NeuroImage 169 (2018) 11–22

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

A spatio-temporal reference model of the aging brain

W. Huizinga [a,*], D.H.J. Poot [a,b,1], M.W. Vernooij [c,d,1], G.V. Roshchupkin [a], E.E. Bron [a], M.A. Ikram [c,d,e], D. Rueckert [f], W.J. Niessen [a,b], S. Klein [a], the Alzheimer's Disease Neuroimaging Initiative[2]

[a] Biomedical Imaging Group Rotterdam, Depts. of Radiology & Medical Informatics, Erasmus MC, Rotterdam, The Netherlands
[b] Quantitative Imaging Group, Dept. of Imaging Physics, Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands
[c] Department of Radiology and Nuclear Medicine, Erasmus MC, Rotterdam, The Netherlands
[d] Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands
[e] Department of Neurology, Erasmus MC, Rotterdam, The Netherlands
[f] Biomedical Image Analysis Group, Department of Computing, Imperial College London, United Kingdom

ARTICLE INFO

Keywords:
Brain morphology
Aging
Percentile curves
Non-rigid groupwise registration
Partial least squares regression
Spatio-temporal atlas

ABSTRACT

Both normal aging and neurodegenerative disorders such as Alzheimer's disease (AD) cause morphological changes of the brain. It is generally difficult to distinguish these two causes of morphological change by visual inspection of magnetic resonance (MR) images. To facilitate making this distinction and thus aid the diagnosis of neurodegenerative disorders, we propose a method for developing a spatio-temporal model of morphological differences in the brain due to normal aging. The method utilizes groupwise image registration to characterize morphological variation across brain scans of people with different ages. To extract the deformations that are due to normal aging we use partial least squares regression, which yields modes of deformations highly correlated with age, and corresponding scores for each input subject. Subsequently, we determine a distribution of morphologies as a function of age by fitting smooth percentile curves to these scores. This distribution is used as a reference to which a person's morphology score can be compared. We validate our method on two different datasets, using images from both cognitively normal subjects and patients with Alzheimer disease (AD). Results show that the proposed framework extracts the expected atrophy patterns. Moreover, the morphology scores of cognitively normal subjects are on average lower than the scores of AD subjects, indicating that morphology differences between AD subjects and healthy subjects can be partly explained by accelerated aging. With our methods we are able to assess accelerated brain aging on both population and individual level. A spatio-temporal aging brain model derived from 988 T1-weighted MR brain scans from a large population imaging study (age range 45.9–91.7y, mean age 68.3y) is made publicly available at www.agingbrain.nl.

Introduction

Magnetic Resonance (MR) imaging plays an important role in diagnosing neurodegenerative diseases due to its depiction of the brain morphology in vivo (Vernooij and Smits, 2012). Interpretation of MR images in the context of dementia diagnosis can be challenging, as early brain abnormalities may be difficult to distinguish from those related to normal aging, especially in the early stages of the disease. Quantitative methods that can distinguish brain morphology due to healthy aging from morphology due to accelerated aging or pathology can therefore aid and possibly improve the diagnosis of neurodegenerative diseases (Brewer, 2009).

Quantitative information on brain morphology is usually obtained by measuring e.g. tissue volumes and regional volumes (Brewer, 2009). However, these measures do not provide fully detailed information about shape differences, since volume is a quantity of an enclosed surface, and shape is a description or outline and therefore potentially much more informative (Davis et al., 2010). Results on hippocampal shape studies

* Corresponding author.
E-mail address: w.huizinga@erasmusmc.nl (W. Huizinga).
[1] Shared second author: D.H.J. Poot supervised the study from the image processing methodology perspective, and M.W. Vernooij from the population imaging perspective.
[2] Data used in preparation of this article were obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

https://doi.org/10.1016/j.neuroimage.2017.10.040
Received 16 June 2017; Received in revised form 12 September 2017; Accepted 19 October 2017
Available online 5 December 2017

suggest that shape may have additional predictive value over volume when used in the prediction of Alzheimer's disease (Achterberg et al., 2010; Costafreda et al., 2011). Therefore, there is an emerging interest in methods for quantifying shape differences and variations in shape in the human brain.

In literature, several methods for estimating models quantifying these shape differences and changes have been proposed. Davis et al. (2010) proposed a kernel regression on image dissimilarities to estimate a brain image representative for each age. Both Serag et al. (2012) and Dittrich et al. (2014) use a similar or more advanced kernel regression to build a spatio-temporal atlas for neonatal and fetal brain development respectively. The latter two methods are especially suited for fetal and neonatal brain development, where the brain rapidly grows with increasing age. Fishbaugh et al. (2017) developed a geodesic shape regression method which uses a sparse representation of diffeomorphisms, describing complex nonlinear changes over time with a small number of model parameters defined by the user. The mentioned methods estimate change in mean morphology of the population with age, but do not model the statistical distribution; the mean, but not the variance of the morphology at a certain age, is modelled. This concern was addressed by Ziegler et al. (2012), who presented numerous approaches that relate aging to differences in brain morphometry. They considered generative models (in which brain morphology is predicted from age) and recognition models (in which age is predicted from brain morphology) in cross-sectional data, and models that estimate individual decline and explain inter-individual variability in aging in longitudinal data.

Rather than predicting age or brain morphology, or classifying healthy and diseased subjects based on brain morphology, we propose a method that generates a reference distribution of healthy brain morphologies as a function of age to which an individual brain can be compared. Like Marquand et al. (2016), Ziegler et al. (2014), and Brewer (2009) we use a normative modeling approach, in which we aim to quantify the variation within a population and assess deviations from that population. Marquand et al. applied normative modeling to assess deviations of brain structure or function as a function of clinical covariates (e.g. cognitive test scores). Brewer et al. assesses deviations of volumetric MR imaging measures as a function of age. Ziegler et al. provide normative voxelwise maps of local gray matter abnormalities and global tissue volume z-scores. In our approach we aim to assess deviations of typical aging patterns of the brain morphology, measured with voxelwise deformations. To generate a distribution of brain morphologies we aim to find brain deformation patterns that are highly correlated with age. To achieve this we choose a data-driven approach using structural MR brain scans of elderly people in a wide age range (46y - 92y). We determine the morphology distribution by applying a regression model to the morphological variation within these brains. The morphological variation is characterized by deformation fields that map each brain image to a common space, which is a standard approach in computational anatomy (Baloch and Davatzikos, 2009). We compute these deformation fields with a groupwise image registration technique. Since we are only interested in the deformations due to aging, we employ a regression technique called partial least squares regression (PLSR). PLSR is especially suitable when there are more predictors than samples and when the predictors are highly collinear or linearly dependent (Wold et al., 2001; de Jong, 1992). It was first evaluated for neuroimaging by Krishnan et al. (2011) to relate brain function to behavior. Ziegler et al. (2013) used multivariate PLS correlation to explore the relationship between cognitive ability patterns and differences in local brain anatomy in the maturing brain. Singh et al. (2014) used PLSR to quantify anatomical shape variation in the brain. They used kernel PLSR to find the relationship between the manifold of diffeomorphisms from atlas to subject domain and global cognitive and functional assesment test scores. Whereas Singh et al. were interested in the PLSR regression coefficient, we aim to find deformation patterns that are most correlated with age, i.e. the PLSR loadings. With each of these loadings comes a corresponding score, and we use these scores to quantify the distribution of brain morphologies due to aging. The density of a distribution can be indicated with percentile values: measures specifying the value below which a given percentage of observations in a group of observations fall. The morphology distribution, however, is a function of age, and therefore we fit percentile *curves* to quantify how this distribution varies with age. To assess if an individual suffers from accelerated brain aging, its score can be compared to these percentile curves.

To validate our method, we use 988 structural MR brain scans from the population-based Rotterdam Scan Study (RSS), a prospective study among community dwelling subjects aged 45 years and over (Ikram et al., 2015). The method's robustness against scanning protocol and its diagnostic value is evaluated using the 988 scans from the RSS and a selection of 509 scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset which is adopted from the study of Cuingnet et al. (2011). The morphology score distribution and the corresponding 4D atlases are made publicly available through a web-based application at www.agingbrain.nl.

Extension of preliminary results

Preliminary results of this method were presented at SPIE Medical Imaging, San Diego 2016 (Huizinga et al., 2016b). In the current work, we considerably expand the previous study. First, in this version we use the displacement field inside the brain instead of the transformation parameters, in order to exclude deformation outside the brain. The control points of the B-spline transformation model that exist outside the brain region influence the deformation field both outside and inside the brain. We do not want deformations in the background to be part of the modeling, which is why we chose the displacement field inside the brain. This voxel-based approach also makes the method more generalizable to other nonparametric registration methods and other voxelwise tissue property maps. Second, we correct for the subject's head position in the scanner by removing rigid body motion inside the brain mask from the deformation field. Third, we introduce a method to determine the number of PLSR components. Fourth, we analyze the residuals that contain deformations due to factors other than aging, e.g. unexpected pathologies. Fifth, we take into account other covariates such as sex and height. Sixth, we added an evaluation of our method on the ADNI dataset.

Methods

We propose a method for modeling brain morphology and its distribution over the population as a function of age. The model is constructed using N training images I_n^{train}, $n \in \{1 \ldots N\}$, from a population-based cross-sectional data collection. This population-based model is then used as a reference to which an individual brain image I_i can be compared. The concept of the proposed method is shown in Fig. 1. The sections below explain all the steps of the proposed framework in detail: 2.1) preprocessing, 2.2) non-rigid groupwise image registration, 2.3) elimination of translation and rotation, 2.4) partial least squares regression, 2.5) percentile curve fitting, 2.6) spatio-temporal atlas construction, and 2.7) individual subject assessment.

Preprocessing

The preprocessing steps are 1) non-uniformity correction of the images using the N3 algorithm (Tustison et al., 2010) and 2) brain extraction using a multi-atlas method described in Bron et al. (2014), with a set of 30 atlases (Gousias et al., 2008; Hammers et al., 2003).

Non-rigid groupwise image registration

The morphological variation is characterized by deformation fields that map each brain image to a common domain, the template domain Ω_{template}. Let $\boldsymbol{x}$ be an image coordinate in $\mathbb{R}^3$. The function $I_n^{\text{train}}(\boldsymbol{x})$ gives

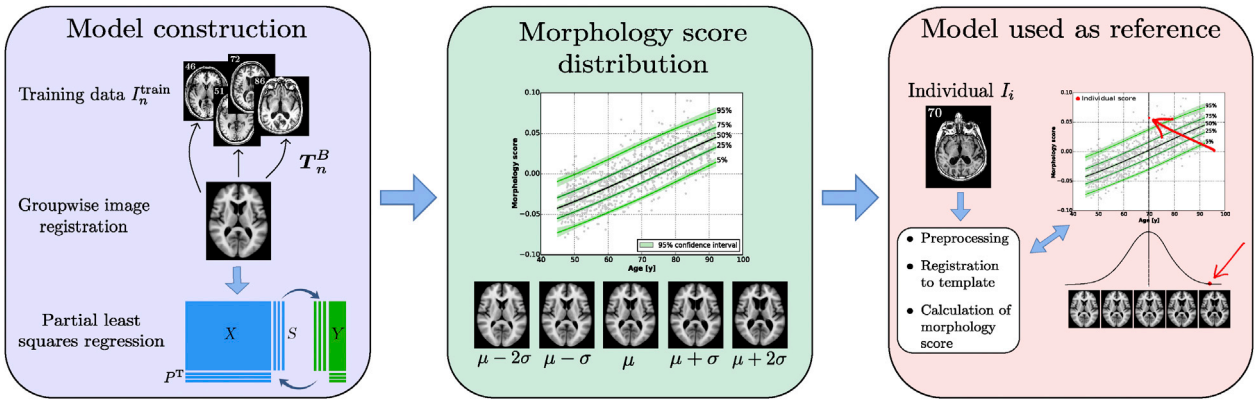


Fig. 1. Illustration of the proposed concept. Using training data I_n^{train} a model of brain morphology as a function of age is constructed. This model is represented by a set of percentile curves and a mode of deformation. Subsequently, an individual I_i is compared to the model that is used as reference.

the intensity of image n at $\boldsymbol{x}$, i.e. $I_n^{\text{train}}(\boldsymbol{x}) : \Omega_n \subset \mathbb{R}^3 \rightarrow \mathbb{R}$. Let N be the total number of images in the training set. The aim of the groupwise registration is to find a set of coordinate transformations $\boldsymbol{T}_n(\boldsymbol{x}) : \Omega_{\text{template}} \rightarrow \Omega_n$, $n \in \{1 \ldots N\}$, such that the warped images $I_n^{\text{train}}(\boldsymbol{T}_n(\boldsymbol{x};\boldsymbol{\mu}_n))$ are aligned with each other in the template domain. The image registration is performed with a parametric approach. The degrees of freedom of the transformation is limited by introducing a parameterization to the transformation: $\boldsymbol{T}_n(\boldsymbol{x};\boldsymbol{\mu}_n)$, where $\boldsymbol{\mu}_n$ is a vector containing the transformation parameters of subject n. In our method, these transformations are determined in two steps. First, we obtain a coarse alignment of all images in the training set via transformations $\boldsymbol{T}_n^A(\boldsymbol{x};\boldsymbol{\mu}_n^A)$. Then we use a non-rigid transformation model, $\boldsymbol{T}_n^B(\boldsymbol{x};\boldsymbol{\mu}_n^B)$, for a more precise alignment of the images. Fig. 2 shows an overview of all transformations. This section explains the transformations and domains shown in the top-half of this figure. As shown in the top-right part of Fig. 2, transformations $\boldsymbol{T}_n^A(\boldsymbol{x};\boldsymbol{\mu}_n^A)$ map $\boldsymbol{x}$ from MNI domain Ω_{MNI} (Mazziotta et al., 2001) to the subject-specific domains Ω_n. Second, as shown in the top-left part of Fig. 2, we obtain $\boldsymbol{T}_n^B : \Omega_{\text{template}} \rightarrow \Omega_{\text{MNI}}$, where $\boldsymbol{T}_n^B$ is parameterized by $\boldsymbol{\mu}_n^B$. The total transformation from Ω_{template} to Ω_n is a composition of $\boldsymbol{T}_n^A$ and $\boldsymbol{T}_n^B$.

The $\boldsymbol{T}_n^A$ are parameterized by an affine transformation model. The parameters $\boldsymbol{\mu}_n^A$ are found by performing a pairwise registration of the brain mask of each subject to the reference brain mask in the MNI domain.

The $\boldsymbol{T}_n^B$ are parameterized by cubic B-splines (Rueckert et al., 1999). We chose cubic B-splines, because their compact support property makes the computation efficient, which is relevant in our large-scale groupwise image registration problems. The spacing of the control points of the B-splines is a setting with which the degrees of freedom of the transformation can be controlled. The parameters of the B-spline transformation model, $\boldsymbol{\mu}_n^B$, are B-spline control point coefficients. To obtain $\boldsymbol{\mu}_n^B$, we use groupwise image registration. During such a registration, Ω_{template} is implicitly defined by constraining the sum of all deformations from the template to each subject to be zero, an approach proposed by Bhatia et al. (2004) and Balci et al. (2007). To achieve this, $\boldsymbol{\mu}_n^B$ must be optimized simultaneously for all $n \in \{1 \ldots N\}$. Advantages of groupwise registration are that the information of all images is taken into account during the registration and, as opposed to pairwise registration, the result is not biased towards any chosen reference image.

For the non-rigid groupwise image registration we use the method of Huizinga et al. (2016a). This method was designed for intra-subject registration of images originating from a quantitative MRI experiment.

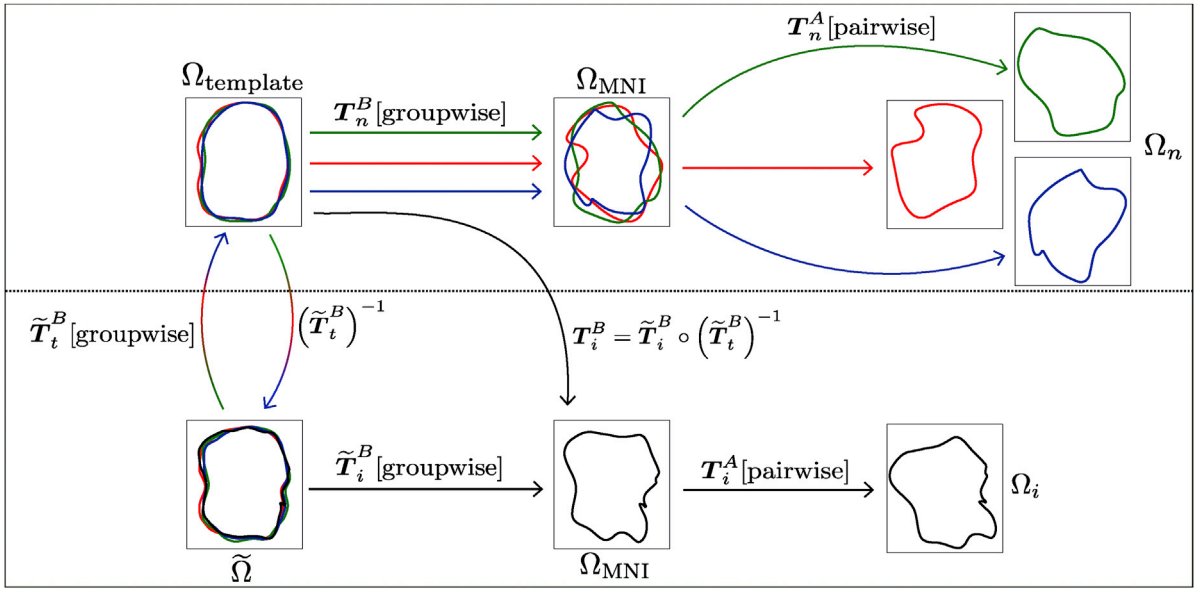


Fig. 2. Scheme showing the different transformations and domains of the registration framework of the proposed method. The top part shows three different training images in red, green and blue in their domains Ω_n, and how the template domain is constructed from the training data. The bottom half shows how an individual I_i in Ω_i is registered to the template domain. After each transformation is indicated if it was obtained using a pairwise or a groupwise registration.

It is therefore robust against arbitrary intensity scaling between the aligned images. In this work we investigate if the method is applicable to inter-subject registration as well. The method assumes that, when images are registered, the intensities can be mapped to a low-dimensional subspace. The dimension of this low-dimensional subspace depends on the model describing the intensity variation in the aligned images. When this subspace is assumed to be one-dimensional, the intensity may vary due to a global scale or shift, which is the case for our application. In this case, the method could be considered as an extension of normalized cross-correlation from pairwise ($N = 2$) to groupwise ($N > 2$) settings.

In our experiments, the image registration was performed with Elastix (Klein et al., 2010). We used a multi-resolution strategy with four resolutions in which the control point spacing of the B-spline transformation model is halved with each resolution step, until a final spacing of 10 mm. The final spacing was determined heuristically.

Individual registration to the template domain

To be able to compare the brain image of an individual, I_i, to the reference model, we need a deformation field that maps I_i to Ω_{template}. Similarly as during the template construction from training images, this is done in two steps. This section explains the transformations and domains shown in the bottom-half of Fig. 2. First, we register I_i to Ω_{MNI} yielding an affine transformation $\boldsymbol{T}_i^A$. Second, we seek a non-rigid transformation $\boldsymbol{T}_i^B$ that maps I_i from Ω_{MNI} to Ω_{template}.

A possible approach to finding $\boldsymbol{T}_i^B$ would be a pairwise registration of I_i with one of the registered training images in Ω_{template} chosen as a reference. However, this could introduce a bias towards this chosen reference. Another possibility would be to perform a pairwise registration of I_i to the mean of the registered training images, $\bar{I}(\boldsymbol{x}) = \frac{1}{N}\sum_n I_n^{\text{train}}(\boldsymbol{T}_n(\boldsymbol{x}))\Big)$ for $\boldsymbol{x} \in \Omega_{\text{template}}$, however, the disadvantage of such an approach is that $\bar{I}$ is blurry at the cortex edges which hampers accurate registration at those locations. Therefore, we use a different approach that is visualized in Fig. 2 and explained below.

We propose to formulate the non-rigid registration of I_i to the template domain as an additional non-rigid groupwise registration, involving $N + 1$ images, namely the N previously registered training images I_n^{train} and the individual image I_i. As shown in the bottom-left part of Fig. 2, this procedure leads to N transformations $\tilde{\boldsymbol{T}}_n^B(\boldsymbol{x};\tilde{\boldsymbol{\mu}}_n^B) : \tilde{\Omega} \rightarrow \Omega_{\text{template}}$ and one transformation $\tilde{\boldsymbol{T}}_i^B(\boldsymbol{x};\tilde{\boldsymbol{\mu}}_i^B) : \tilde{\Omega} \rightarrow \Omega_{\text{MNI}}$, where $\tilde{\Omega}$ is the common domain of I_n^{train} and I_i. Since the training images I_n^{train} had already been aligned before the registration, it is safe to assume that their transform parameters $\tilde{\boldsymbol{\mu}}_n^B$ associated with $\tilde{\Omega}$ are approximately equal. A single, unbiased transformation, from $\tilde{\Omega}$ to Ω_{template} $\tilde{\boldsymbol{T}}_t^B : \tilde{\Omega} \rightarrow \Omega_{\text{template}}$, is obtained by averaging $\tilde{\boldsymbol{\mu}}_n^B$ over all n, obtaining a single transform parameter vector $\tilde{\boldsymbol{\mu}}_t^B$:

$$\tilde{\boldsymbol{\mu}}_t^B = \frac{1}{N}\sum_n \tilde{\boldsymbol{\mu}}_n^B \tag{1}$$

To bring I_i to Ω_{template} we finally compute $\boldsymbol{T}_i^B : \Omega_{\text{template}} \rightarrow \Omega_{\text{MNI}}$ as:

$$\boldsymbol{T}_i^B = \tilde{\boldsymbol{T}}_i^B\left(\left(\tilde{\boldsymbol{T}}_t^B\right)^{-1}(\boldsymbol{x};\tilde{\boldsymbol{\mu}}_t^B);\tilde{\boldsymbol{\mu}}_i^B\right) \tag{2}$$

where $\boldsymbol{T}^{-1}(\boldsymbol{x};\boldsymbol{\mu})$ is the inverse of $\boldsymbol{T}(\boldsymbol{x};\boldsymbol{\mu})$, obtained using the procedure described in Metz et al. (2011). In total one pairwise registration, the affine registration to MNI space, and one groupwise registration, the non-rigid registration to Ω_{template}, are required to analyze a new image I_i.

Elimination of translation and rotation

As global brain shrinkage could be (partially) captured by the affine transformation $\boldsymbol{T}^A$, we consider the composition of the affine and non-rigid transformations, $\boldsymbol{T}^A(\boldsymbol{T}^B(\boldsymbol{x}))$. To focus on brain morphology only, we propose to extract the rigid body motion, like the arbitrary orientation of the subject's head in the scanner, from this composition. In this way, scale and skew transformations are preserved, and any rigid body motion present in $\boldsymbol{T}^B$ is also eliminated.

For any rigid transformation $\boldsymbol{T}_n^R$ parameterized by $\boldsymbol{\mu}_n^R$, the residual deformation is defined by:

$$\boldsymbol{d}_n^*(\boldsymbol{x};\boldsymbol{\mu}_n^R) = \boldsymbol{T}_n^R\left(\boldsymbol{T}_n^A\left(\boldsymbol{T}_n^B(\boldsymbol{x};\boldsymbol{\mu}_n^B);\boldsymbol{\mu}_n^A\right);\boldsymbol{\mu}_n^R\right) - \boldsymbol{x}. \tag{3}$$

The rigid transformation parameters of each subject are estimated by:

$$\widehat{\boldsymbol{\mu}}_n^R = \arg\min_{\boldsymbol{\mu}_n^R} \frac{1}{|\Omega_{\text{mask}}|}\sum_{\boldsymbol{x}\in\Omega_{\text{mask}}} \left\|\boldsymbol{d}_n^*(\boldsymbol{x};\boldsymbol{\mu}_n^R)\right\|^2, \tag{4}$$

where Ω_{mask} is the domain containing voxels inside the brain mask in Ω_{template} to ensure that we only evaluate the deformation inside the brain. In the statistical analysis that follows we will be using $\boldsymbol{d}_n^*(\boldsymbol{x}) \widehat{=} \boldsymbol{d}_n^*(\boldsymbol{x};\widehat{\boldsymbol{\mu}}_n^R)$ for $\boldsymbol{x} \in \Omega_{\text{mask}}$. The same approach is applied to the individual deformation field $\boldsymbol{d}_i$.

Partial least squares regression

We aim to correlate the deformations obtained by image registration with age. Let N be the number of subjects in the training set and let $|\Omega_{\text{mask}}| = M$, then $\boldsymbol{X}$ is the $N \times 3M$ matrix of which each row contains $\boldsymbol{d}_n^*(\boldsymbol{x})$ for all $\boldsymbol{x} \in \Omega_{\text{mask}}$. Let $\boldsymbol{Y}$ be the $N \times 1$ vector containing the corresponding ages of each subject, then PLSR finds the directions in the deformation space that explain the maximum covariance with age. Let $\boldsymbol{X}_0 = (\boldsymbol{X} - \overline{\boldsymbol{X}})$ and let $\boldsymbol{Y}_0 = (\boldsymbol{Y} - \overline{\boldsymbol{Y}})$, where $\overline{\boldsymbol{X}}$ and $\overline{\boldsymbol{Y}}$ are the column-wise data means replicated to all N rows. In PLSR $\boldsymbol{X}_0$ is decomposed into:

$$\boldsymbol{X}_0 = \boldsymbol{S}\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{E}. \tag{5}$$

Let L be the number of components used in the PLSR, then $\boldsymbol{S}$ are the $N \times L$ scores and $\boldsymbol{P}$ are the $3M \times L$ orthonormal loadings. Matrix $\boldsymbol{E}$ is the $N \times 3M$ residual matrix. The decomposition of $\boldsymbol{X}_0$ is made by maximizing the covariance between $\boldsymbol{Y}_0$ and a weighted sum of $\boldsymbol{X}_0$. The weight vector $\boldsymbol{W}^j$ for each component $1 \leq j \leq L$ is estimated by solving the following optimization problem:

$$\max_{\boldsymbol{W}^j}\left[(\boldsymbol{X}_0\boldsymbol{W}^j)^{\mathrm{T}}\boldsymbol{Y}_0\right] \tag{6}$$

subject to $(\boldsymbol{X}_0\boldsymbol{W}^j)^{\mathrm{T}}\boldsymbol{X}_0\boldsymbol{W}^j = 1$. The weight vector is found by a singular value decomposition on $\boldsymbol{R} = \boldsymbol{X}_0^{\mathrm{T}}\boldsymbol{Y}_0$ (Höskuldsson, 1988; de Jong, 1992). The PLSR scores for component j are defined as $\boldsymbol{S}^j = \boldsymbol{X}_0\boldsymbol{W}^j$, and the loadings of component j are defined as $\boldsymbol{P}^j = \boldsymbol{X}_0^{\mathrm{T}}\boldsymbol{S}^j$. Covariance $\boldsymbol{R}$ is deflated with each iteration of the algorithm to obtain $\boldsymbol{W}^j$, $\boldsymbol{S}^j$ and $\boldsymbol{P}^j$ for $j \geq 2$.

The number of components used in PLSR is a tradeoff between overfitting (too many components) or losing valuable information (not enough components). To determine the optimal number of components we propose to use the randomization test of Wiklund et al. (2007). In this test a null-distribution of $(\boldsymbol{S}^j)^{\mathrm{T}}\boldsymbol{Y}_0$ is determined by randomly permuting $\boldsymbol{Y}$. When the probability of finding the observed $(\boldsymbol{S}^j)^{\mathrm{T}}\boldsymbol{Y}_0$ is smaller than α, the component is significant. Since PLSR is prone to overfitting, we chose a conservative significance level of $\alpha = 0.01$. Results of the randomization test showed that only the first PLSR component was significant (see Sec. 4.1). We therefore only used the scores of the first PLSR component, $\boldsymbol{S}^1$, to describe the brain morphology distribution as a function of age, and from here on, we omit the superscript 1.

To compare an individual subject to the reference population, the morphology score of this individual subject, S_i, is required. To this end, the individual deformation field $\boldsymbol{d}_i^*$ is projected on $\boldsymbol{W}$ to obtain S_i:

$$S_i = \left(\boldsymbol{d}_i^* - \overline{\boldsymbol{X}}\right)\boldsymbol{W}. \tag{7}$$

where $\overline{\boldsymbol{X}}$ is a vector of size $3M$ with the column-wise means of $\boldsymbol{X}$.Since an individual's brain morphology may not only change due to aging but also due to pathology we also propose to compare the individual's residual to that of the model. Each row of the residual matrix $\boldsymbol{E}$ contains the non-age-related deformations for a subject in I_n^{train}. The residual norm of I_n^{train} is defined as:

$$\|\boldsymbol{E}\|_n = \sqrt{\sum_{m=1}^{3M} E_{nm}^2} \tag{8}$$

where E_{nm} is element (n, m) of matrix $\boldsymbol{E}$. The residual of an individual is defined as:

$$\boldsymbol{E}_i = \boldsymbol{d}_i^* - S_i \boldsymbol{P}^{\mathrm{T}} \tag{9}$$

Inspection of $\boldsymbol{E}_i$ is important since an individual could have a different morphology due to other factors than aging. If this is the case, it would not be visible by the individual's morphology score. We therefore propose to compare the individual residual norm $\boldsymbol{E}_i$ to the distribution of $\boldsymbol{E}_n$ in I_n^{train}. If $\boldsymbol{E}_i$ is significantly different from the distribution of $\boldsymbol{E}_n$ further inspection of $\boldsymbol{d}_i^*$ is necessary.

To perform the PLSR, we implemented the SIMPLS algorithm (de Jong, 1992) in Python.

Percentile curve fitting

The score of subject n in I_n^{train} is referred to as S_n. To visualize the distribution of scores S_n as function of age we fit percentile curves. These curves show both the distribution of the morphology scores and how they vary with age. We refer to the p^{th}% percentile curve at age a as $s(a, p)$.

For fitting of percentile curves to the morphology score data, we use the LMS method (Cole and Green, 1991). The LMS method assumes that the data is standard normally distributed after applying the Yeo-Johnson transformation, which is an extension of the Box-Cox transformation proposed by Cole and Green (1991). This method estimates the $\lambda-$parameter of the Yeo-Johnson transformation (Yeo and Johnson, 2000) (L), the median (M) and coefficient of variation (S) for the appropriate morphology score at each age. With the parameters L, M, and S, percentiles can be computed at each age to obtain a smooth curve. The smoothness of the fitted curves is influenced by the degrees of freedom δ, a user-defined parameter. In our experiments, we set the smoothness parameter δ to a value of 2 and we deployed the R-package VGAM (Yee, 2010) for the percentile curve fitting.

The value of the morphology score may also be influenced by other covariates than age, e.g. sex or height, since those covariates may influence the head size and may therefore affect brain scaling in the deformation fields that are used in the PLSR. It is therefore desirable to correct the reference distribution for these covariates. We model the correction for these two covariates as a linear shift in the morphology score distribution.

The precision of the estimated percentile curves depends on the number of datapoints in the appropriate age range. If the data is non-uniformly distributed over age, it could be that the curve estimation is not precise in the part where there are very few datapoints. To assess the precision of the fitted curves, we use a bootstrapping procedure, by random sampling subjects with replacement and re-estimating the percentile curves. A distribution of possible curves is collected, from which confidence intervals at any significance level can be estimated (Carpenter and Bithell, 2000).

Spatiotemporal atlas construction

The deformation having the highest covariance with age is contained in the loading vector $\boldsymbol{P}$. To be able to interpret the morphology score it is necessary to know what this deformation looks like, and therefore we aim to visualize $\boldsymbol{P}$ as a spatio-temporal atlas. An estimate of the age-related morphology of subject n is obtained by multiplying score S_n with $\boldsymbol{P}$. Instead of multiplying with just one specific score S_n, we multiply $\boldsymbol{P}$ with $s(a, p)$ for a chosen range of a and a specific p. We choose to show the aging trajectory of $\bar{I}$. We convert the deformation field in $s(a, p)\boldsymbol{P}$ to B-spline transformations and invert these to obtain $\boldsymbol{T}_{\text{atlas}}^{a,p}(\boldsymbol{x})$. The spatio-temporal atlas for percentile p and a chosen range of a is then constructed by warping $\bar{I}$ for each a:

$$\bar{I}^{a,p}(\boldsymbol{x}) = \bar{I}(\boldsymbol{T}_{\text{atlas}}^{a,p}(\boldsymbol{x}; \boldsymbol{\mu})), \tag{10}$$

Individual subject assessment

Given the distribution of morphology scores and residual magnitudes of the reference data, outliers from these distributions can be assessed in terms of percentiles, which is analogous to the use of growth charts to map child development in terms of height and weight as a function of age (de Onis et al., 2006). Let p_i be the percentile at which the individual morphology score S_i at age a_i can be found in the reference morphology score distribution, after correction for available covariates. Let $p_{\|\boldsymbol{E}\|_i}$ be the percentile at which the individual residual magnitude can be found in the reference residual magnitude distribution, then the status of the brain morphology of patient i could, for example, be assessed according to the following set of rules:

1. No accelerated brain aging: $0.05 < p_i < 0.95$ and $0.05 < p_{\|\boldsymbol{E}_i\|} < 0.95$
2. At risk of accelerated brain aging: $p_i > 0.95$ and $0.05 < p_{\|\boldsymbol{E}\|_i} < 0.95$
3. At risk of unknown pathology: $p_{\|\boldsymbol{E}\|_i} > 0.95$

In addition, the individual scan I_i should be compared to the spatio-temporal atlas $\bar{I}^{a_i,p_i}$, for a qualitative assessment of the morphology due to aging. In the third case, further investigation of I_i, $\boldsymbol{E}_i$, and the individual deformation $\boldsymbol{d}_i^*$ is necessary.

Experiments

In our experiments we used two image databases:

1. The Rotterdam Scan Study (RSS), containing brain scans of non-demented, asymptomatic subjects,
2. The Alzheimer's Disease Neuroimaging Initiative (ADNI) database, containing both asymptomatic and symptomatic subjects.

First, the model was applied and validated on the RSS database. Next, the difference in scores between asymptomatic and symptomatic subjects was evaluated in the ADNI dataset. Finally, the robustness of the scores across databases was evaluated.

Data

Rotterdam Scan Study

We used 988 T1w scans (433 male, age = 68.3 ± 13.0 (mean ± SD)) from the population-based RSS, a prospective longitudinal study among community dwelling subjects aged 45 years and over (Ikram et al., 2015). Participants with dementia at the time of MRI were excluded (Schrijvers et al., 2012). All brain scans were acquired on a single 1.5T MRI system (GE Healthcare, US). The T1w imaging protocol was a 3-dimensional fast radiofrequency spoiled gradient recalled acquisition in steady state with an inversion recovery pre-pulse sequence (Ikram et al., 2015). The voxel size was $0.5 \times 0.5 \times 0.8$ mm^3. Besides age and sex, height information of the partipants was available as well. For six participants, the height variable was missing, for which we substituted the average height (= 170 cm). We will refer to this dataset as RSS988.

ADNI

Data used in the preparation of this article were obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimers disease (AD). For up-to-date information, see www.adni-info.org.

The ADNI cohort used in this article is adopted from the study of Cuingnet et al. (2011), consisting of an AD patient group, an MCInc group (mildly cognitive impaired but not converted to AD within 18 months), an MCIc (mildly cognitive impaired and converted to AD within 18 months), and a cognitive normal group (CN). The inclusion criteria for participants were defined in the ADNI-GO protocol.[3] The AD group consisted of 137 patients (67 male, age = 76.0 ± 7.3 years) (AD137), the MCInc group of 134 participants (84 male, age = 74.4 ± 7.2 years) (MCInc134), the MCIc group of 76 participants (43 male, age = 74.7 ± 7.4 years) (MCIc76), and the CN group of 162 participants (76 male, age = 76.2 ± 5.4 years) (CN162). Acquisition had been performed according to the ADNI acquisition protocol (Jack et al., 2008). The brain scans were acquired on 1.5T MRI systems (GE Healthcare, Philips Medical Systems, Siemens Medical Solutions) and the T1w imaging protocol was a 3-dimensional magnetization prepared rapid acquisition gradient echo sequence. The voxel size was approximately 1 mm^3, with a maximum of 1.5 mm in any direction. Besides age and sex, height information of the partipants was available as well. For one participant, the height variable was missing, for which we substituted the average height of the subjects in RSS988 (= 170 cm). We will refer to the entire dataset as ADNI509.

Morphology distribution RSS988

We trained the model on RSS988 to visualize the main age-related deformations in a healthy reference population. Sex and height were used as covariates.

After the preprocessing step, the brain images were cropped to the bounding box of the mask and resampled to $1.5 \times 1.5 \times 1.5$ mm^3 spacing. The 988 obtained deformation fields $\boldsymbol{d}^*$, which are used as input of the PLSR, were downsampled to a $3 \times 3 \times 3$ mm^3 spacing and their size was $52 \times 66 \times 55$ voxels. Since the number of parameters with which the deformation field was generated was much smaller than the number of voxels in the downsampled field, the downsampling will not influence the result.

The number of PLSR components was determined by applying the randomization test to RSS988, using 1000 randomizations of $\boldsymbol{Y}$. In addition, we inspected the score distributions, the explained variance in age, and deformation modes (PLSR loadings) of the first ten components.

Morphology distribution ADNI509

We trained the model ADNI509 and fitted percentile curves on the scores from each class separately to see if the morphology distribution is different for the various classes in the ADNI database. Sex and height were used as covariates.

After the preprocessing step, the brain images were cropped to the bounding box of the mask and resampled to $1.5 \times 1.5 \times 1.5$ mm^3 spacing, to reduce computation time and memory consumption of the groupwise registration. The obtained 509 deformation fields, which are used as input for the PLSR, were downsampled to a $3 \times 3 \times 3$ mm^3 spacing and their size was $59 \times 66 \times 55$ voxels.

[3] http://www.adni-info.org/Scientists/doc/ADNI_Go_Protocol.pdf.

Model validation

We validated the model using data from the RSS database, in which we trained the model on a subset of 888 images, RSS888, and tested on the remaining 100 images, RSS100. We also validated the model on data from different databases, in which we trained the model on RSS988 and tested on the AD137 and CN162.

Leave-100-out validation (1)

The groupwise image registration on RSS988 resulted in 988 brain deformation fields. In this experiment we evaluate the generalizability of the PLSR outside the training data. Therefore we randomly selected 888 deformation fields as input of the PLSR. The resulting PLSR weights are used to compute individual scores on the remaining 100 deformation fields using Equation (7).

Leave-100-out validation (2)

In this experiment we trained the entire model only on RSS888: I_n^{train} = RSS888. The remaining 100 subjects were treated as entirely new individuals, I_i = RSS100. RSS100 was registered to Ω_{template}, which was constructed from RSS888, and 100 S_i were computed. We evaluated if the scores from the leave-100-out experiment (1) could be reproduced.

Individual subject comparison

In this experiment the training dataset was I_n^{train} = RSS988, and the individual subjects were I_i^{CN} = CN162 and I_i^{AD} = AD137, respectively. We compared the individual scores S_i^{CN} and the S_i^{AD} to the morphology score distribution of RSS988, while taking into account the covariates sex and height. The goals of this experiment were:

- To evaluate if the individual comparison can be performed when individual subjects are scanned on different scanners with different scanning protocols.
- To evaluate if healthy subjects from different populations have the same brain morphology distribution.
- To evaluate if the AD subjects have different morphology scores than the healthy subjects from a different population.

Results

Morphology distribution RSS988

Results from the randomization test indicated that only the first component was significantly different from the null-distribution ($p = 2.6{\cdot}10^{-10}$). The second component was not significant ($p = 1.4{\cdot}10^{-2}$). Fig. 3 shows the explained variance in age of the first ten PLSR components. This figure shows that the first component explains most variance in age (∼60%) and that the following components do not add much information, which is in agreement with the results of the randomization test. The score distributions of components two to ten showed very little to no relation with age and the deformation modes did not contain clear patterns that can be expected in aging. We therefore only used the scores of the first PLSR component to describe the brain morphology distribution as a function of age.

Fig. 4(a) shows S_n of the first PLSR component and the fitted percentile curves $s(a,p)$ for $p \in \{5, 25, 50, 75, 95\}$, for male subjects of height 170 cm, the average height of all subjects in RSS988. Fig. 4(b) shows S_n of the first PLSR component and the fitted percentile curves for female subjects of height 150 cm, to show the effect of the covariates on the morphology score distribution. The distributions have a clear relation with age. As expected, the scores of short, female subjects are higher than the scores of male subjects of average height. Fig. 4(c)-(g) shows the main mode of deformation in $\boldsymbol{P}$ applied to $\bar{I}$ for RSS988. This mode shows that the higher S_n, the larger the ventricles, (cortical) atrophy, and brain shrinkage.

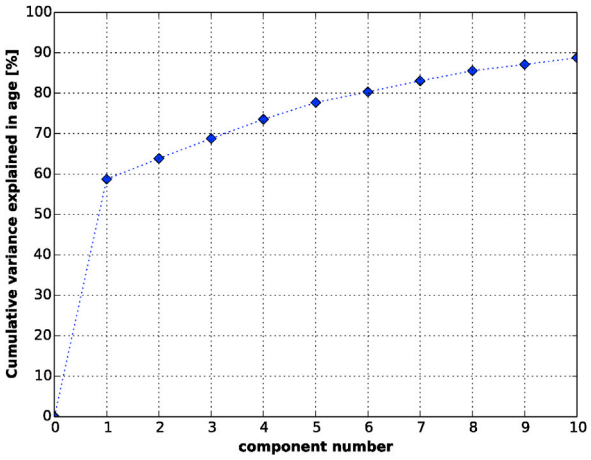


Fig. 3. Cumulative explained variance [%] of the first ten PLSR components.

Morphology distribution ADNI509

Fig. 5 shows S_n and $s(a,p)$ for $p \in \{5, 25, 50, 75, 95\}$, adjusted for covariates sex and height, of the model trained on ADNI509. The scores of CN162 are the lowest, and MCIc76 and AD137 have the highest scores. Interestingly, a morphology difference between MCInc134 and MCIc76 is visible. The confidence bounds of the median curves of AD137 and CN162 do not overlap, indicating that the median curve of the two groups are significantly different. The confidence bounds of the median curves of MCInc134 and MCIc76 group overlap slightly, possibly due to the lower number of subjects in MCIc76.

Distributions of $\|\boldsymbol{E}\|$ in ADNI509

The distributions of $\|\boldsymbol{E}\|$ are shown in Fig. 5(f). The Welch's two-sample t-test was performed to test if the distributions of $\|\boldsymbol{E}\|$ are significantly different. Tests were performed between all possible group pairs. All p-values were higher than the significance level of 0.05, and therefore no significant difference between the distributions was observed.

Model validation

Leave-100-out validation (1)

Fig. 6 shows the morphology scores for RSS888 gray and RSS100 test subjects in red. The morphology scores of the test subjects fall within the morphology score distribution of the training subjects.

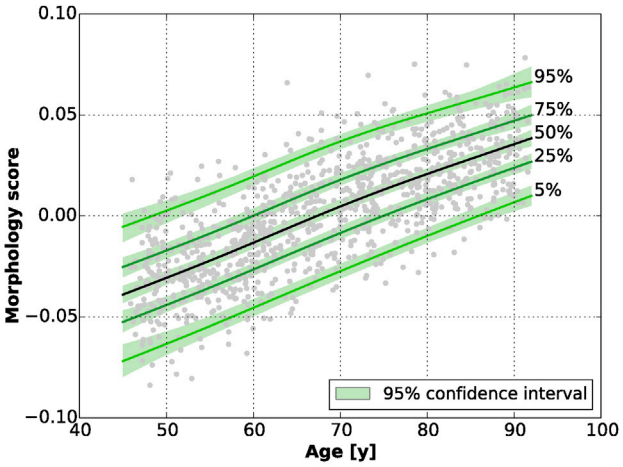
Leave-100-out validation (2)

Fig. 7 shows a scatterplot of the morphology scores of the leave-100-out (1) experiment versus the leave-100-out (2) experiment. The morphology scores show a high correlation (Pearson's $r = 0.996$), but a small bias is present. The distribution of $\boldsymbol{E}$ of RSS888 is not significantly different from the distribution of $\boldsymbol{E}$ of RSS988 according to the Welch's two-sample t-test, as shown in Fig. 8.

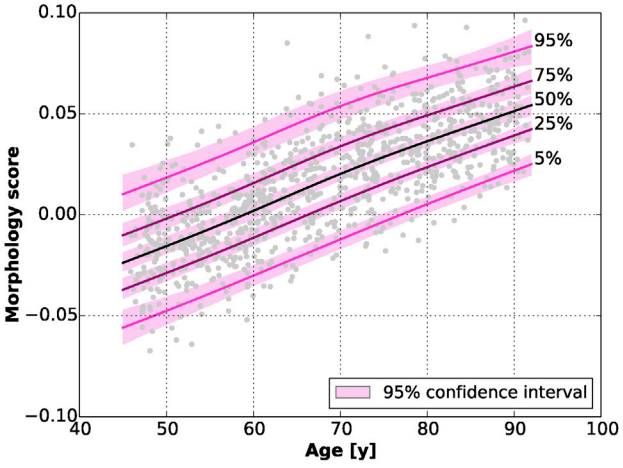
Individual subject comparison

Fig. 9 shows the S_i^{CN} and S_i^{AD} projected onto the model of RSS988 (male, 170 cm). Fig. 9(a) shows that 94% of the CN162 lie below the 95% percentile lines of the RSS distribution. From Fig. 9(b) we see that 34% of AD137 lie above the 95% percentile, and 76% above the 75% percentile, indicating that their morphology shows more atrophy than that of cognitive normals at the same age.

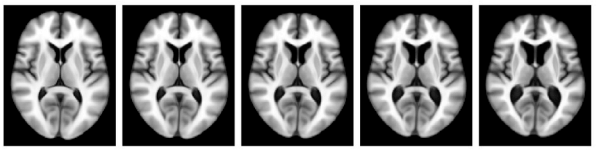
Fig. 10 shows the distributions of $\|\boldsymbol{E}\|$ of RSS988 subjects and CN162 (a) and AD137 (b). The distributions have overlap, but are significantly different according to the Welch's two-sample t-test.



(a) Morphology score distribution for RSS988 (male, 170 cm).



(b) Morphology score distribution for RSS988 (female, 150 cm).



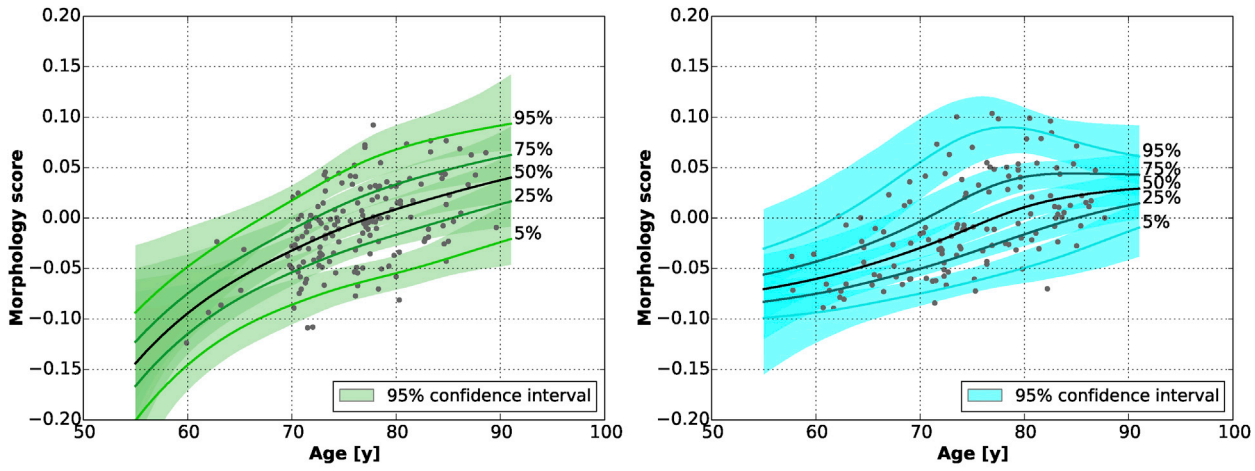
(c) −0.064 (d) −0.032 (e) 0 (f) 0.032 (g) 0.064

Fig. 4. Morphology scores S_n for all n and $s(a,p)$ for $p \in \{5, 25, 50, 75, 95\}$ for (b) male subjects of height 170 cm of RSS988, and (b) female subjects of height 150 cm of RSS988. The 95% confidence intervals were determined with 1000 bootstraps. (c)–(g) The main mode of deformation where S equals -2σ, $-\sigma$, 0, σ, and 2σ, with σ the standard deviation of S_n for $n \in \{1 \ldots 988\}$.

Discussion

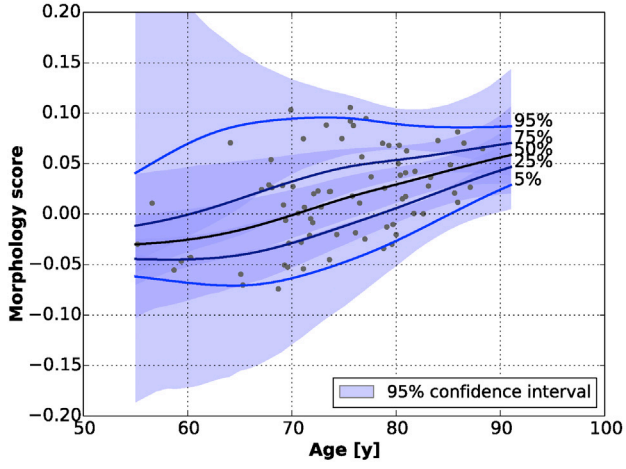
We proposed a method for developing a spatio-temporal model of morphological differences in the brain due to normal aging, to which an individual's brain morphology can be compared. We applied the framework to a set of 988 images of non-demented aging subjects from a large population imaging study and on a set of 509 subjects from the case-control study ADNI. The main mode of deformation due to aging, $\boldsymbol{P}$, shows the expected deformation patterns due to aging: larger ventricles, (cortical) atrophy, and brain shrinkage. We performed various experiments to validate the proposed method. The results of these experiments are encouraging and show clinical potential.

The leave-100-out experiment (1) showed that the PLSR generalizes for subjects that were not part of the training set. The leave-100-out experiment (2) showed that we can reproduce the morphology scores

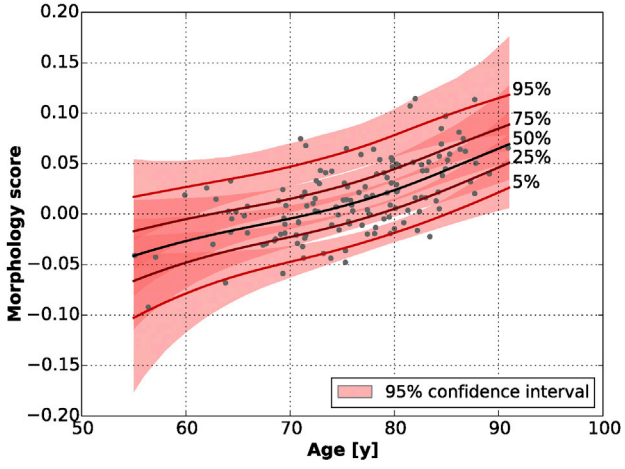


(a) Morphology score distribution of CN162 (male, 170 cm).

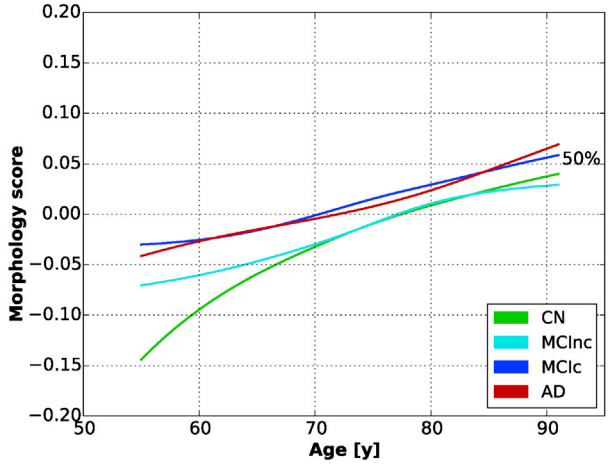
(b) Morphology score distribution of MCInc134 (male, 170 cm).



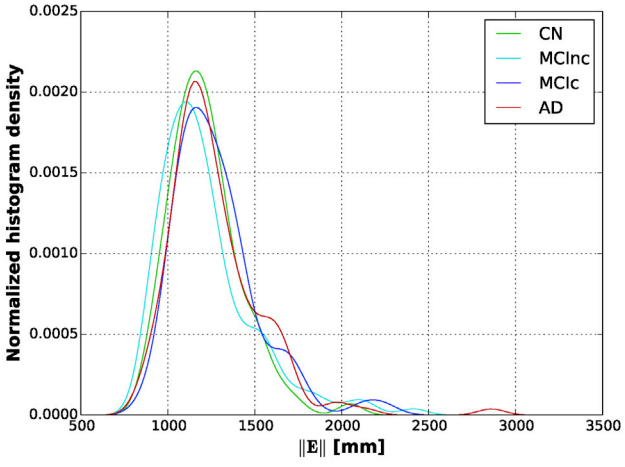
(c) Morphology score distribution of MCIc76 (male, 170 cm).



(d) Morphology score distribution of AD137 (male, 170 cm).



(e) Median percentile all groups in ADNI509 (male, 170 cm).



(f) Distributions of $\|\boldsymbol{E}\|$ of the groups in ADNI509.

Fig. 5. In figures (a)-(d) we see the morphology score distribution of each subgroup in ADNI509. The 95% confidence intervals are estimated using 1000 bootstraps. At the age extremities not many data points are available resulting in wide confidence intervals. For comparison of the four subgroups, Figure (e) shows the median percentile of each subgroup in one graph, without confidence intervals. This graph shows that the CN and MCInc group have similar morphologies, and that the AD and MCIc group have similar morphologies. The scores of the latter two are generally higher than the scores of the CN/MCInc group. Figure (f) shows the distributions of $\boldsymbol{E}$ of CN162, MCInc134, MCIc76 and AD137.

from leave-100-out experiment (1). This indicates that the deformations resulting from the individual subject registration to the template are similar to the deformations following from the groupwise registration, and that the method can be used to compute individual morphology scores. We did observe a slight bias when comparing the scores of leave-100-out experiment (1) and leave-100-out experiment (2). This is due to the fact that the starting point of the registration in leave-100-out experiment (1) is different from the starting point of the registration in leave-100-out experiment (2), leading to slightly different deformation fields.

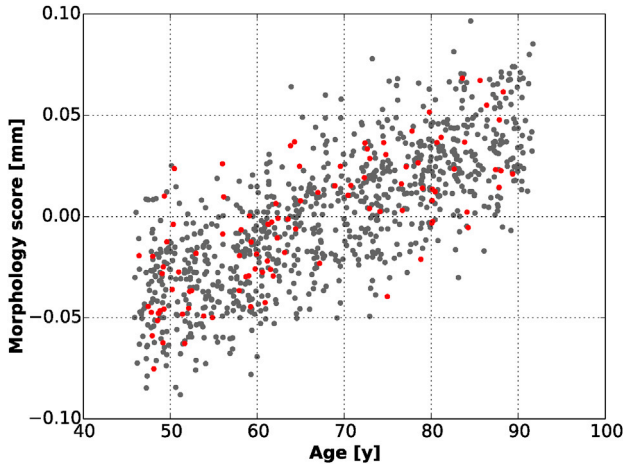


Fig. 6. Morphology scores for 888 training subjects in gray and 100 test subjects in red, obtained in leave-100-out validation (1).

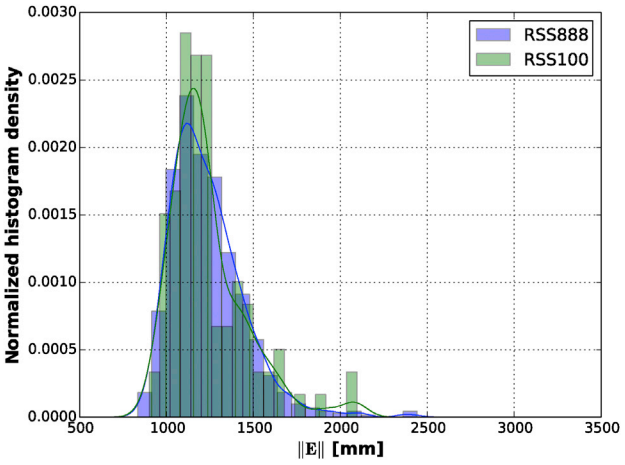


Fig. 8. Distributions of $\|\boldsymbol{E}\|$ of training set RSS888 (blue) and test set RSS100 (green).

The percentile curves computed on the ADNI database showed a difference between the morphology score distributions of the CN and AD subgroups (median difference significant, i.e. no overlapping confidence bands, for the age range 60–72y), and, interestingly, also between MCInc and MCIc (median difference not significant, i.e. confidence bands overlap). Although it was not specifically trained to separate the groups, the model showed that the AD and MCIc patients have, on average, higher morphology scores than the CN and MCInc groups. The distributions of the norm of the PLSR residual, $\boldsymbol{E}$, between the four groups was not significantly different, indicating that the non-age related deformations were of similar magnitude.

To evaluate the reproducibility across datasets we computed scores of both AD and cognitive normal individuals from the ADNI database using the RSS subjects as reference data. The cognitive normal subjects from ADNI fell nicely within the morphology score distribution of the cognitive normal subjects from RSS. The AD subjects had on average higher morphology scores than the cognitively normal subjects. These results suggest that the proposed method is able to compute individual morphology scores of subjects from one population when trained on subjects from another population. We did observe that $\|\boldsymbol{E}\|$ is different for the two databases of scans, which we did not see when registering subjects from the same population. Possible causes for this are the presence of non-age-related morphology differences between the two populations, e.g. differences in scanner type or protocol, inclusion criteria, demographics, environmental factors, etc.

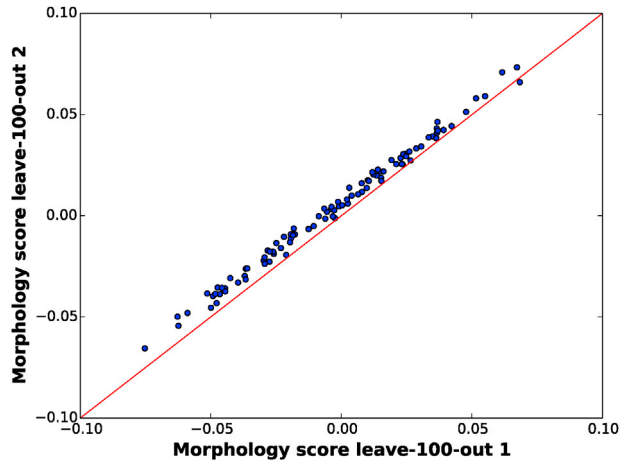


Fig. 7. Morphology scores of the 100 test subjects from leave-100-out validation (1) vs the scores from leave-100-out validation (2).

Suggestions for future work

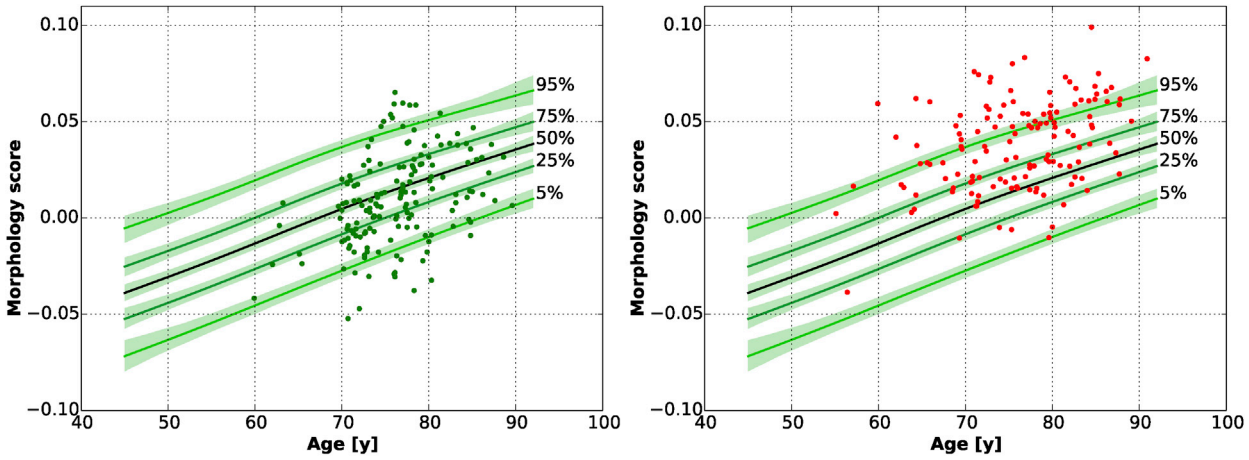
Several improvements can be made to the proposed model to increase sensitivity and performance. Also, the proposed method can be used in a variety of applications.

The uncertainty in the percentile curves was quantified by bootstrapping, obtaining confidence intervals for each curve. The uncertainty on individual morphology scores due to registration errors was not taken into account in the presented method. However, this could be estimated by perturbation of the individual deformation field, according to estimations of registration uncertainty. In this way, a distribution of possible morphology scores for this individual could be obtained. Obtaining reliable estimates of registration uncertainty is, however, still an active topic of research (Folgoc et al., 2016; Sokooti et al., 2016; Kybic, 2010).

Overall, the experiments showed that the proposed method is valid when comparing individual brain morphologies to a (cognitively) healthy reference population. Diagnostic value of the morphology score alone, however, is limited due to the high variability between individuals. In this study, we applied our method to T1-weighted MRI brain scans, but in principle, it could be applied to scans of other sequences or modalities. Future work may investigate age and other demographics related changes in other voxelwise maps, such as diffusion or perfusion imaging derived maps. This can be accomplished by replacing the deformation field $\boldsymbol{d}^*$ with such a voxelwise map in the template domain. This could possibly improve discriminative ability between people at risk of accelerated brain aging and people without risk.

Besides application of this method to voxelwise maps, it can also be applied to selected points on the surface of a segmented anatomical structure of interest, for example the hippocampus. In that case, a spatio-temporal model and its variation in the population of that specific structure can be studied in more detail and individual anatomical structures can be compared to a reference shape distribution.

It is clinically relevant to follow how a person's brain morphology changes with age. The proposed method can be used to compute morphology scores of individual scans at baseline and at follow-up. Plotting the baseline and follow-up scores in the percentile curves allows comparing the aging trajectory of an individual to the (cross-sectional) reference population. Following individual morphology scores over time allows estimation of the physiological variation in subject specific aging trajectories. On an individual level, inspection of these aging trajectories could be relevant in clinical trials to, for example, investigate if changes in e.g. medication or lifestyle have any effect. A clinical evaluation of the proposed deformation-based framework would be necessary to prove its value in the clinic.



(a) S_i^{CN} are shown in green. (b) S_i^{AD} are shown in red.

Fig. 9. Morphology score distribution (male, 170 cm) of RSS988 in five percentile curves and S_i^{CN} (a) and S_i^{AD} (b). The S_i^{CN} and S_i^{AD} are corrected and projected onto the reference distribution.

Limitations

A limitation of using cross-sectional data, which was also mentioned by (Ziegler et al., 2012), is that effects of different birth cohorts are not excluded.

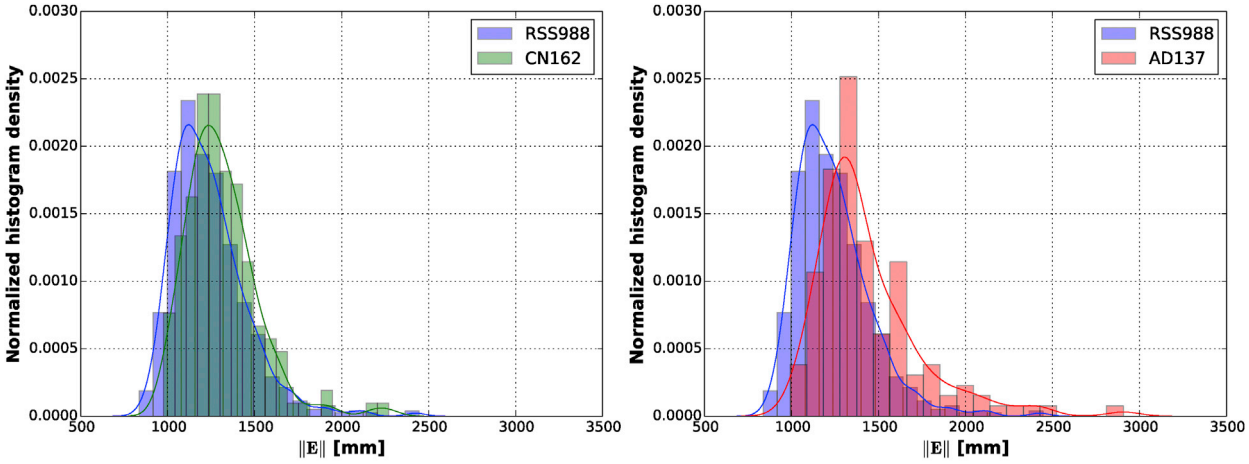
Because PLSR tries to maximize the covariance between the morphology scores and age, and the number of predictor variables is high ($= 3M \gg N$), it is very likely that it will find a linear relationship. Therefore, when modeling a population, and the relation with age is not expected to be linear, this may be a limitation of PLSR.

To compute the score of an individual brain image an additional groupwise registration of all images in the training set plus the individual brain image(s) has to be performed. This additional groupwise registration is time-consuming and therefore this is a practical limitation of our proposed framework. In the current implementation, the registration takes about two days and has to be performed on a machine capable of reserving 100 GB of RAM to complete the task. In the future we will search for possibilities to register an individual brain image to the template domain without having to perform an additional groupwise registration with the already registered training data. In the presented work, however, the groupwise registration is preferred over a pairwise registration for obtaining an unbiased result. This is supported by preliminary experiments which are not shown due to lack of space.

Conclusion

We developed a spatio-temporal model of morphological differences in the brain due to normal aging. The method provides a representative distribution of brain morphologies as a function of age instead of a single population mean morphology. Our method reduces high dimensional morphology to a single score. This score can be interpreted using the spatio-temporal atlas showing which deformation due to aging belongs to that score.

The framework was tested using data from two different datasets. Experiments showed that the proposed method extracts the expected deformation patterns due to aging and they showed that on a group level there is a morphology difference between cognitively normal and AD subjects, which manifests as accelerated aging, indicating the potential at least for clinical group studies. The spatio-temporal model can be used to compare an individual's brain morphology to a cognitively healthy reference population. Smooth percentile curves showing the brain morphology changes as a function of age as well as spatio-temporal atlases derived from the cognitively healthy reference population (RSS988) are publicly available via an interactive web application at www.agingbrain.nl. We believe that this framework has the potential to be used clinically as an indicator of accelerated brain aging.



(a) Normalized histogram of $\|\boldsymbol{E}\|$ for RSS988 and CN162 (b) Normalized histogram of $\|\boldsymbol{E}\|$ for RSS988 and AD137

Fig. 10. Distributions of $\boldsymbol{E}$ where the RSS images are I_n^{train} for $n \in \{1, \ldots, 988\}$ and the ADNI images were I_i^{CN} for $i \in \{1, \ldots, 162\}$ and I_i^{AD} for $i \in \{1, \ldots, 137\}$.

Conflicts of interest

Wiro J. Niessen, PhD, co-founder, chief scientific officer and shareholder of Quantib BV, which develops software to extract quantitative imaging biomarkers from medical imaging data.

Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme FP7/2007 - 2013, project VPH-DARE@IT (grant agreement no: 601055) and from the European Unions Horizon 2020 research and innovation programme, project DynaHEALTH (grant agreement no: 633595), project EuroPOND (grant agreement no: 666992), and from the European Research Council (ERC), project ORACLE (grant agreement no: 678543).

This research is supported by the Dutch Technology Foundation STW (grant agreement no: 12723), which is part of the NWO, and which is partly funded by the Ministry of Economic Affairs.

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

References

Achterberg, H., van der Lijn, F., den Heijer, T., van der Lugt, A., Breteler, M., Niessen, W., de Bruijne, M., 2010. Prediction of dementia by hippocampal shape analysis. In: MICCAI 2010: Medical Image Computing and Computer-assisted Intervention, Machine Learning in Medical Imaging. Springer, pp. 23–30.

Balci, S., Golland, P., Shenton, M., Wells, M., 2007. Free-form B-spline deformation model for groupwise registration. In: Proceedings of Medical Image Computing and Compututer-assisted Intervention. Springer, pp. 23–30.

Baloch, S., Davatzikos, C., 2009. Morphological appearance manifolds in computational anatomy: groupwise registration and morphological analysis. NeuroImage 45, S73–S85.

Bhatia, K., Hajnal, J.V., Puri, B., Edwards, A., Rueckert, D., 2004. Consistent groupwise non-rigid registration for atlas construction. In: Proc. IEEE Int Symp on Biomed Imaging: Nano to Macro, pp. 908–911.

Brewer, J., 2009. Fully-automated volumetric MRI with normative ranges: translation to clinical practice. Behav. Neurol. 21, 21–28.

Bron, E., Steketee, R., Houston, G., Oliver, R., Achterberg, H., Loog, M., van Swieten, J., Hammers, A., Niessen, W., Smits, M., Klein, S., 2014. Diagnostic classification of arterial spin labeling and structural MRI in presenile early stage dementia. Hum. Brain Mapp. 35, 4916–4931.

Carpenter, J., Bithell, J., 2000. Bootstrap confidence intervals: when, which, what? Statistics Med. 19, 1141–1164.

Cole, T., Green, P., 1991. Smoothing reference centile curves: the LMS method and penalized likelihood. Stat. Med. 11, 1305–1319.

Costafreda, S., Dinov, I., Tu, Z., Shi, Y., Liu, C., Kloszewska, I., Mecocci, P., Soininen, H., Tsolaki, M., Vellas, B., Wahlund, L., Spenger, C., Toga, A., Lovestone, S., Simmons, A., 2011. Automated hippocampal shape analysis predicts the onset of dementia in mild cognitive impairment. Neuroimage 56, 212–219.

Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Leh'ericy, S., Habert, M., Chupin, M., Benali, H., Colliot, O., 2011. Automatic classification of patients with Alzheimers disease from structural MRI: a comparison of ten methods using the ADNI database. NeuroImage 56, 766781.

Davis, B., Fletcher, P., Bullitt, E., Joshi, S., 2010. Population shape regression from random design data. Int. J. Comput. Vis. 90, 255–266.

Dittrich, E., Raviv, T., kasprian, G., Donner, R., Brugger, P., Prayer, D., Langs, G., 2014. A spatio-temporal latent atlas for semi-supervised learning of fetal brain segmentations and morphological age estimation. Med. Image Anal. 18, 9–21.

Fishbaugh, J., Durrleman, S., Prastawa, M., Gerig, G., 2017. Geodesic shape regression with multiple geometries and sparse parameters. Med. Image Anal. 39, 1–17.

Folgoc, L.L., Delingette, H., Criminisi, A., Ayache, N., 2016. Quantifying registration uncertainty with sparse Bayesian modelling. IEEE Trans. Med. Imaging 36, 607–617.

Gousias, I., Rueckert, D., Heckemann, R., Dyet, L., Edwards, J.B.A., Hammers, A., 2008. Automatic segmentation of brain MRIs of 2-year-olds into 83 regions of interest. Neuroimage 40, 672–684.

Hammers, A., Allom, R., Koepp, M., Free, S., Myers, R., Lemieux, L., Mitchell, T., Brooks, D., Duncan, J., 2003. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. Hum. Brain Mapp. 19, 224–247.

Höskuldsson, A., 1988. PLS regression methods. J. Chemom. 2, 211–228.

Huizinga, W., Poot, D., Guyader, J.M., Klaassen, R., Coolen, B., van Kranenburg, M., van Geuns, R., Uitterdijk, A., Polfliet, M., Vandemeulebroucke, J., Leemans, A., Niessen, W., Klein, S., 2016a. PCA-based groupwise image registration for quantitative MRI. Med. Image Anal. 29, 65–78.

Huizinga, W., Poot, D., Roschchupkin, G., Bron, E., Ikram, M., Vernooij, M., Rueckert, D., Niessen, W., Klein, S., 2016b. Modeling the brain morphology distribution in the general aging population. In: Proc. SPIE 9788, Medical Imaging 2016: Biomedical Applications in Molecular, Structural, and Functional Imaging, 9788 – 9788 – 7.

Ikram, M., van der Lugt, A., Niessen, W., Koudstaal, P., Krestin, G., Hofman, A., Bos, D., Vernooij, M., 2015. The Rotterdam Scan Study: design update 2016 and main findings. Eur. J. Epidemiol. 30, 1299–1315.

Jack, C., Bernstein, M., Fox, N., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P., Whitwell, J., Ward, C., Dale, A., Felmlee, J., Gunter, J., Hill, D., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C., Krueger, G., Ward, H., Metzger, G., Scott, K., Mallozzi, R., Blezek, D., Levy, J., Debbins, J., Fleisher, A., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., Weiner, M., 2008. The Alzheimerś disease neuroimaging initiative (ADNI): MRI methods. J. Magnetic Reson. Imaging 27, 685–691.

de Jong, S., 1992. SIMPLS: an alternative approach to partial least squares regression. Chemom. Intelligent Laboratory Syst. 18, 251–263.

Klein, S., Staring, M., Murphy, K., Viergever, M., Pluim, J., 2010. elastix: a toolbox for intensity based medical image registration. IEEE Trans. Med. Imaging 29, 196–205.

Krishnan, A., Williams, L., McIntosh, A., Abdi, H., 2011. Partial least squares (PLS) methods for neuroimaging: a tutorial and review. NeuroImage 56, 455–475.

Kybic, J., 2010. Bootstrap resampling for image registration uncertainty estimation without ground truth. IEEE Trans. Image Process. 19, 64–73.

Marquand, A., Rezek, I., Buitelaar, J., Beckmann, C., 2016. Understanding heterogeneity in clinical cohort using normative models: beyond case-control studies. Biol. Psychiatry 80, 552–561.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., Goualher, G.L., Boomsma, D., Cannon, T., Kawashima, R., Mazoyer, B., 2001. A probabilistic atlas and reference system for the human brain: international consortium for brain mapping (ICBM). Philosofical Trans. R. Soc. Lond. 356, 1293–1322.

Metz, C., Klein, S., Schaap, M., van Walsum, T., Niessen, W., 2011. Nonrigid registration of dynamic medical imaging data using nD+t B-splines and a groupwise optimization approach. Med. Image Anal. 15, 238–249.

de Onis, M., Onyango, A., Borghi, E., Siyam, A., Pinol, A., et al., 2006. WHO Child Growth Standards: Length/height-for-age, Weight-for-age, Weight-for-length, Weight-for height and Body Mass Index-for-age: Methods and Development. Technical Report. WHO Department of Health and Nutrition.

Rueckert, D., Sonoda, L., Hayes, C., Hill, D., Leach, M., Hawkes, D., 1999. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Trans. Med. Imaging 18, 712–721.

Schrijvers, E., Verhaaren, B., Koudstaal, P., Hofman, A., Ikram, M., Breteler, M., 2012. Is dementia incidence declining?: Trends in dementia incidence since 1990 in the Rotterdam Study. Neurology 78, 1456 – 1163.

Serag, A., Aljebar, P., Ball, G., Counsell, S., Boardman, J., Rutherford, M., Edwards, D., Hajnal, J., Rueckert, D., 2012. Construction of a consistent high-definition spatio-temporal atlas of the developing brain using adaptive kernel regression. NeuroImage 59, 2255–2265.

Singh, N., Fletcher, P., Preston, J., King, R., Marronb, J., Weinerc, M., Joshi, S., Alzheimers Disease Neuroimaging Initiative (ADNI), 2014. Quantifying anatomical shape variations in neurological disorders. Med. Image Anal. 18, 616–633.

Sokooti, H., Saygili, G., Glocker, B., Lelieveldt, B., Staring, M., 2016. Accuracy estimation for medical image registration using regression forests. In: MICCAI 2016: Medical Image Computing and Computer-assisted Intervention. Springer, pp. 107–115.

Tustison, N., Avants, B., Cook, P., Zheng, Y., Egan, A., Yushkevich, P., Gee, J., 2010. N4ITK: improved N3 bias correction. IEEE Trans. Med. Imaging 29, 1310–1320.

Vernooij, M., Smits, M., 2012. Structural neuroimaging in aging and Alzheimer's disease. Neuroimaging Clin. N. Am. 22, 33–55.
Wiklund, S., Nilsson, D., Eroksson, L., Sjöström, M., Wold, S., Faber, K., 2007. A randomization test for PLS component selection. J. Chemom. 21, 427–439.
Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemom. Intelligent Laboratory Syst. 58, 109–130.
Yee, T., 2010. The VGAM package for categorical data analysis. J. Stat. Softw. 32, 1–34.
Yeo, I., Johnson, R., 2000. A new family of power transformations to improve normality or symmetry. Biometrika 87, 954–959.
Ziegler, G., Dahnke, R., Winkler, A., Gaser, C., 2013. Partial least squares correlation of multivariate cognitive abilities and local brain structure in children and adolescents. NeuroImage 82, 284–294.
Ziegler, G., Dhanke, R., Gaser, C., 2012. Models of the aging brain structure and individual decline. Front. Neuroinformatics 6, 1–16.
Ziegler, G., Ridgway, G., Dahnke, R., Gaser, C., for The Alzheimer's Disease Neuroimaging Initiative, 2014. Individualized Gaussian process-based prediction and detection of local and global gray matter abnormalities in elderly subjects. NeuroImage 97, 333–348.