \frac{1}{4} dist^2_{Log}(\Sigma, Id) && (4) \\ &= \frac{1}{4} dist^2_{Eucl}(\log(\Sigma), \log(Id)) && (5) \\ &= \frac{1}{4} \int ||\log(\Sigma)||^2 \end{aligned}$$

A statistical elastic registration can be implemented as follows. The elastic algorithm with Reg_{RE} can be used to run a first round of registration on a dataset. Then, the covariance and the mean of the Σ's can be computed using the Log-Euclidean metrics. Reg_{RE} can then be generalized to its statistical form using Mahalanobis distance on these tensors [57]:

$$Reg_{SRE}(\vec{q}) = \frac{1}{4} \int Vect(\Sigma_L - \bar{\Sigma}_L) Cov^{-1} Vect(\Sigma_L - \bar{\Sigma}_L)^T \quad (6)$$

where $Vect(\Sigma)^T = (\sigma_{11}, \sigma_{22}, \sigma_{33}, \sqrt{2}\sigma_{12}, \sqrt{2}\sigma_{13}, \sqrt{2}\sigma_{23})$, $\Sigma_{Li} = log(\Sigma_i)$, and the means and covariances of the Σ's are represented by $\bar{\Sigma}_L = \frac{1}{N} \sum_i log(\Sigma_i)$ and $Cov = \frac{1}{N} \sum Vect(\Sigma_{Li} - \bar{\Sigma}_L) Vect(\Sigma_{Li} - \bar{\Sigma}_L)^T$, respectively. The new regularizer Reg_{SRE} is used on the original (non-registered) images to perform a second round of registration, this time taking into account known statistical information on the data through the covariances and means.

2) The Fluid formulation: The general problem in image registration is to find the transformation q that minimizes the dissimilarity between the images, usually as part of a compound cost functional that also considers the regularity or smoothness of the deformation VF. Here and in our previous work [9], we use the sum of squared intensity differences (SSD) criterion $Cost = Sim(I, K \circ q) = \int (I(x) - K(q(x))^2 dx$, but any other image similarity criterion might be used. In [9], we used a fluid matching approach - which, unlike the elastic case, allows for large deformations. We combined the fluid and Riemannian frameworks to create a so-called isotropic Riemannian fluid registration algorithm. At each voxel, for each of the time steps Δt, the regularizer and image similarity cost terms were optimized to find the velocity $\dot{q}$ according to the equation [1]:

$$\frac{d\dot{q}(x,t)}{dt} = \nabla_q Cost - \alpha \nabla_{\dot{q}} Reg_{Riem}(\dot{q}) - \beta \dot{q} \quad (9)$$

α and β are the weights of the regularizing terms with regard to the cost function. $\dot{q}$ is obtained from equation (9) at each time step Δt and integrated over time to find the displacement q. Following eq. (4), the fluid Riemannian regularizing term is

$$Reg_{Riem}(\dot{q}, t) = \int \frac{\mu}{4} Tr(\log(\Sigma_{\dot{q}}^2)) + \frac{\lambda}{8} Tr(\log(\Sigma_{\dot{q}}))^2 \quad (10)$$

with $\Sigma_{\dot{q}} = (\nabla \vec{\dot{q}} + Id)^T (\nabla \vec{\dot{q}} + Id)$.

Reg_{Riem} constrains the deformation of one image into another by acting on the rate of strain $\Sigma_{\dot{q}}$ rather than on the DT Σ (as was the case in equation (4)). The introduction of Reg_{Riem} was a first step toward the implementation of a statistical fluid algorithm, as the aim was to integrate statistical information on the Σ's within this regularizing term.

B. Redefining the fluid algorithm with the Lagrangian formulation

In this section, we use the Lagrangian theory of section (II) to reformulate the isotropic Riemannian fluid registration algorithm (equation 9). In particular, given the definition of the kinetic energy and acknowledging the conservative properties of the similarity term (see paragraph III-A2), the different terms of this modified Navier-Stokes equation (9) can be defined as follows:

- Kinetic energy: $T = \frac{1}{2} ||\dot{q}_j||_2^2$
- Potential Energy: $V = Cost(q)$
- Nonconservative energy:

$$V_{\vec{F}} = V_{\vec{F}1} + V_{\vec{F}2} = \frac{1}{2} \beta ||\dot{q}||_2^2 + \alpha Reg_{Riem}(\dot{q})$$

In this case, following the expression of the dynamic Lagrangian equation (equation (1)), equation (9) may be rewritten with the following equalities:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = \frac{d\dot{q}(q,t)}{dt} = \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}}\left(\frac{1}{2}||\dot{q}||_2^2\right)\right) \quad (11)$$

[1]Note: As in [16], we assume small deformations at each time step, thus the kinematic nonlinearities of the VF q are not taken into account. Consequently, we have

$$\frac{d\vec{q}(x,t)}{dt} = \frac{\partial \vec{q}}{\partial t} + \sum \dot{q} \frac{\partial \vec{q}}{\partial x} \approx \frac{\partial \vec{q}}{\partial t} \quad (7)$$

Similarly, we will consider that

$$\frac{d\dot{q}(x,t)}{dt} \approx \frac{\partial \dot{q}}{\partial t} \quad (8)$$

$$\frac{\partial L}{\partial q} = \nabla_q Cost(I, K, q) \tag{12}$$

$$\vec{F}\left(\frac{\partial \vec{r}}{\partial q}\right) = \alpha \nabla_{\dot{q}} Reg_{Riem}(\dot{q}) + \beta \dot{q} \tag{13}$$

C. Incorporating statistics on the deformation matrices

Given a dataset, we execute a first round of registrations to compute the statistics that will be needed for the statistical regularization, i.e., that will be incorporated in the regularizing (non-conservative) terms. For each image from this dataset, we apply the non-statistical algorithm (see equation (9)) to obtain a distribution of vector fields, from which we compute the covariance of the DTs, Σ, and the covariance of the displacement fields, $\vec{q}$. According to the elastic version (see equation (6)), the first non-conservative regularizing term can be defined as:

$$Reg_{stat} = \frac{1}{4}\int Vect(W_{\dot{q}} - \bar{W}_{\dot{q}}) Cov^{-1} Vect(W_{\dot{q}} - \bar{W}_{\dot{q}})^T \tag{14}$$

Here $W_{\dot{q}} = log(\Sigma_{\dot{q}})$, and to avoid any bias, we choose to keep the average rate of strain $\bar{W}_{\dot{q}} = \frac{1}{N}\sum_i log(\Sigma_{\dot{q}_i})$ equal to zero at all times. The covariance Cov is computed using DTs Σ_q. Using $\Sigma_{\dot{q}}$ would be equivalent because q and $\dot{q}$ are collinear to each other. This new form Reg_{stat} can replace Reg_{Riem} in $V_{\vec{F}}$.

D. Incorporating statistics on the displacement fields

The other non-conservative term can also be modified to embody the covariance of the displacements. The Euclidean norm $||.||_2$ is replaced by a Mahalanobis distance in $V_{\vec{F}1}$:

$$V_{\vec{F}1} = ||\dot{q}_j||^2 = \dot{q}_j^T cov_{q_j}^{-1} \dot{q}_j \tag{15}$$

with $cov_{q_j} = \frac{1}{N}\sum_i (q_i - \bar{q}_j)^T (q_i - \bar{q}_j)$, the covariance of the displacements q at a voxel j across the images i. Strictly speaking, this non-conservative term may be interpreted as a Rayleigh dissipation term, as explained in Appendix C. In fact, it is proportional to the quadratic velocity.

The algorithm can include both of the previous types of statistical data (vector- + tensor-based statistical version), the information on the Σ's only (tensor-based version), information on the displacement fields q only (vector-based version) or neither of them (non-statistical version of the algorithm).

E. Conclusion on the Lagrangian Structure

The Lagrangian definition is well suited for solving the minimization problem, and is used to find $\dot{q}$ at each time step. To transform the registration problem into a dynamical system, a gradient descent is performed to find the optimal registration parameters. For instance in [49], this can be achieved by dissipating the kinetic energy acquired at each step according to the Navier-Stokes equation. In our algorithm, we dissipate more energy when we are going into an expected direction (that agrees with the prior statistical information) or retain the energy when we are going in unlikely directions. Practically, this is possible because the regularizing term V_{F1} and Reg_{stat} can be modulated by the prior information. They can either be enhanced (relative to what they would be without the statistical constraint) if the inverse of the covariance (of the vector fields and diffusion tensors, respectively) is high, or they may decrease if the inverse of the covariance is small. These modulated values help the algorithm to surpass some local minima and keep moving towards a statistically plausible solution.

F. Implementation

The use of Lagrangian mechanics allows a clear explanation of the forces exerted on the mechanical system. While this formulation is fundamental for a complete understanding of our algorithm, we go back to Newton equations to explain the resolution of the problem. In fact, our final goal is to find the displacement q and thus the velocity $\dot{q}$ of each particle at each time step δt. We use a multiresolution algorithm. As a first approximation, the second-order terms are neglected. All the first-order derivatives are computed using finite differences. Thus equation (9) is modified into:

$$\nabla_q Cost - \alpha \nabla_{\dot{q}} Reg(\dot{q}) - \beta \hat{\dot{q}} = 0$$

Reg can either be the Riemmanian regularizing term (equation (10)) or the statistical one (see III-C). Similarly, $\beta\hat{\dot{q}}$ is either the velocity (as presented in the initial equation 9) or the gradient of the Mahalanobis distance (see III-D). The regularizing part of this equation contains non-linear terms, hence the computation of $\dot{q}$ using a gradient descent method based on *Levenberg-Marquardt* optimization. At each time δt, the cost function is calculated between the moving image and the template. The velocity is found and integrated to find the displacement. A supplementary step is needed to prevent singularities. If the Jacobian falls below a threshold (here 0.5), a regridding step is performed [16]. The algorithm is as follows:

1) Define a grid on the template and an initial resolution; initialize $t = 0$ and $\dot{q}(\vec{x}, t = 0) = \vec{0}$
2) Calculate the force, i.e. the gradient of the mean square difference $\nabla_q Cost$ at this given resolution.
3) Solve the PDE to find the velocity at the same resolution, at each point in the grid, using gradient descent (RK4). We chose $\dot{q} = \gamma G \circ \vec{F}$ with $\gamma = 0.3$ and G is a Gaussian function that can be modified according to the resolution. The general slope of the gradient descent is: $S_1 = (\beta\hat{\dot{q}}_n - \vec{F} + \alpha\nabla Reg)$ ($\beta = 1$ and α varies with the resolution). S_i's ($i \in [2:4]$) are computed iteratively. q_{n+1} is computed the following way: $\dot{q}_{n+1} = \dot{q}_n - \frac{1}{6}\epsilon(S_1 + \frac{1}{2}S_2 + \frac{1}{2}S_3 + S_4)$.
4) Find a time step that is consistent with the maximal flow allowed (which is predetermined).
5) Integrate $\dot{q}$ to find q, using an explicit algorithm ($q_{t+\delta t} = q_t + \dot{q}_t.\delta t$)
6) Compute the Jacobian of the displacements. If the Jacobian determinant falls below 0.5, then re-grid the template and return to Step 4.
7) Obtain the new displacement field once the Jacobian value is acceptable.

The gradient descent scheme used here may converge on local minima, especially when the number of degrees of freedom is large, as is the case here. This is a generic problem for all deformable registration algorithms, and the cost functions often have large numbers of local minima. Here, we do not use a traditional gradient descent but an explicit iterative method (RK4, [62]). This was chosen so that the resolution is less sensitive to local minima than a first-order method. Interestingly, as observed below (see Results, paragraph V-B and figure 5), including vector statistics in the deformation reduces the variance of the deformation matrix without affecting registration accuracy (**Figure 4**). Our statistically regularized method promotes solutions with a reasonable magnitude of deformation, and with spatial derivatives typical of smooth anatomical mappings. This avoids very high frequency solutions or local minima that optimize the data fidelity term only.

IV. Evaluation of the algorithm: methodology

A. Data and preprocessing

1) Twin dataset: We scanned 23 pairs of monozygotic (MZ) (11 male and 12 female pairs) and 23 pairs of same-sex dizygotic (DZ) twins (10 male and 13 female pairs), recruited as part of a 5-year research project that will eventually evaluate 1150 twins. Additionally we also scanned one healthy subject with the same protocol to be used as template for the registration. This template image was used strictly as a common space and was not included in the data set for the genetic analysis. The age range for the subjects was $22-25$ years old for all the subjects, including the template (mean age: 23.8 ± 1.8 SD years). Each subject was informed of the goals of the study and signed a formal consent. The study was approved by the appropriate Institutional Review and Research Ethics Boards. Zygosity was established objectively by typing nine independent DNA microsatellite polymorphisms (Polymorphism Information Content > 0.7), using standard polymerase chain reaction (PCR) methods and genotyping. These results were cross-checked with blood group (ABO, MNS and Rh), and phenotypic data (hair, skin and eye color), giving an overall probability of correct zygosity assignment of greater than 99.99%. All subjects underwent physical and psychological screening to exclude cases of pathology known to affect brain structure. None of the twin subjects reported a history of significant head injury, a neurological or psychiatric illness, substance abuse or dependence, or had a first-degree relative with a psychiatric disorder.

2) Image Acquisition and preprocessing: All MR images were collected using a 4 Tesla Bruker Medspec whole body scanner (Bruker Medical, Ettingen, Germany) at the Center for Magnetic Resonance (University of Queensland, Australia). Three-dimensional $T1$-weighted images were acquired with a magnetization prepared rapid gradient echo (*MP-RAGE*) sequence to resolve anatomy at high resolution. Acquisition parameters were: inversion time (TI) /repetition time (TR) /echo time (TE) $= 1500/2500/3.83$ $msec$; flip angle $= 15^o$; slice thickness $= 0.9$ mm with a 256x256x256 acquisition matrix. Extracerebral (non-brain) tissues were manually deleted from the MRI images using the 3D interactive program, Display (Montreal Neurological Institute, McGill University, Canada). All scans were then aligned to the standard Colin27 brain template using a 9-parameter global registration (i.e., translational and rotational alignment, allowing scaling in 3 independent directions), which may be found in the *FMRIB*'s Linear Image Registration Toolbox, *FLIRT* [44]. For all subjects, corpus callosum outlines were manually traced in the mid-sagittal plane, using BrainSuite [65] and rigidly aligned to the target corpus callosum (using 2 translations, 2 rotations, and scaling in 2 independent directions), using a least-squares cost function on the binary mask data.

B. Data analysis

All subjects' 3D scans and $2D$ binary corpus callosum images were non-linearly registered using SAFIRA's four versions as well as the traditional fluid method [49] ($3D$ only). In each case, VFs, and their corresponding Jacobian matrices J were computed at each voxel, resulting in a scalar value for the Jacobian determinant, $det(J)$. The $det(J)$'s express the local differences in volume ($3D$) or area ($2D$) between each subject and the target image: $detJ(\vec{u}) > 1$ indicates a local excess in the image being studied in comparison to the template while $detJ(\vec{u}) < 1$ indicates a local deficit.

C. Accuracy of volume quantification

As our algorithm was primarily developed to study volume and shape differences between subjects and groups in morphological studies such as TBM, we first estimated how the incorporation of different types of statistics influenced the results with regard to volume estimation. We based this analysis on the method developed by Klein et al. in their validation study [45], which used the $LPBA40$ brain MRI dataset, which is based on data from 40 healthy volunteers (20 men and women) [66]. We randomly chose a subject from this $LPBA40$ dataset as a template and registered all the other MRI scans to this template using the 3 independent versions of our algorithm (non-statistical, tensor-based and vector-based) in order to see the independent effect of each statistical prior on the accuracy of the registration. Then, we applied the VF, obtained from each subject's structural MRI registration to the template, to the corresponding labeled image (for each of the subjects, 56 structures (or regions of interest) were delineated manually - see http://www.loni.ucla.edu/Atlases/LPBA40 for more details). Each registered labeled image was compared to the manually segmented labeled template, which served as the ground truth segmentation. The volume differences between the template T_r and each subject S_r were reported for each region and summed across the population. The volume similarity coefficient V_s was defined as follows:

$$V_s = 2\frac{\sum(|S_r - T_r|)}{\sum(|S_r| + |T_r|)}$$

For this measure, smaller values denote a more accurate quantification of substructure volumes. We repeated this test using the traditional fluid algorithm.

D. Measure of the smoothness and regularity of the deformation

In general, if two registration algorithms match the features in two images equally well, the one that produces the smoother (regular) deformation is usually considered better, as smoother deformations are usually more biologically plausible. Secondly, a smoother deformation usually requires fewer degrees of freedom to model, so, to obtain a more parsimonious model, a less complex transformation is usually preferred over a more high-dimensional one. The smoothness of the deformation can be illustrated for each subject in several ways, by showing deformed grids carried through the transformation or voxelwise maps of $det(J)$. $det(J)$ shows the amount of deformation when registering each image to the template, while the deformation grid gives an indication of the level of regularization. More complicated measures, such as the voxel-wise norm of the deformation matrix Σ can be determined:

$$N_\Sigma = \sqrt{(trace(log(\Sigma)^2))}$$

This quantity measures the smoothness of the deformation between a subject's image and the template in the log-Euclidean framework [2]. The voxel-wise variance can be averaged across the whole population (composed of one twin per pair randomly selected) to show the variability the registration mappings.

E. Genetic influence on brain structure

The statistical power of the different versions of SAFIRA and the traditional fluid, for use in a morphometry study, was compared by computing genetic measures from the dataset. To measure the resemblance between twin pairs, we first computed the intraclass correlation coefficient (ICC) for both the MZ and the DZ groups in the cerebrum and the corpus callosum, according to the equation:

$$ICC = \frac{\sigma_b^2}{(\sigma_b^2 + \sigma_w^2)} \tag{16}$$

σ_b^2 is the pooled variance between pairs and σ_w^2 is the variance within pairs [64].

These ICC measures were computed from the Jacobian determinant at each voxel in the registered maps, which is an index of the regional volume of specific structures, relative to the standard template. As such, the meaning of the ICC is just a map of how similar brain structure volumes are between twins of various kinds, with higher values meaning that volumes are more correlated across members of a twin pair.

Heritability is an estimate of the proportion of the variation in a measurement that is attributable to genetic differences among individuals. We computed Falconer's heritability statistic, h^2, defined as twice the difference in correlation between MZ and DZ pairs

$$h^2 = 2(r(MZ) - r(DZ))$$

where $r(MZ)$ and $r(DZ)$ are the intraclass correlation values for the MZ and DZ groups, respectively [23]. This is a fairly standard measure of heritability, although it may be estimated using other methods, such as structural equation models and path analysis, which we have used in other reports [10], [14].

We did not want to assume that the data, $det(J)$, was normally distributed across subjects, so we computed p-values at each voxel with a voxelwise permutation test, to establish a null distribution for the ICC statistics at each voxel [56]. The null hypothesis for the intraclass correlation was $ICC = 0$ (no correlation). At each permutation, every subject's scalar value, $det(J)$, is randomly assigned to another subject; the r-values computed from this randomly-generated distribution are compared to the r-values for our data; the resulting permutation-based (non-parametric) p-value is defined as the quantile of the empirical null distribution where the real data falls. To control the standard error of p, we performed 5000 permutations at each voxel [22].

F. Definition of the distances between images

We also defined four metrics to define a global (overall) distance between one image and its target, based on the deformation mappings. Two of the distances were based on displacements and the two others on DTs.

- Distance on q:

$$d_{q1} = \int_{image} ||\vec{q_j}||^2,$$

- Statistical distance on u:

$$d_{q2} = \int_{image} Vect(\vec{q_j}) cov_{q_j}^{-1} Vect(\vec{q_j})^T,$$

- Distance on the Σ's (defined using the Log-Euclidean framework [2]):

$$d_{s1} = \int_{image} N_\Sigma^2,$$

- Statistical distance on the Σ's equivalent to the energy Reg_{stat}:

$$d_{s2} = \int_{image} Vect(W_q - \bar{W}_q) Cov^{-1} Vect(W_q - \bar{W}_q)^T$$

These distances were computed on the binary $2D$ images for all the MZ twin pairs. From these single values per subject, for each algorithm, and each distance, we compute the ICC for the MZ group. The corpus callosum is a highly heritable subcortical structure of the brain [43], so the premise of this approach is that the ICC should be high for all these distances, and that any registration error should tend to deplete the correlation between identical twins. As such, in the absence of ground truth on what the deformation should be, the ICC is a fair measure of the registration accuracy, offering a useful but not sufficient criterion for good registration.

A subset of the monozygotic twins were then fluidly registered to their twin sibling and to the rest of the population. For the four algorithms, the obtained vector ($\vec{q}$) and DT (Σ) fields were input into the distances.

V. Results

A. Preliminary results

Figure 2 illustrates the influence of the variations of the non-conservative terms on the deformation. A circle is transformed into an ellipse using the non-statistical version of the $2D$ registration algorithm (green grid) (see eq. 9). The same transformation was made with the same number of iterations (time steps) with a strong dissipation and with a small dissipation (corresponding to small and large covariance matrices on the displacement incorporated into $V_{\vec{F}1}$). These two cases are represented by the red grid (small dissipation) and the blue grid (strong dissipation). For an equivalent number of time steps, the greater the dissipation, the slower the progression of the transformation. In the case of biological data, the dissipation decreases as the global value of the covariance matrices computed from the displacement fields increases. For local areas where the dissipation is stronger (and the covariance smaller) in the image deforms less than the other areas within a time step. This is logical, because a small covariance matrix means that the anatomical variation is small across subjects and the target structure is expected to be found within a small deformation distance of the starting position. The tensor-based version follows the same kind of behavior. As the global value of the covariance matrices computed from the Σ's increases, the progression of the transformation increases.

Figure 3 compares the regularity of the deformation, obtained from the transformation of one corpus callosum into the template, using the four different versions of the algorithm. The smoothness of the deformation is also given by $det(\Sigma)$ maps. When prior information on the displacements (vector-based statistics) is incorporated in the registration, smoother transformations tend to occur. The Jacobian determinant map for the non-statistical version (*bottom - left*) shows that without any *a priori* information, the larger deformations are concentrated around the edges of the structures, especially where there are irregularities (such as in the isthmus and splenium). This pattern is observable in all cases. Even so, the effects are much more localized with the vector-based statistics algorithm (*bottom - middle left*) whereas they propagate to the whole white matter structure when the tensor-based statistics technique is used (*bottom - middle right*). Introducing both statistics combines the influences of the two previous methods (*bottom - far right*). The dissipation is stronger with the vector-based statistics where the covariance of the displacement field is smaller (i.e., inside the structure). This smoother effect is not seen with the tensor-based prior.

B. Estimating the volume conservation

As explained in paragraph (IV-C), the volumetric quantification error was computed across the $LPBA40$ population for each of the 56 delineated labels (see **Figure 4**). Our algorithm was primarily developed for use in tensor-based morphometry studies, the deformation of one brain volume onto another is essentially measuring volume differences between structures, so its accuracy in labeling subvolumes of the brain can be evaluated using manually labeled standard brain datasets. In other words, the differences between the manually labeled volumes and those obtained by deforming a labeled template onto them can be regarded as a measure of registration accuracy.

Figure 4 shows the volume quantification error, V_s, for all 56 regions of interest. Blue colors indicate no volume difference between the registered label and the manually defined ground truth label, whereas red colors indicate a large difference between the volume of the deformed segmentation and the manually defined ground truth. The results are shown for three versions of the algorithm (non-statistical, and the versions using vector-based and tensor-based statistics) and for the traditional fluid. Overall, incorporating vector-based statistics on the deformation field during the registration improves volumetric matching, and improves the accuracy of volume quantification. In particular, this is especially clear for the subcortical gray matter structures, such as the caudate and the putamen (near the bottom of Figure 3). In those cases, the vector-based statistical algorithm was more accurate for volume quantification, which is advantageous for large-scale volumetric studies.

C. Estimating the smoothness of the deformation

Figure 5 shows the voxel-wise variance of the DTs Σ measured over a population composed of one twin randomly selected per pair, for each version of the algorithm (see paragraph IV-D). Blue indicates a small variance. In all cases where a statistical regularization was applied, the variance of the deformation matrices was smaller than in the non-statistical case. This means that each subject's transformation is closer to the other ones when a statistical prior is used. Whether or not these deformations are closer to the truth depends on whether the population variance in the Jacobian reflects true biological variation or methodological noise. In other words, it is not necessarily better to have a low variance in the Jacobian maps for morphometric studies, it might be advantageous to have a high variance in the Jacobian maps so long as they are really encoding true features of the anatomy. Otherwise, low-dimensional warps would always be favored, when really the optimal dimension and frequency content of the registration field depends on the detail or complexity in the true signal. Even so, these maps do show that the priors do limit the variance in the deformations, and are likely to favor transformations that are more likely to occur.

For the algorithm using vector-based statistics (*middle left*), the covariance of q is smaller in large homogeneous regions, such as the white matter. This is to be expected, as constraining q affects the magnitude of the deformation field, rather than just its derivatives (which can be low even for a large-deformation mapping). Even so, the differences between the non-statistical and the tensor-based methods were less noticeable (*middle right*), as the tensor constraints tend to influence only the smoothness of the deformation (1^{st} order-statistics compared to 0^{th}-order statistics). When the two statistics were included in the transformation (*far right*), the effect of the displacement-based prior information (constraining q) was more influential.

An additional test was performed to better understand the evolution of the covariance of the VFs, when the statistical

approach is iterated until convergence. It is of interest to see the effects a recurring registration scheme on the deformations (see **Figure 6**). We ran a first round of registrations, computed the covariance on the VFs, included this information in a second round of registrations, computed the covariance of these latter VFs and re-injected it again into another round of registrations. The number of iterations was fixed for all registrations. The voxelwise trace of the covariance of the VFs is presented here in a logarithmic scale. Blue colors indicate regions of low covariance and red colors correspond to a high covariance. The first image (top left) represents the covariance of the non-statistical -unconstrained- VFs (i.e., using the standard homogeneous regularizer as the prior).

For each round of registration, the deformation of the image is more constrained in regions of low covariance (see equation 15). This means that the constraint on the magnitude of the deformation will be higher in subcortical regions (**Figure 6**, *top left*) and thus the displacement in these regions will be smaller. This will result in a smaller covariance at the level of these regions and will be echoed in the following round of registration and so forth. In fact, as the covariance becomes smaller, the deformation is more constrained and there is less variability in the resulting vector fields. The outcome of this experiment suggests that repeated rounds of registration, using an updated statistical prior, may not be advantageous, as the damping due to the statistical prior increase over the iterations and drive the solutions towards the identity. Given this, it may be optimal to base the prior on one round of unconstrained registration without statistical damping rather than use the approach recursively.

D. Statistical Analysis

1) Corpus callosum: **Figure 7** shows the results of the Tensor-based Morphometry study of the corpus callosum with each method. The intraclass correlations were computed for both the MZ and DZ groups, as well as their significance and the corresponding heritability. All versions of the algorithm show a similar pattern: the $2D$ structure is shown to be influenced by genetic factors in several regions: the extreme part of the splenium (corresponding to white matter fibers projecting primarily to the occipital lobes [39]), the anterior third (projecting primarily to frontal and prefrontal cortices) and the midbody (projecting to the mid-cingulate cortex and other limbic areas). Intriguingly, the non-statistical version gave slightly more powerful results, in the sense that the effect sizes for intraclass correlations were greater. Both MZ and DZ groups showed higher intraclass correlation and significance (i.e., higher effect sizes at the voxel level) with the non-statistical technique compared to the others. While the pattern presented in the DZ group and its significance are comparable with the first three algorithms (although less strong in the statistical versions), the incorporation of both statistical priors in the registration decreases the signal in the DZ group, removing part of the effects found along the midbody.

2) Cerebrum: The ICC, its significance and the heritability are displayed as 3D maps for the whole cerebrum in **Figures 8** to **10**. Again in the $3D$ case, the anatomical pattern is consistent overall across all methods for the three types of maps. Subcortical structures, white matter and ventricles are shown to be influenced by genetic differences across individuals. However, the effect sizes given by each method follows a different rank order than that seen in the $2D$ case (see V-D1). In both MZ and DZ groups, similar overall patterns of resemblance are mapped by each of the four versions of the algorithm and by the traditional fluid method. Even so, the ICC, its significance and the heritability maps present a more powerful signal when the vector-based algorithm is used in the registration step of the TBM study. In all cases, SAFIRA outperforms the traditional fluid algorithm. To better quantify the difference in power, cumulative distribution functions were plotted for each version in the MZ group (**Figure 11**), based on the pvalues for the ICC. This is also consistent with the previous results; including prior information on the deformation fields increases the detection sensitivity (the black and blue lines are merged together; they lie on top of each other). However, the incorporation of tensor statistics, on their own, very slightly depletes the statistical power to detect differences (yellow line) when compared to the non-statistical version (purple line).

E. Distances

The statistical maps and cumulative distribution functions showed differences in effect sizes between the different versions of the algorithm. We also examined several global distances between twins images (see IV-F), on the premise that any registration errors would tend to diminish the correlation (ICC) between twins.

Table I shows the ICC and its significance for distances defined on the the corpus callosum. As this subcortical structure is known to be under strong genetic control, we estimate the robustness of the algorithms by comparing the ICC and p-values. The higher the ICC the more the registration has picked up the similarities between the images; registration errors tend to deplete the correlation between identical twins. For each of the measures, the more significant statistics were determined (red color values in the table). While none of the algorithms gave statistically significant results with d_{s2} (as all the p-values are > 0.05), the vector-based distances d_{q1} and d_{q2} favored the versions containing the statistical information on the displacement fields. The non-statistical version benefited from d_{s1}. Overall, across the distances, the version incorporating vector statistics gave higher similarities across twins (mean of the ICC).

Figure 12 illustrates the distances measured on a subset of the MZ group for the method-distance combinations that showed highest effect sizes (lowest p-values) in **Table I** (red-colored values in the table). Each twin is represented by an integer on the x-axis. The distance to its twin sibling is represented by a filled blue circle. In most cases, members of a twin pair were less distant from each other than they were to the other subjects. This suggests the biological validity of these metrics and the registration from which they are derived.

VI. Discussion

A. Results

Here we combine the advantages of a large-deformation fluid matching approach with empirical statistics on population variability in anatomy to build SAFIRA, a Statistically-Assisted Fluid Registration Algorithm. SAFIRA was mathematically formulated using a *non-conservative Lagrangian* approach, which allows one to regard the $3D$ image volume as being embedded in a deforming mechanical system, subjected to conservative and non-conservative forces. Two types of prior information from the dataset can meaningfully be incorporated in the regularization terms. The covariance on the displacement fields q was included in the first term (Rayleigh's dissipation term) and the covariance of the deformation matrices Σ was embodied in the second term (or Riemannian term), using Log-Euclidean metrics. In both cases, Mahalanobis distances were used. The statistical dissipation acts during the registration process by slowing down or favoring the deformation at certain regions in the image. The medium is consequently considered as non-homogeneous, following prior papers that have ascribed non-uniform deformability to the medium (e.g., [52]).

From these results, we can conclude that all four versions of SAFIRA outperformed the traditional fluid registration approach, using various performance metrics. Preliminary observations in $2D$ showed that deformation patterns vary somewhat depending on the technique used. While the overall transformation was smoother with the vector-based statistics version of the algorithm, the inclusion of tensor information in the registration process gave the opposite effect. As the corpus callosum is a fairly narrow $2D$ structure at midline, the gradient of the displacement fields (∇q) is relatively high and so the tensor information tends to be quite variable in this region. In $3D$, both the vector- and tensor-based regularizer give smoother deformations than the non-statistical one, even though there is not as much evidence in the second case. In fact, the non-statistical prior has a stationary and isotropic autocorrelation which is the same everywhere in the image, by the definition of a homogeneous material. The empirically-derived priors, however, will have an autocorrelation that matches that of the data. Consequently, in large homogeneous regions, which are less subject to deformations, such as the white matter, the vector-based prior is likely to give smoother deformations than in the regions with more anatomical detail in the data (higher spatial autocorrelation in homogeneous regions). This effect can be improved by performing several rounds of registration to include a more accurate estimate of the covariance.

This autocorrelation effect has been further investigated in various cortical areas in [25]. In this paper, the authors used a large set of sulcal lines to evaluate the spatial correlation between any pair of cortical points, which they extended to interhemispheric correpondences. Naturally, they found high correlation between points that are spatially close to each other, but they also found long range local maxima in the correlation fields, revealing somewhat surprising connections between anatomical variation in distant brain regions. In an earlier study [29], the author similarly used the correlation between two points of the images, to guide an intra-subject registration method based on Bayesian statistics.

We also relied on the heritability of brain structures to quantify the power and effect sizes of each algorithm, in an application to a genetic study of brain structure. When computing statistics in $2D$, the statistical versions of the algorithm were slightly less powerful. Results differed in $3D$. First, it is worth noting that all four versions of SAFIRA improve the detection sensitivity compared to the traditional fluid. Secondly, the incorporation of the empirical information on q clearly resulted in more powerful results, whereas adding 1^{st}-order statistics did not influence the results compared to the non-statistical version.

The difference in power between the $2D$ and $3D$ results noted for the q-based method might be explained by a smaller influence of the statistical prior in the case of a $2D$ binary structure, compared to a grayscale $3D$ volume. In $2D$, registrations are obtained from the information found at the border of the structure only. For registration of binary data, the gradient data fidelity term is a very strong constraint and, as the structure is quite elongated, is quite densely defined in the image. The amount of information that is re-injected in the statistical registrations is thus relatively limited and the impact of the statistical regularizer compared to the non-statistical one may be somewhat minimal. However, in both dimension (and more so in $3D$), the information on the displacement seems to have more impact on the results and on the registration. This is likely due to the fact that a penalty on the deformation gradients Σ does not constrain the deformation field as much as a constraint on the displacement vectors, which restricts the magnitude of the deformation. The Σ's are computed from the derivatives of q and made symmetric. The strength of the statistics and thus its influence on the registration is smaller as a result, when the tensor statistics are used on their own.

The biological plausibility of all methods was also investigated. First, the volumes of different gray matter structures were compared between a registered image and the template image, and these automated labeling experiments favored the vector-based over the traditional fluid, the non-statistical and tensor-based versions. Furthermore, as the corpus callosum is a heritable structure, then under certain reasonable assumptions, the closer the intraclass correlation in the MZ population, the more accurate the registration method is. One of the assumptions of this metric is that there is no confound that would tend to lead to the registration errors being more similar for members of the same twin pair than for a randomly selected pair of subjects, or for twin pairs of different zygosities (MZ versus DZ). As all registrations are performed to a common template, which is based on a different subject (not one of the twins), then there should be no registration error x zygosity interaction, so this assumption is reasonable.

When global distances between images were examined, the vector-based statistical method consequently showed the greatest improvement in detection sensitivity versus the non-statistical Riemannian Fluid code. Theoretically, constraining the deformations via prior information on the dataset is equivalent to decreasing the dissipation in regions where the displacements are more likely to happen. Smoother transfor-

mations were also achieved using the vector-based statistical method, and these may also be more biologically plausible.

These findings offer some guidance on which algorithm to use. Relative to standard methods, priors that incorporate empirical statistics will tend to help, whenever the local deformation statistics (here the displacement field or its Jacobian matrix) are significantly non-stationary or non-isotropic. In other words, they should perform a more standard deformation model by a greater amount in brain regions where the directional biases of anatomical variation are greater.

When choosing the order of the statistics to include in the prior (on 0^{th}-order displacement fields or 1^{st}-order local deformation tensors), the statistics on the displacement field seem to provide greatest benefit; statistics on the deformation vectors are also more stable to compute than the local deformation tensors, which rely on spatial derivatives.

B. Limitations and Future Work

1) On the registration algorithm: This algorithm was built so that the deformations produced by all versions remain diffeomorphic. In fact, any singularity is prevented through the regridding step (see section III-F). However, SAFIRA does not ensure inverse consistency. At each step, a numerical optimization method is solved through a modified gradient-descent method and consequently local minima may be reached instead of global minima. An additional consequence is that the path found by this optimization scheme is not geodesic. However, in certain cases, such as in atlas construction, dependence upon the directionality of the registration can introduce a bias. One solution to this problem would be to optimize the transformation in a symmetric-fashion, such as in [5], [17], [47], [76], [77]. This could be implemented in the non-statistical case but could become computationally demanding when using the statistical versions.

Our formulation of the registration problem as a Lagrangian mechanical system places it in the class of methods that find a natural path of a deforming physical system over time, with forces that may depend on the path taken. As such, it is conceptually distinct from some other registration methods that optimize the transformation over the space of paths. For example, some methods regularize an energy on the space-time interval K x $[0, 1]$, and find the optimal path among the natural paths. Recent work on 'geodesic shooting' has noted that many optimization problems initially formulated on the space-time interval can be reformulated in terms of the initial momentum only, greatly reducing computational burden. In general, it is worth noting that these two approaches to the energy minimization problem have slightly different conceptual foundations.

2) On the validation method: The differences between the $2D$ and $3D$ results deserves comment. In $2D$, the registration problem is relatively simple, as it consists of registering binarized images of the corpus callosum. In that case, most algorithms with sufficiently many degrees of freedom will produce an accurate match, and the data fidelity term reduces to a difference between binary functions. In that simpler context, the non-statistical approach produced the most statistically sensitive measures in $2D$, but not in $3D$. It could be that this difference in power (and the relative improvement given by vector statistics in $3D$) occurs because binary shapes are easy to register even without statistical information, and not because the image is 2D. Statistical regularization may help to 'fix' the limitations of the similarity metric. That is, the performance of all algorithms is good in the $2D$ case, where the intensity difference metric is a good model of the underlying matching goal. In $3D$, however, the performance of the non-statistical approach may suffer because the intensity difference metric may not be a good model for the joint histogram of correctly registered T1-weighted MR images. It has to be recognized that, to some extent, the statistical prior may be overcoming limitations of the similarity metric. This is also true of any regularization method, as regularized solutions can be most beneficial in cases where (1) the signal to noise of the data is low, or (2) the information provided by the data fidelity term is poor or unreliable. A more agnostic interpretation of the results is that the statistical priors may help overcome a poorly chosen intensity model by reducing its impact in regions where the similarity metric is failing or introducing noise. This effect may explain the increased ICC in the prior-based algorithm and may be a more plausible explanation than the statistics being capable of capturing something with more biological meaning.

The use of the intraclass correlation, as one of our methods of validation can be discussed as well. When using twin data to study registration results, one goal is to both maximize ICC for the twins and maximize differences in the ICC groups of twins as well. It is worth nothing that this goal is somewhat different than the usual goal of TBM, which is to identify group differences in anatomy. The premise of the twin approach is that, in general (except for some confounding situations noted below) registration errors will tend to deplete the ICC obtained for twin pairs, and they will also tend to deplete the differences in ICCs between types of twins (MZ and DZ), which are the basis for studying genetic effects. In other words, ICCs may be considered as a possible measure of registration error. There are two caveats regarding this approach. The first is that the use of high ICCs to suggest good registration accuracy is just a guide, and cannot be treated as the sole criterion of registration accuracy. For example, it could be that in reality, twins differ in anatomy at a very fine scale, but resemble each other at a more global scale. If this biological scenario were true, then an algorithm with more degrees of freedom or with better recovery of correspondences at high frequencies may in fact lead to a lower ICC when run on the same set of twin scans. A similar argument may be proposed to suggest that the ICC between twins may be higher with decreasing deformation. However that is not the case, because every twin is independently registered to an independent template (not to each other), so there is no reason for smaller deformations to be more correlated among twins. Second, the registration errors tend to deplete the ICC, so they will be proportionally higher (lowering the ICC) when the deformation magnitude is too low. In the limiting case of zero deformation (poor registration), the ICC is not defined. Even so, the ICC is just one of many metrics that may be useful

for assessing a successful registration algorithm. In this paper, we also made sure that registration accuracy was assessed via the more traditional approach of quantifying manually labeled volumes (**Figure 4**), and the statistical version of the algorithm performed very well. Clearly, in a real TBM study of group differences, one would hope that the registration method could detect group differences with the highest possible effect size, or with the smallest minimal sample sizes (maximum efficiency). However, one of the issues with comparing effect sizes on data representing group differences is that we have no ground truth on how much difference there really is: and in fact an algorithm that finds smaller group differences may be more accurate. This is related to the confound that twins may resemble each other on a gross scale but not a fine scale. Overall, the agreement of deformation maps with independent methods for volumetric quantification is the primary criterion of success (see [45] for other metrics).

3) Future work: In this paper, we focused on considering the sum of squared intensity differences (SSD) as a data fidelity term, as it is perhaps the most commonly used. Even though SSD was proved to perform as well as other similarity measures for brain MRI [45] (see, e.g., http://picsl.upenn.edu/ANTS/metric.php), it would be worth comparing different similarity metrics such as cross-correlation or information-theoretic measures, when used in conjunction with the statistical formulation here. Without substantially altering the overall mathematical formulation here, more complex image similarity measures could replace the SSD, such as cross-correlation or information-theoretic measures. These data-dependent terms only affect the body force added to the cost function, and could easily be swapped in, in a modular way. Registration methods developed for more precise cortical pattern matching (e.g., [72]) could alternatively be used to better match cortical areas. For applications requiring more accurate registration of the cortical mantle, hybrid surface and volume registration methods have recently been proposed to enable precise simultaneous registration of both subcortical and cortical regions [50], [59].

Future work will also examine the value of including statistical information for longitudinal MRI studies, tracking brain change over time. In clinical trials, for example, it is often vital to maximize the statistical power to detect brain changes, but power falls off dramatically when the inter-scan interval is shorter [42]. In this context, a statistical prior may have additional value for registering serial images, as Bayesian theory suggests that the relative value of the empirical statistics increases when the image data are noisier or less informative.

ACKNOWLEDGMENTS

Algorithm development for this work was supported by the NIA, NIBIB, the National Library of Medicine, and the National Center for Research Resources (grants AG016570, EB01651, LM05639, RR019771 to P.T.) and by INRIA funding (to X.P.). We are grateful to the NSF Equipes Associées program (to Victor Vianu, UC San Diego) for supporting the scientific collaboration between our teams at UCLA and INRIA. The scanning of twins and genetic data analysis was supported by NIH grant R01 HD050735 and the National Health and Medical Research Council, Australia grant 496682 (to M.W. and P.T.). We thank one of our reviewers for their help in proposing explanations of the benefits of the statistical prior.

VII. APPENDIX A: COMPUTATION OF THE GENERALIZED LAGRANGIAN

For a dynamic system subject to generalized (i.e., conservative and non-conservative) forces, the action $\mathcal{S}$ can be expressed as:

$$\mathcal{S} = \int_{t_0}^{t_1} L + W dt \qquad \text{with} \qquad \delta W = \vec{F}.\delta\vec{r}$$

δW is the work created by the non-conservative forces $\vec{F}$ during the virtual displacement $\delta\vec{r}$. The path followed by a mechanical system minimizes the action, hence the variational principle with generalized forces,

$$\delta\mathcal{S} = \int_{t_0}^{t_1} \delta L + \delta W dt = 0$$

Given that the Lagrangian $L(q, \dot{q})$ is defined for the kinetic and potential energies,

$$\delta\mathcal{S} = \int_{t_0}^{t_1} \left(\delta\dot{q}\frac{\partial L}{\partial \dot{q}} + \delta q \frac{\partial L}{\partial q} + \delta q \vec{F}.\frac{\partial \vec{r}}{\partial q} \right) dt$$

$\delta q(t_0) = \delta q(t_1) = 0$ as t_0 and t_1 are the initial and final times. Using integration by parts, the first term in the equation becomes:

$$\int_{t_0}^{t_1} \delta\dot{q} \left(\frac{\partial L}{\partial \dot{q}} \right) dt = - \int_{t_0}^{t_1} \delta q \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} \right) dt$$

Thus,

$$\delta\mathcal{S} = \int_{t_0}^{t_1} \delta q \left(\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} + \vec{F}.\frac{\partial \vec{r}}{\partial q} \right) dt$$

hence,

$$\left(\frac{\partial L}{\partial q} \right) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}} \right) + \vec{F}.\left(\frac{\partial \vec{r}}{\partial q} \right) = 0$$

VIII. APPENDIX B: VARIATION OF THE ENERGY OF THE SYSTEM

The variation of the energy of the system with time is given by $\frac{\partial H}{\partial t}$. According to the definition,

$$H = p\dot{q} - L \qquad \text{with} \qquad p = \left(\frac{\partial L}{\partial q} \right)_{q\dot{q}}$$

The new Hamiltonian equation is found from the Lagrangian L, and modified to include the generalized forces:

$$\dot{p} = \frac{d}{dt}\left(\frac{\partial L}{\partial q} \right)_{q\dot{q}} = \left(\frac{\partial L}{\partial q} \right) + \vec{F}\left(\frac{\partial \vec{r}}{\partial q} \right) \tag{17}$$

So,

$$\frac{dH}{dt} = \frac{d}{dt}(p\dot{q} - L) = \dot{p}\dot{q} + p\ddot{q} - \frac{\partial L}{\partial q}\dot{q} - \frac{\partial L}{\partial \dot{q}}\ddot{q} - \frac{\partial L}{\partial t}$$

As in our case, $\frac{\partial L}{\partial t} = 0$

$$\frac{dH}{dt} = \dot{p}\dot{q} - \frac{\partial L}{\partial q}\dot{q} \tag{18}$$

Combining equations 18 and 17, we obtain

$$\frac{dH}{dt} = \vec{F}\left(\frac{\partial \vec{r}}{\partial q}\right)\dot{q}$$

IX. Appendix C: dissipative forces

If non-conservative forces are dissipative, then a term D can be included in the derivative of L such that

$$\left(\frac{\partial L}{\partial q}\right) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) + \frac{\partial D}{\partial \dot{q}} + \vec{F}.\left(\frac{\partial \vec{r}}{\partial q}\right) = 0$$

Here, $\vec{F}$ represents all the other non-conservative forces. In fact, as dissipative forces are inversely proportional to the velocity $\dot{q}$, the work becomes

$$\delta W = \vec{f}.\delta q = -\alpha \dot{q}^2 dt$$

and the power of this force is

$$P = -\frac{dW}{dt} = \alpha \dot{q}$$

The dissipation is such that $D = \frac{1}{2}P$, so

$$D = \frac{1}{2}\alpha \dot{q}^2$$

Consequently,

$$f = -\frac{\partial D}{\partial \dot{q}}$$

The first term of the nonconservative energy $V_{\vec{F}1}$ (see paragraph III-B) is a dissipative term.

References

[1] Allassonnière, S., Trouvé, A., Younes, L., *Geodesic Shooting and Diffeomorphic Matching Via Textured Meshes*, Energy Minimization Methods in Computer Vision and Pattern Recognition, (2005), 365–381

[2] Arsigny V., Fillard, P., Pennec, X., Ayache, N., *Log-Euclidean metrics for fast and simple calculus on diffusion tensors*, Mag Res Med, **56** (2), (2006) 411–421

[3] Avants, B.B., Gee, J.C., *Geodesic estimation for large deformation anatomical shape averaging and interpolation*, NeuroImage, 23(1), (2004) S139-S150

[4] Avants, B.B., Schoenemann, P.T., Gee, J.C., *Lagrangian frame diffeomorphic image registration: Morphometric comparison of human and chimpanzee cortex* Med Image Anal, **10** (3), (2006) 397–412

[5] Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C., *Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain*, Med Image Anal, **12** (1), (2008) 26–41

[6] Bajcsy, R., Kovačič, S., *Multiresolution elastic matching*, Computer vision, graphics and image processing, 46, (1989) 1-21

[7] Beg, M.F., Khan, A., *Computation of Average Atlas using LDDMM and Geodesic Shooting* IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Special Session on the use of shape in biomedical image formation, modeling, and analysis, April, 2006

[8] Bro-Nielsen, M., Gramkow, C., *Fast fluid registration of medical images*, 4th International conference on Visualization in Biomedical Computing, Hamburg, Germany, (1996) 272-276

[9] Brun, C., Lepore, N., Pennec, X., Chou, Y.-Y, Lee, A.D., de Zubicaray, G.I., McMahon, K., Wright, Barysheva, M., M., Toga, A.W., Thompson, P.M., *A new registration method based on Log-Euclidean tensor metrics and its application to genetic studies*, ISBI, Paris, France (2008)

[10] Brun C.C., Leporé, N., Pennec, X., Lee, A.D., Barysheva, M., Madsen, S.K., Avedissian, C., Chou, Y.-Y., de Zubicaray, G.I., McMahon, K., Wright, M.J., Toga, A.W., Thompson, P.M., *Mapping the Regional Influence of Genetics on Brain Structure Variability - A Tensor-Based Morphometry Study*, NeuroImage, 48(1) (2009), 37–49

[11] Cachier, P., et al., *Symmetrization of the Non-Rigid Registration Problem using Inversion-Invariant Energies: Application to Multiple Sclerosis*, MICCAI, Pittsburgh, PA, USA, (2000) 472-481

[12] Chiang, M.-C., Dutton, R.A., Hayashi, K.M., Toga., A.W., Lopez, O.L., Aizenstein, H.J., Becker, J.T., Thompson P.M., *Fluid registration of medical images using Jensen-Rényi divergence reveals 3D-profile of brain atrophy in HIV/AIDS*, Proceedings of the 3rd IEEE International Symposium on Biomedical Imaging, Arlington, Virginia, USA, 6-9 April (2006), 193–196

[13] Chiang, M.C., Reiss, A.L., Lee, A.D., Bellugi, U., Galaburda, A.M., Korenberg, J.R., Mills, D.L., Toga, A.W., Thompson, P.M., *3D pattern of brain abnormalities in Williams syndrome visualized using tensor-based morphometry*, Neuroimage, 36(4), (2007) 1096–1109

[14] Chiang, M.C., Barysheva, M., Lee, A.D., Madsen, S.K., Klunder, A.D., Toga, A.W., McMahon, K.L., de Zubicaray, G.I., Wright, M.J., Srivastava, A., Balov, N., Thompson, P.M., *Genetics of Brain Fiber Architecture and Intelligence*, Journal of Neuroscience, 29(7), (2009) 2212–2224

[15] Christensen, G.E., Rabbit, R.D., Miller, M.I., Joshi, S.C., Grenander, U., Coogan, T., Van Essen, D.C., *Topological properties of smooth anatomic maps* In Bizais, Braillot, and Di Paola, editors, Information Processing in Medical Imaging, Kluwer Academic Publishers, Boston, (3), (1995) 101–112

[16] Christensen, G.E., Rabbitt, R.D., Miller, M.I., *Deformable templates using large deformation kinematics*, IEEE Trans. Image Process. 5, (1996) 1435–1447

[17] Christensen, G.E, Johnson, J.H., *Consistent Image Registration*, IEEE Transactions on Medical Imaging, 20(7), (2001) 568–582

[18] Collins, D.L., Zijdenbos, A.P., Kollokian, V., Sled, J.G., Kabani, N.J., Holmes, C.J., Evans, A.C. *Design and construction of realistic digital brain phantom*, IEEE-TMI, 17(3), (1998) 463–468

[19] Commowick, O., Stefanecu, R., Fillard, P., Arsigny, V., Ayache, N., Pennec, X., Malandain, G., *Incorporating Statistical Measures of Anatomical Variability in Atlas-to-Subject Registration for Conformal Brain Radiotherapy*, MICCAI, 2(3750), (2005), 927–934

[20] Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J., *Active shape models - their training and application*, Computer Vision and Image Understanding, 61(1), (1995), 38–59

[21] Durrleman, S., Pennec, X., Trouvé, A., Ayache, N., *Statistical models of sets of curves and surfaces based on currents* Medical Image Analysis, 13(5), (2009), 793–808

[22] Edgington, E.S., *Randomization Tests*, 3rd Edition Marcel Dekker, (1995)

[23] Falconer. D.S., *Introduction to Quantitative Genetics*, 3^{rd} Ed. Essex, UK: Longman, (1989)

[24] Fillard, P., Pennec, X., Thompson, P.M., Ayache, N., *Evaluating brain anatomical correlations via canonical analysis of sulcal lines*, Proceedings, Information Processing in Medical Imaging (IPMI), (2005)

[25] Fillard, P., Arsigny V., Pennec, X., Thompson, P.M., Ayache, N., *Extrapolation of sparse tensor fields: Application to the modeling of brain variability*, Inf Process Med Imaging, 19, (2005) 27–38

[26] Fillard, P., Arsigny, V., Pennec, X., Hayashi, K.M., Thompson, P.M., Ayache, N., *Measuring Brain Variability by Extrapolating Sparse Tensor Fields Measured on Sulcal Lines*, NeuroImage, 34(2), (2007) 639–650

[27] Friston, K.J., *Statistical Parametric Mapping: Ontology and Current Issues*, Journal of Cerebral Blood Flow and Metabolism, 15, (1995), 15, 361–370

[28] Gee, J.C., Barillot, C., Briquer, L.L., Haynor, D., Bajcsy, R., *Matching structural images of the human brain using statistical and geometrical image features*, SPIE, Visualization in Biomedical Computing, 2359, (1995) 191–204

[29] Gee, J.C., *On matching brain volumes*, Pattern Recognit, (1999)

[30] Gee, J.C., Bajcsy, R., *Elastic Matching: Continuum Mechanical and Probabilistic Analysis*, In Toga, A., Brain Warping, Academic Press, (1999) 183–197

[31] Gogtay, N.*, Lu, A.*, Leow, A.D., Klunder, A.D., Lee, A.D., Chavez, A., Greenstein, D., Giedd, J.N., Toga, A.W., Rapoport, J.L., Thompson, P.M., *Three-dimensional brain growth abnormalities in childhood-onset schizophrenia using Tensor-Based Morphometry*, Proc Natl Acad Sci, 105(41), (2008), 15979–15984 [* equal contribution]

[32] Goldstein, H., Classical Mechanics, Addison-Wesley Ed, July 1980

[33] Gramkow C., *Registration of 2D and 3D medical images*, Master's thesis, Danish Technical University, Copenhagen, Denmark (1996).

[34] Grenander, U., Miller, M.I., *Computational anatomy: An emerging discipline*, Quart of App Maths 56, (1998) 617-694

[35] Grenander, U., Srivastava, A., Saini, S., *Characterization of biological growth using iterated diffeomorphisms*, International Symposium on Biomedical Imaging, (2006), 1136–1139.

[36] Grenander, U., Srivastava, A., Saini, S., *A Pattern-Theoretic Characterization of Biological Growth*, IEEE Trans. Med. Imaging, **26**(5), (2007), 648–659.

[37] Hernandez, M., Olmos, S., Pennec, X. *Comparing algorithms for diffeomorphic registration: Stationary LDDMM and Diffeomorphic Demons*, Workshop on Mathematical Foundations of Computational Anatomy, (MICCAI), New-York, USA (2008)

[38] Holden, M., *A review of geometric transformations for nonrigid body registration*, IEEE Transactions Medical Imaging, 27(1), (2008), 111–127

[39] Hofer, S., Frahm, J., *Topography of the human corpus callosum revisited - Comprehensive fiber tractography using diffusion tensor magnetic resonance imaging*, NeuroImage, 32(3),(2006), 989–994

[40] Hua, X., Leow, A.D., Parikshak, N., Lee, S., Chiang, M.C., Toga, A.W., Jack, C.R., Weiner, M.W., Thompson, P.M., *Tensor-Based Morphometry as a Neuroimaging Biomarker for Alzheimers Disease: An MRI Study of 676 AD, MCI, and Normal Subjects*, NeuroImage, 43(3), (2008) 458–469

[41] Hua, X., Leow, A.D., Levitt, J.G., Caplan, R., Thompson, P.M., Toga, A.W., *Detecting brain growth patterns in normal children using tensor-based morphometry*, Human Brain Mapping, 30(1), (2009), 209–219

[42] Hua, X., Lee, S., Hibar, D.P., Yanovsky, I., Leow, A.D., Toga, A.W., Jack, C.R.Jr, Bernstein, M.A., Reiman, E.M., Harvey, D.J., Kornak, J., Schuff, N., Alexander, G.E., Weiner, M.W., Thompson, P.M., *Mapping Alzheimer's Disease Progression in 1309 MRI Scans: Power Estimates for Different Inter-Scan Intervals*, NeuroImage, 51(1), 63–75, (2009)

[43] Hulshoff Pol, H.E., Schnack, H.G., Posthuma, D., Mandl, R.C.W., Baaré, W.F., Van Oel, C.J., Van Haren, N.E., Collins, D.L., Evans, A.C., Amunts, K., Bürgel, U., Zilles, K., de Geus, E., Boomsma, D.I., Kahn R.S., *Genetic Contributions to Human Brain Morphology and Intelligence*, J Neurosci, 26(40), (2006) 10235–10242

[44] Jenkinson, M., Bannister, P.R., Brady, J.M., Smith, S.M., *Improvised optimization for the robust and accurate linear registration and motion correction of brain images*, Neuroimage, 17(2), (2002) 825–841

[45] Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B.B., Chiang, M.-C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., Song, J.H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R.P., Mann, J.J., Parsey, R.V., *Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration*, NeuroImage, 46(3), (2009) 786–802

[46] Lee, A.D., Leow, A.D., Lu, A., Reiss, A.L., Hall, S., Chiang, M.-C., Toga, A.W., Thompson, P.M., *3D Pattern of Brain Abnormalities in Fragile X Syndrome visualized using Tensor-Based Morphometry*, Neuroimage, 34(3), (2007) 924–938

[47] Leow, A.D., Chiang, M.-C., Yanovsky, I., Lee, A.D., Lu, A., Klunder, A.D., Becker, J.T., Davis, S.W., Toga, A.W., Thompson, P.M., *Statistical properties of Jacobian maps and the realization of unbiased large-deformation nonlinear image registration*, IEEE Transactions on Medical Imaging, **26**(6), (2007) 822–832

[48] Leporé, N., Brun, C., Chou, Y.-Y, Chiang, M.-C., Dutton, R.A., Hayashi, K.M., Lueders, E., Lopez, O.L., Aizenstein, H.K., Toga, A.W., Becker, J.T., Thompson, P.M., *Generalized Tensor-Based Morphometry of HIV/AIDS using multivariate statistics on deformation tensors*, IEEE Transactions on Medical Imaging, 27(1), (2008) 129–14

[49] Leporé, N., Chou, Y.-Y., Lopez, O.L., Aizenstein H.J., Becker, J.T., Toga, A.W., Thompson, P.M., *Fast 3D Fluid Registration of Brain Magnetic Resonance Images*, SPIE Medical Imaging conference, San Diego, CA, February 17-21 (2008)

[50] Leporé, N., Joshi, A., Leahy, R., Brun, C.C., Chou, Y.-Y., Pennec, X., Lee, A.D., Barysheva, M., de Zubicaray, G.I., Wright, M., McMahon, K., Toga, A.W., Thompson, P.M., *Surface and volume registration of brain magnetic resonance images*, SPIE, (2009)

[51] Leporé N., Voss P., Chou, Y.-Y., Fortin, M., Gougoux, F., Leporé F., Lee, A.D., Brun, C.C., Lassonde, M., Madsen, S.K., Toga, A.W., Thompson, P.M., *Brain Structure Changes Visualized in Early- and Late-Onset Blind Subjects*, NeuroImage, 2009 Jul 28. [Epub ahead of print].

[52] Lester, H., Arridge, S.R., Jansons, K.M., Lemieux, L., Hajnal, J.V., Oatridge, A., *Non-linear registration with the variable viscosity fluid algorithm*, Proceedings, Information Processing in Medical Imaging (IPMI), Visegràd, Hungary (1999)

[53] Marsland, S., Twining, C.J., *Constructing an atlas for the diffeomorphism group of a compact manifold with boundary with application to the analysis of image registrations*, JCAM, 222(2), (2008) 411–428

[54] Modersitzki, J., *Numerical methods for image registration*, Numerical mathematics and scientific computation. Oxford University Press. (2004)

[55] Miller, M.I., *Computational anatomy: shape, growth, and atrophy comparison via diffeomorphism*, Neuroimage, 23(1), (2004) 19–33

[56] Nichols, T.E. and Holmes, A.P., *Non parametric permutation tests for functional neuroimaging: a primer with examples*, Hum Brain Map, 15(1), (2002) 1–25

[57] Pennec. X., Stefanescu, R., Arsigny, V., Fillard, P., Ayache, N., *Riemannian elasticity: A statistical regularization framework for non-linear registration*, MICCAI, Palm Springs, CA, USA, (2005) 943-950

[58] Portman, N., Grenander, U., Vrscay, E.R., *Direct Estimation of Biological Growth Properties from Image Data Using the 'GRID' Model*, International Conference on Image Analysis and Recognition, (2009), 832–843

[59] Postelnicu*, G.M., Zöllei*, L., Fischl, B., *Combined Volumetric and Surface Registration*, IEEE Transactions on Medical Imaging (TMI), 28(4), (2009) 508–522 [* equal contribution]

[60] Subhankar, R., and Shamanna, J., *On virtual displacement and virtual work in Lagrangian dynamics*, Eur J Phys, 27(2), (2006) 311–323

[61] Rueckert, D., Frangi, A.F., Schnabel, J.A., *Automatic construction of 3D statistical deformation models of the brain using non-rigid registration*, IEEE Transactions on Medical Imaging, 22(8), (2003) 1014–1025

[62] Runge, C., *Über die numerische Auflösung von Differentialgleichungen*, Math Ann, 46, (1895) 167–178

[63] Sarrut, D., Miguet, S., *Fast 3D Image Transformations for Registration Procedures*, 10^{th} International Conference on Image Analysis and Processing (ICIAP), Venice, Italy, September 1999

[64] Scout PE, Fleiss JL, *Intraclass correlations: Uses in assessing rater reliability*, Psychol Bull 2, (1979) 420-428

[65] Shattuck D.W. and Leahy R.M., *BrainSuite: an automated cortical surface identification tool*, Med Image Anal 8, (2002) 129–141

[66] Shattuck D.W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K.L., Poldrack, R.A., Bilder, R.M., Toga, A.W., *Construction of a 3D probabilistic atlas of human cortical structures*, NeuroImage 39(3), (2008) 1064–1080

[67] Sowell ER, Peterson BS, Thompson PM, Welcome SE, Henkenius AL, Toga AW. *Mapping cortical change across the human life span*, Nature Neuroscience, 6(3), (2003) 309–315

[68] Stefanescu, R., Pennec, X., Ayache, N., *Grid powered nonlinear image registration with locally adaptive regularization*, Medical Image Analysis, 8(3), (2002) 325–342

[69] Thompson, P.M, MacDonald, D., Mega, M.S., Holmes, C.J., Evans, A.C., Toga, A.W., (1997) *Detection and Mapping of Abnormal Brain Structure with a Probabilistic Atlas of Cortical Surfaces*, Journal of Computer Assisted Tomography, 21(4), 567–581, (1997)

[70] Thompson, P.M., Mega, M.S., Narr, K.L., Sowell, E.R., Blanton, R.E., Toga, A.W., *Brain images and atlas construction*, Proceedings, SPIE Medical Imaging Conference, San Jose, CA, USA, (2000)

[71] Thompson, P.M., Giedd, J.N., Woods, R.P., MacDonald, D., Evans, A.C., Toga, A.W., (2000). *Growth Patterns in the Developing Brain Detected By Using Continuum-Mechanical Tensor Maps*, Nature, 404(6774), (2000) 190–193

[72] Thompson, P.M., Hayashi, K.M., Sowell, E.R., Gogtay, N., Giedd, J.N., Rapoport, J.L., de Zubicaray, G.I., Janke, A.L., Rose, S.E., Semple, J., Doddrell, D.M., Wang, Y., van Erp, T.G., Cannon, T.D., Toga, A.W., *Mapping cortical change in Alzheimer's disease, brain development, and schizophrenia*, Neuroimage, 23(1), (2004) 2–18

[73] Trouvé, A., *Diffeomorphisms Groups and Pattern Matching in Image Analysis*, Int J Comput Vis, 28(3), (1998) 213–221

[74] Tveter, F.T., *Hamilton's equations of motion for non-conservative systems*, Celestial Mechanics and Dynamical Astronomy, 60(4), (1994) 409–419

[75] Vercauteren, T., Pennec, X., Malis, E., Perchant, A., Ayache, N., *Insight into efficient image registration techniques and the Demons algorithm*, Proceedings, Information Processing in Medical Imaging (IPMI), The Netherlands (2007)

[76] Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., *Symmetric Log-Domain Diffeomorphic Registration: A Demons-based Approach*. Proc. Medical Image Computing and Computer Assisted Intervention (MICCAI), New York, USA, 1(5241), September 2008, 754–761.

[77] Yanovsky, I., Leow, A.D., Lee, S., Osher, S.J., Thompson, P.M., *Asymmetric and Symmetric Unbiased Image Registration: Statistical Assessment of Performance*, Medical Image Analysis, 13(5), (2009) 679–700

Fig. 1. Example of a virtual displacement δr. The system consists of a bead on a moving inclined plane (moving at the speed u). The virtual displacement can be seen as the difference between two displacements that are allowed by the constraint forces (vdt and $v'dt$) between t and $t + dt$.

Fig. 2. Circle deformed into an ellipse using the non-statistical algorithm (green grid), and the statistical algorithm with small dissipation (red grid - *left panel*) and large dissipation (blue grid - *right panel*). A large dissipation delays the progression of the transformation.

Fig. 3. **Top row:** Deformation grids obtained from the registration of one of the subjects's $2D$ corpus callosum to the template image using the four versions of the algorithm, showing the regularity of the deformation. From *left* to *right*: non-statistical regularization, vector-based statistics, tensor-based statistics, vector and tensor-based statistics versions. **Bottom row**: Maps of the determinants computed from the deformation matrices (Σ's), themselves computed from the displacement fields ($\vec{q}$). These maps give a measure of the smoothness of the deformation (see IV-B).

Fig. 4. Volume Quantification Error. Dark blue colors denote more accurate measurements: the volumetric labeling error is computed as the difference between the ground truth and estimated volumes, divided by their mean volume. Low values (dark blue) therefore denote more accurate volume quantification. These measures of registration accuracy were computed across the LPBA40 population (a standard manually labeled brain MRI dataset) for 56 structures (shown here on the y-axis). Results are shown for the non-statistical version of the registration algorithm, and for the versions of the algorithm that include vector- or tensor-based statistics during the registration and lastly for the traditional fluid registration (from left to right). Dark blue shades indicate that the difference in volume is very small whereas red colors indicate a large difference between the two volumes (i.e., poor volumetric overlap). In most cases and across all regions, the registration method that incorporates vector-based empirical statistics consistently performs better for volume quantification. Areas that are difficult to label include the right cingulate, inferior gyri of the occipital lobe, and the insular cortex. In these regions, gyral patterning is notoriously variable across subjects, and labeling is more likely to be successful with a registration approach that also includes cortical constraints.

Fig. 5. The smoothness of a transformation can be illustrated using the deformation matrix Σ. The images represent the variance of the norm of Σ measured across the whole population for the four algorithms (IV-D). Blue colors indicate regions where this variance is small ($Var = 0$), whereas red colors represent a higher relative variance ($Var = 0.05$). The corresponding anatomical images are shown in Fig. 10. The variance of the deformation is smaller when using statistical regularization. The variance in the deformation fields is quite heavily reduced when prior information on the displacement vectors is included.

Fig. 6. Maps of the covariance of the vector fields obtained from the 0- non-statistical algorithm, 1- the vector-stat version, 2- the vector-stat version including the information from 1, 3- the vector-stat version including prior information from 2. The scale is logged. Blue corresponds to small covariance and red colors correspond to regions of high covariances. The overall covariance decreases from 0 to 3. At a given step, the vector-statistical algorithm controls the deformation in regions of low covariance. Resulting vector fields are more homogeneous in these regions, and thus the resulting covariance is lower.

Fig. 7. Intraclass correlation, heritability and significance maps are presented for TBM analysis performed using each of the four registration methods. From *left* to *right*: non-statistical, vector-based statistics, tensor-based statistics, vector and tensor-based statistics versions. In the intraclass correlation maps (1^{st} and 2^{nd} row), blue corresponds to 0 (i.e. no correlation) while red colors indicate a high correlation ($ICC = 0.75$). The same scale applies to the heritability maps (3^{rd} column). Blue colors indicate regions where there are little or no detectable genetic influences, whereas red colors show regions in which genetic influences on brain structures are relatively high. The bottom two rows show the significance of the intraclass correlation computed separately for the MZ and DZ groups (3^{rd} and 4^{th} row, respectively). Green regions indicate regions that are significant (p-value< 0.05) - see IV-E. All registration methods generated a similar pattern for all of the statistical measures.

Fig. 8. Intraclass correlation computed in the MZ and DZ groups from Jacobian determinants from the four registration methods. From *left* to *right*: non-statistical, vector statistics, tensor statistics, vector and tensor statistics versions and Traditional Fluid. Blue colors correspond to 0 (i.e., no correlation) while red colors indicate a high correlation ($ICC = 0.75$). The different slices represented here correspond to the anatomical images in Figure 10. In the *top* and the *bottom* panels, the first row corresponds to A, the second and third to B and C, respectively. The black circles show the main regions of differences between all the methods. The incorporation of the vector statistics in the registration gives more powerful results in the TBM statistical analysis.

Fig. 9. Maps showing the voxelwise significance of the intraclass correlation (the correlation values are shown in Figure 8). These p-values were computed using voxel-wise permutation (see IV-E). From *left* to *right*: non-statistical, vector statistics, tensor statistics, vector and tensor statistics algorithms and Traditional Fluid algorithm. Red colors indicate significant effects at the voxel level (p-value < 0.05), while blue regions are not significant (p-value > 0.05). The scale has been logarithmically transformed so that the significant regions are found from 0 (red) to 0.05 (yellow). The slices from the top to the bottom row correspond to A-C (see *bottom panel*, Figure 10). Black circles show the main regions of differences across all the registration methods. Incorporation of vector-based statistics during the registration process also boosts statistical power in the TBM statistical analysis, although all methods give largely consistent results.

Fig. 10. *Top panel*: Maps of heritability coefficients (which show the degree to which brain structure depends on individual genetic differences) are computed from the maps of intraclass correlations for both MZ and DZ groups (see IV-E). From *left* to *right*: non-statistical, vector statistics, tensor statistics, vector and tensor statistics algorithms and Traditional Fluid algorithm. Blue colors indicate regions where no genetic influences are detected $h^2 = 0$ whereas red colors indicate a relatively high heritability, $h^2 = 0.75$. *Top panel*: Anatomical view of the template brain (target used in the non-fluid registration). A, B and C show the slices presented in all the panels in Figure 5 and Figures 8 to 10, from the *top row* to the *bottom row*. Black circles show the main regions of differences across all the registration methods. Again, incorporation of vector-based statistics during the registration process also boosts statistical power in the TBM statistical analysis, although all methods give largely consistent results.

Fig. 11. Cumulative distribution functions are shown for the observed p-values in $3D$ for the MZ group intraclass correlation, versus the corresponding p-value under the null hypothesis for the non-statistical (magenta line), vector-based (black line), tensor-based (yellow line), and joint vector- plus tensor-based (dark blue line) statistical versions of the algorithm. The light blue color line represents the traditional fluid. Note that the dark blue line is on top of the black line, and shows that using vector statistics, or both vector and tensor statistics improves the effect sizes. This was confirmed by the computation of the areas under the curve (A). The green dotted line shows the expected distribution of p-values under the null hypothesis (here, the null hypothesis would be that members of MZ twin pairs would show no statistical similarity in brain structure). The detection power is higher in all versions of SAFIRA compared to the traditional fluid but appears to be slightly depleted by the addition of a prior on deformation matrices (tensors) in the registration process (yellow line). However, it is greatly improved by the addition of prior information on deformation fields (black and blue lines). This type of CDF plot is fundamental to tensor-based morphometry analyses that use the false discovery rate method as a criterion to decide whether the overall pattern of findings in a statistical map is significant, after multiple comparisons correction. As such, steeper plots are generally the ones that can be thresholded in such as way as to control the false discovery rate in the supra-threshold set, and would therefore be regarded as significant in a voxel-based brain mapping study.

Fig. 12. Plot showing the distances between subjects' corpus callosum outlines, for the distance-algorithm combinations highlighted in red in **Table I**. From *left* to *right*: d_{q1} and Vector and tensor-based statistics version - d_{q1} and vector-based statistics version - d_{s1} and non-statistical version. Each integer on the x-axis corresponds to a twin subject. The blue filled circles represent the distance between two members of a twin pair, whereas the other colored circles represent the distance between one twin and the rest of the subset (i.e., other subjects unrelated to that twin). In general, members of a twin pair are less distant from each other than they are from the other subjects, which is in line with the notion that anatomy is highly heritable. In addition to the measures for volume quantification, these plots show that the registration methods produce distance metrics between anatomies that are sensitive to known biological information, such as the high heritability of brain structure in human populations.

		$\mathbf{d_{q_1}}$	$\mathbf{d_{q_2}}$	$\mathbf{d_{s_1}}$	$\mathbf{d_{s_2}}$
(0)	**ICC**	0.61	0	0.56	0.07
	***p*-value**	0.002	1	0.036	0.34
(1)	**ICC**	0.77	0.42	0.46	0.17
	***p*-value**	0.0002	0.02	0.011	0.22
(2)	**ICC**	0.50	0	0.14	0.10
	***p*-value**	0.009	1	0.24	0.29
(3)	**ICC**	0.76	0.27	0.08	0
	***p*-value**	0.0001	0.09	0.30	1

TABLE I

INTRACLASS CORRELATION AND ITS SIGNIFICANCE (p-VALUE) COMPUTED FROM THE FOUR DISTANCES d_{q1}, d_{q2}, d_{s1} AND d_{s2} (SEE PARAGRAPH IV-F) APPLIED TO THE CORPUS CALLOSUM FOR EACH ALGORITHM. (0) NON STATISTICAL VERSION - (1) VECTOR-BASED STATISTICS VERSION - (2) TENSOR-BASED STATISTICS VERSION - (3) VECTOR AND TENSOR STATISTICS VERSION. THE MORE SIGNIFICANT RESULTS ARE SHOWN IN RED. IN PARTICULAR, d_{q1} AND d_{q2} FAVORS THE VECTOR-BASED VERSIONS WHILE d_{s1} IS A BETTER FIT FOR THE NON-STATISTICAL TECHNIQUE. OVERALL, THE BEST APPROACH FOR ALL THE DISTANCES IS THE REGISTRATION METHOD USING VECTOR-BASED STATISTICS ONLY.