NeuroImage xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

NeuroImage

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

# Estimating the age of healthy subjects from $T_1$-weighted MRI scans using kernel methods: Exploring the influence of various parameters

Katja Franke [a,*], Gabriel Ziegler [a], Stefan Klöppel [b], Christian Gaser [a] and the Alzheimer's Disease Neuroimaging Initiative [1]

[a] Structural Brain Mapping Group, Department of Psychiatry, University of Jena, Jena, Germany
[b] Department of Psychiatry and Psychotherapy, Section of Gerontopsychiatry and Neuropsychology, Freiburg Brain Imaging, University Hospital Freiburg, Freiburg, Germany

ARTICLE INFO

*Article history:*
Received 11 September 2009
Revised 1 December 2009
Accepted 5 January 2010
Available online xxxx

*Keywords:*
MRI
Relevance vector machines (RVM)
Support vector machines (SVM)
Regression
Aging
Brain disease

ABSTRACT

The early identification of brain anatomy deviating from the normal pattern of growth and atrophy, such as in Alzheimer's disease (AD), has the potential to improve clinical outcomes through early intervention. Recently, Davatzikos et al. (2009) supported the hypothesis that pathologic atrophy in AD is an accelerated aging process, implying accelerated brain atrophy. In order to recognize faster brain atrophy, a model of healthy brain aging is needed first. Here, we introduce a framework for automatically and efficiently estimating the age of healthy subjects from their $T_1$-weighted MRI scans using a kernel method for regression. This method was tested on over 650 healthy subjects, aged 19–86 years, and collected from four different scanners. Furthermore, the influence of various parameters on estimation accuracy was analyzed. Our age estimation framework included automatic preprocessing of the $T_1$-weighted images, dimension reduction via principal component analysis, training of a relevance vector machine (RVM; Tipping, 2000) for regression, and finally estimating the age of the subjects from the test samples. The framework proved to be a reliable, scanner-independent, and efficient method for age estimation in healthy subjects, yielding a correlation of $r = 0.92$ between the estimated and the real age in the test samples and a mean absolute error of 5 years. The results indicated favorable performance of the RVM and identified the number of training samples as the critical factor for prediction accuracy. Applying the framework to people with mild AD resulted in a mean *brain age gap estimate* (BrainAGE) score of +10 years.



## Introduction

During the normal aging process, the brain changes due to progressive (e.g., cell growth and myelination) and regressive neuronal processes (e.g., cell death and atrophy). Brain development and healthy aging have been found to follow a specific pattern. Using a semiautomated approach based on a very crude geometrical method for the segmentation of the MRI data, Pfefferbaum et al. (1994) showed that gray matter (GM) volume increases from birth until the age of four and thereafter decreases continuously until subjects reach their 70 s. White matter (WM) volume increases steadily until around the age of 20 when it plateaus. Cerebrospinal fluid (CSF) exhibits a complementary pattern, remaining constant until about 20 years of age and increasing steadily thereafter (Pfefferbaum et al., 1994). A similar, but more recent study conducted a fully automated voxel-based morphometry (VBM) study with 465 normal subjects aged 17–79 years to explore global and regional effects of age (Good et al., 2001). The results of this cross-sectional VBM study also suggested a linear decline in GM to be predominant in normal ageing as well as a linear increase of CSF with age. Furthermore, local areas of accelerated GM decline and microstructural changes in WM were reported, suggesting a heterogeneous and complex pattern of atrophy across the adult life span (Good et al., 2001). Evidence for a region-specific and non-linear pattern of neurodegenerative age-related changes in GM volume was also provided by cross-sectional morphometric analyses (Terribilli et al., 2009) as well as longitudinal data comparison (Resnick et al., 2003). These results support the hypothesis of normal age-related GM decline being inversely related to the phylogenetic origin of each respective region, with younger structures being the last to mature as well as being more vulnerable to neurodegeneration (see also Terribilli et al., 2009; Toga et al., 2006).

Diseases such as Alzheimer's disease (AD) or schizophrenia alter brain structures in diverse and abnormal modes (Ashburner et al., 2003; Meda et al., 2008). Developing a fully automated, reliable, and sufficiently sensitive as well as specific method for the early identification of such pathologic brain developments – even before

* Corresponding author.
*E-mail address:* katja.franke@uni-jena.de (K. Franke).

[1] Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). 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. Complete list of ADNI investigators is available at www.loni.ucla.edu/ADNI/Collaboration/ADNI_Manuscript_Citations.pdf.

1053-8119/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.01.005

the onset of clinical symptoms – has been given great emphasis during the last years (Ashburner, 2009; Davatzikos et al., 2009). Pathologic brain development patterns have been explored and subsequently a variety of classification methods have been employed to separate one or more groups of patients from healthy controls (Davatzikos et al., 2005, 2008a, 2008b; Fan et al., 2008a, 2008b; Klöppel et al., 2008a, 2008b, 2009; Liu et al., 2004; Teipel et al., 2007; Vemuri et al., 2008, 2009a,b). Most of these studies used a processing sequence that started with segmenting and spatially normalizing MRI data, then applied some kind of feature selection or dimensionality reduction (e.g., principal component analysis (PCA)), trained a classifier based on Support Vector Machines (SVM), and finally estimated the classification accuracy with (jackknife) cross-validation. Typically, the sample sizes of these classification studies were rather small, thus entailing the risk of overfitting, which could potentially produce considerable underperformance of the trained classifier when it is applied to a completely new sample. In order to increase sensitivity and reliability of the classification methods, Ashburner (2009) advocated the initiation and usage of multi-scanner data sets tracking a large number of subjects. Integrating data from different scanners in a linear SVM classification study, Klöppel et al. (2008b) reported rates for correctly classified AD patients versus healthy controls of around 90%. This suggests that kernel methods like SVM have the capability to generalize on data obtained from various scanners.

Recently, Davatzikos et al. (2009) showed the longitudinal progression of AD-like patterns in brain atrophy in the normal aging subjects and furthermore an accelerated AD-like atrophy in subjects with mild cognitive impairment (MCI). These results support the hypothesis of AD being a form of accelerated aging, implying accelerated brain atrophy (Driscoll et al., 2009; Fotenos et al., 2008; Sluimer et al., 2009; Spulber et al., 2008; Wang et al., 2009; for a controversial view, see Ohnishi et al., 2001). In case of schizophrenia, a similar hypothesis of the disease being a syndrome of accelerated aging has been presented (Kirkpatrick et al., 2008). If these hypotheses hold true in future research, accelerated and thus pathologic brain atrophy should be recognizable quite early and before the onset of clinical symptoms. In order to recognize faster brain atrophy, a model of healthy and normal brain aging is needed. A straightforward and efficient solution is to model age regression based on normal brain anatomy such that an individual's age can be accurately estimated from its brain scan alone.

Until recently, only a few studies were published that perform age estimation or prediction based on MRI scans. Lao et al. (2004) tested an SVM-based classification method by assigning their elderly subjects into one of four age groups and reached an accuracy rate of 90%. In order to demonstrate the performance of his algorithm for diffeomorphic image registration, Ashburner (2007) estimated the age of subjects based on their brain images utilizing a relevance vector machine (RVM) for regression (Tipping, 2000, 2001). As a measure for prediction accuracy, a root mean squared error (RMSE) of 6.5 years was reported. Another method used quantitative brain water maps to predict age and gender of 44 healthy volunteers aged 23 to 74 years (Neeb et al., 2006). A linear discriminant analysis with jackknife cross-validation for age prediction resulted in a median absolute deviation between real and predicted age of $\pm 6.3$ years.

Although a number of approaches exist that model the pattern of healthy neuronal aging using MRI data, to our knowledge neither the influences of different processing parameters on age estimation were explored, nor was it used for early detection of abnormal aging processes. Large discrepancies between the true and estimated age could indicate pathologic structural changes. Therefore, this work could help to contribute to an early diagnosis and better understanding of neurodegenerative diseases as well as to a more specific and earlier intervention.

In this paper, we present a framework for automatically and efficiently estimating the age of healthy subjects from $T_1$-weighted MRI scans using RVM-based regression. To avoid overfitting as well as to increase sensitivity and reliability, we combine data from the IXI database (http://fantail.doc.ic.ac.uk) and a second sample (Gaser et al., 1999). In total, data from over 650 healthy subjects aged between 19 and 86, collected from four different scanners, were included. To explore the influence of various parameters on the age estimation framework, several analyses on this large database were conducted. We sought to identify the optimal set of processing parameters when the age of data coming from a new scanner had to be estimated. Another goal of this study was a comparison of the performance of well-established SVM with RVM-based regression. SVM require the optimization of a number of parameters (described in more detail in the Methods section). We therefore expect RVM to be more stable and less vulnerable to parameter selection errors than SVM. Due to the "curse of dimensionality", we expect the age estimation to be more accurate if the dimensionality of the preprocessed data is reduced by a dimension reduction method like PCA.

Finally, the age estimation framework will be applied to a clinical sample from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI), which includes $T_1$-weighted images of people with mild AD as well as healthy elderly control subjects. Compared to the group of healthy subjects, we hypothesized that the AD group would have a systematically larger gap between the estimated brain age and the true age due to accelerated brain aging that is presumed to be responsible for the diseased state.

## Methods

### *Subjects/database*

To train and test the age estimation framework with respect to prediction accuracy and reliability, we used brain MR images of healthy subjects from the publicly accessible IXI database (http://fantail.doc.ic.ac.uk) and from our own sample. In February 2009, the IXI database contained $T_1$ images from 550 normal subjects aged 19–86 years, which were collected on three different scanners (Philips 1.5T, General Electric 1.5T, Philips 3T). The subjects were pseudo-randomly split into a training sample, which was used to generate the regression models in relevance vector regression (RVR) and support vector regression (SVR), and a test sample: after sorting the subjects by age, every fourth subject entered the test sample. Since three subjects, for whom no age was given, had to be excluded, the training sample "TRAIN1-3" consisted of 410 subjects, and the first test sample ("TEST1-3") consisted of the remaining 137 subjects from the IXI database, acquired on the three different scanners mentioned above. The second test sample ("TEST4") originally served as a control group in a clinical study (Gaser et al., 1999). TEST4 contained $T_1$ images from 108 healthy subjects aged 20–59 years, which were obtained on a fourth scanner (Philips 1.5T).

The characteristics of the three groups are given in Table 1, and the distribution of age within the training sample and both test samples are shown in Fig. S1.

### *Preprocessing of structural data*

Preprocessing of the images was done using the SPM8 package; SPM 8, 2009) and the VBM8 toolbox (http://dbm.neuro.uni-jena.de). All $T_1$-weighted images were corrected for bias-field inhomogeneities, then spatially normalized and segmented into GM, WM, and CSF within the same generative model (Ashburner and Friston, 2005). The segmentation procedure was further extended by accounting for partial volume effects (Tohka et al., 2004), by applying adaptive maximum a posteriori estimations (Rajapakse et al., 1997), and by applying hidden Markov random field model (Cuadra et al., 2005) as described by Gaser (2009). Only GM images were used for the TRAIN1-3 sample and to test the age estimation model. To make this age

Please cite this article as: Franke, K., et al., Estimating the age of healthy subjects from $T_1$-weighted MRI scans using kernel methods: Exploring the influence of various parameters, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.01.005

**Table 1**
Characteristics of the subjects in the training groups (TRAIN1-3) and both test samples (TEST1-3 and TEST4). TRAIN1-3 and TEST1-3 were collected from the IXI database utilizing three different scanners, whereas the MRI data of the TEST4 sample were collected on a fourth scanner and were not used for training. The characteristics of the two groups used in the application of the age estimation framework (AD and NO) are given in italics.

| | IXI database (scanners 1–3) | | Own sample (scanner 4) | *ADNI database* | |
|---|---|---|---|---|---|
| | TRAIN1-3 | TEST1-3 | TEST4 | *AD (CDR = 1)* | *NO (CDR = 0)* |
| No. subjects | 410 | 137 | 108 | *102* | *232* |
| Males/females | 184/226 | 58/79 | 68/40 | *47/55* | *119/113* |
| Age mean (SD) | 48.16 (16.61) | 47.99 (16.66) | 32.16 (9.99) | *75.85 (8.25)* | *76.01 (5.12)* |
| Age range | 20–86 | 19–83 | 20–59 | *55–88* | *60–90* |

estimation framework fast and efficient, the images were additionally processed with affine registration (AF) and smoothed with an 8-mm full-width-at-half-maximum (FWHM) smoothing kernel (S8). In order to reduce data size the spatial resolution was set to 8 mm (R8), resulting an image size of about 3700 voxels per subject.

Furthermore – for comparison – the images were registered non-linearly (NL), a 4-mm FWHM smoothing kernel (S4) was used, and spatial resolution was set to 3 mm (R3) and 4 mm (R4). As non-linear spatial normalization, the approach implemented in the New Segment toolbox in SPM8 was used.

*Data reduction*

Usually, there are high spatial correlations in voxel-based structural images, which probably lead to redundant voxels. Moreover, not every single voxel is equally relevant for age prediction. Because of that and due to the "curse of dimensionality", data reduction or feature selection might be necessary to obtain meaningful results from the pattern recognition analysis (Ashburner, 2009; Duchesnay et al., 2007; Guyon and Elisseeff, 2003). Commonly, PCA is conducted to reduce the dimensionality of the data.

Using the "Matlab Toolbox for Dimensionality Reduction" (version 0.7b; van der Maaten, 2007, 2008), PCA was applied to the preprocessed images of the training sample. Then the two test samples were reduced using the resulting PCA transformation. Corresponding to the number of subjects in the training sample, the data finally had a size of 410 principal components per subject.

*Support vector regression (SVR)*

The main idea behind SVMs is the transformation of training data from input space into high-dimensional space – the *feature space* – via a mapping function $\Phi$ (Bennett and Campbell, 2003; Schölkopf and Smola, 2002). For the purpose of classification, the hyperplane that best separates the groups is computed within this feature space, resulting in a non-linear decision boundary within the input space. The best separating hyperplane is found by maximizing the margin between the two groups. The data points lying on the margin boundaries are called *support vectors* since only these are used to specify the optimal separating hyperplane. In the case of overlapping class distributions, some training data points are allowed to be misclassified, resulting in some support vectors lying within the margin or on the wrong side of the margin boundary (soft-margin classification; Bishop, 2006).

For the case of real-valued output functions (rather than just binary outputs as used in classification), the SV algorithm was generalized to regression estimation (Bennett and Campbell, 2003; Schölkopf and Smola, 2002). In SVR, a function has to be found that fits as many data points as possible. Analogous to the soft margin in classification, the regression line is surrounded by a tube. Data points lying within that tube do not influence the course of the regression line. Data points lying on the edge or outside that tube are called *support vectors* (Fig. 1a). The expansion of the tube can be determined in a variety of ways, with $\varepsilon$-SVR and $\nu$-SVR being the most common approaches. In $\varepsilon$-SVR, the *a priori* specified constant $\varepsilon$ defines the width of the linear $\varepsilon$-insensitive tube around the regression line. Data points falling within this $\varepsilon$-insensitive tube are not penalized, and are therefore not taken as support vectors. In $\nu$-SVR, the *a priori* specified sparsity parameter $\nu$ defines the upper bound on the fraction of support vectors, i.e., data points lying outside an $\varepsilon$-insensitive tube that is automatically adjusted in width. To control the behavior of $\varepsilon$-SVR and $\nu$-SVR, the type of kernel has to be chosen, along with two more parameters: $C$, which controls for model complexity, and $\varepsilon$ or $\nu$, respectively. A short overview of SVM can be found in Bennett and Campbell (2003). More details can be found in Bishop (2006) or Schölkopf and Smola (2002).

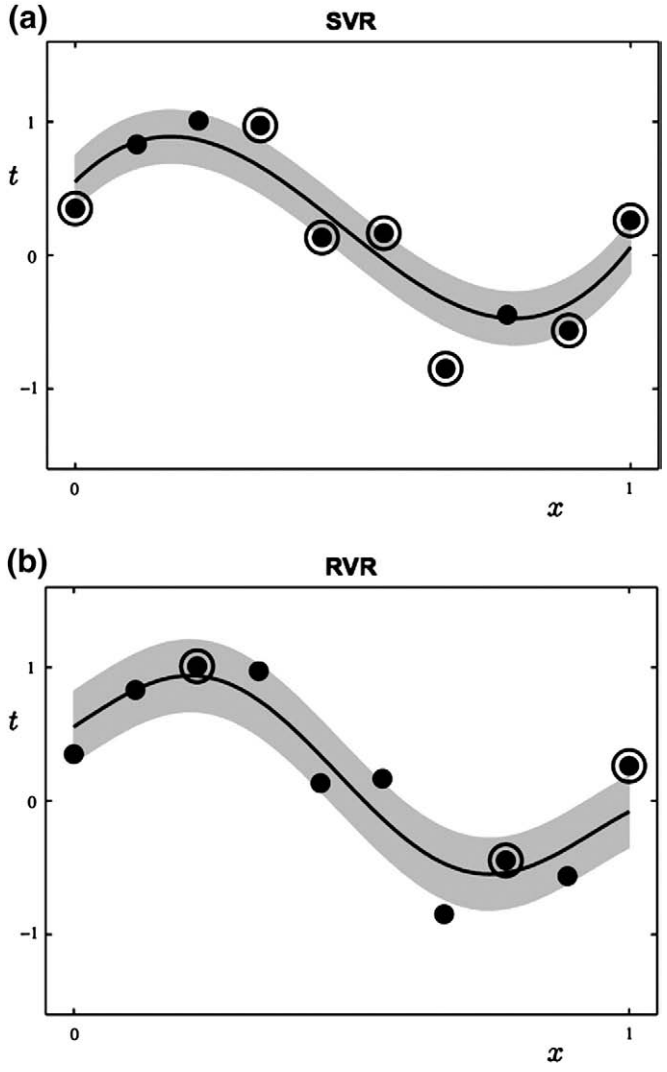


**Fig. 1.** Illustration of (a) SVR and (b) RVR (modified from Bishop, 2006, pp. 344, 349). Data points are shown as black dots; circles indicate (a) support vectors and (b) relevance vectors, respectively.

Please cite this article as: Franke, K., et al., Estimating the age of healthy subjects from $T_1$-weighted MRI scans using kernel methods: Exploring the influence of various parameters, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.01.005

*Relevance vector regression (RVR)*

RVMs were introduced by Tipping (2000) as a Bayesian alternative to SVMs for obtaining sparse solutions to pattern recognition tasks. Moreover, they do not suffer from some limitations of the SVM as their predictions are being probabilistic rather than binary and do not need the determination of additional parameters. In contrast to the support vectors in SVM, the relevance vectors in RVM appear to represent the prototypical examples within the specified classification or regression task instead of solely representing separating attributes.

Furthermore severe overfitting associated with the maximum likelihood estimation of the model parameters was avoided by imposing an explicit zero-mean Gaussian prior (Ghosh and Mujumdar, 2008; Zheng et al., 2008). This prior is a characteristic feature of the RVM, and its use results in a vector of independent hyperparameters that reduces the data set (Faul and Tipping, 2002; Tipping and Faul, 2003; Tipping, 2000). Therefore, in most cases the number of relevance vectors is much smaller than the number of support vectors (Fig. 1b).

To control the behavior of the RVR, only the type of kernel has to be chosen. All other parameters are automatically estimated by the learning procedure itself. More details can be found in Bishop (2006), Schölkopf and Smola (2002), or Tipping (2000, 2001).

*Computing the age estimation model*

We used the freely available toolbox *The Spider* (Version 1.71; Weston et al., 2006) running under MATLAB 7.4.0 to compute the final age regression model.

The $T_1$-weighted MRI data of the training sample TRAIN1-3 and both test samples TEST1-3 and TEST4 were preprocessed by applying affine registration, followed by smoothing with an FWHM kernel of 8 mm and resampling with spatial resolution of 8 mm *(AF_S8_R8)*. The preprocessed data were reduced using PCA, and the RVR age estimation model was trained using this reduced data set. The type of kernel was set to be a polynomial of degree 1, due to its fast convergence rate. We also tested the performance of non-linear kernels. Age estimation did not improve (results not shown), despite adding at least one more parameter (e.g., kernel width). Finally, the ages of the subjects in TEST1-3 and TEST4 were estimated (Fig. 2, box ①).

To measure the accuracy of the age estimations, we used the mean absolute error:

$$\mathrm{MAE} = 1/n^* \sum_i |g_i' - g_i|, \quad (1)$$

with $n$ being the number of subjects in the test sample, $g_i$ the real age, and $g_i'$ the age estimated by the regression model. We found MAE to be the most meaningful measure for assessing the influence of different parameters. For comparison, the root mean squared error:

$$\mathrm{RMSE} = \left[1/n^* \sum_i (g_i' - g_i)^2\right]^{1/2} \quad (2)$$

as well as the correlation coefficient were calculated. Because of the restricted age range in the sample TEST4 and a resulting underestimation of the correlations between the real age and the predicted age, the correlations were corrected following Holmes (1990).

*Systematic analyses of different parameters influencing the age estimation model*

We first compared the age estimation accuracies when testing the age estimation model with data from "known" scanners (i.e., TEST1-3) versus when testing with data from a "new" scanner (i.e., TEST4; see Fig. 2, box ①).

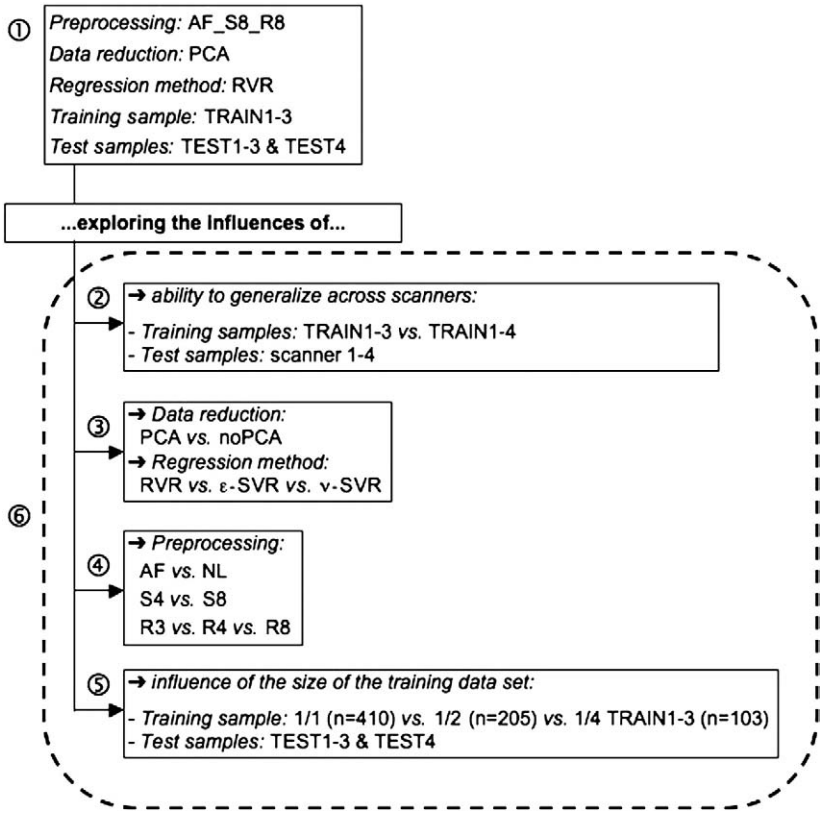


**Fig. 2.** Shown is an overview of the six analyses conducted within this age estimation study to explore the influences of various parameters on age estimation accuracy (AF: affine registration; NL: non-linear registration; S4/S8: smoothing kernel = 4 mm/8 mm; R3/R4/R8: spatial resolution = 3 mm/4 mm/8 mm; PCA: principal component analysis; TRAIN1-3: training sample; TEST1-3 and TEST4: test samples; RVR: relevance vector regression; SVR: support vector regression).

Please cite this article as: Franke, K., et al., Estimating the age of healthy subjects from $T_1$-weighted MRI scans using kernel methods: Exploring the influence of various parameters, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.01.005

Secondly, in order to explore the ability to generalize across scanners, we included data from the fourth scanner into the training sample (see Fig. 2, box ). To test for the effect of scanners on prediction accuracy, the whole IXI data set as well as TEST4 was randomly and separately split into four groups. This resulted in a training set that included 410 randomly selected subjects from scanners 1–3 (IXI) plus 81 randomly selected subjects from scanner 4, and a test set including the remaining 137 subjects from the IXI sample as well as the remaining 27 subjects from scanner 4. The age estimation framework was trained two times: In the first run, the RVR was trained with 410 randomly selected subjects from the IXI sample (scanners 1–3) only. Then the age of the remaining 137 subjects from the IXI sample and of the 27 randomly selected subjects of TEST4 was estimated. In the second run, the RVR was trained with the same 410 IXI subjects as in the first training run plus the randomly selected training sample from TEST4. Again, age was estimated for the actual test subjects from all four scanners. After repeating the whole procedure 20 times, the results were averaged over the trials.

Thirdly, the influence of data reduction and different kernel regression methods was tested (Fig. 2, box ). For comparison, the age estimation model was also computed using $\varepsilon$-SVR and $\nu$-SVR. As before, a polynomial kernel of degree 1 was chosen. Here, the cost parameter $C$ and the width of the $\varepsilon$-tube or $\nu$ for $\varepsilon$-SVR and $\nu$-SVR, respectively, also have to be set. Instead of performing an exhaustive grid search and cross-validation to find these model parameters, we followed Cherkassky and Ma (2004) in choosing the size of the $\varepsilon$-SVR parameters, resulting in $C=98$ and $\varepsilon=0.064$. With respect to $\nu$-SVR, we followed Chalimourda et al. (2004), resulting in $C=20500$ and $\nu=0.54$. Furthermore, we also used the default values of the toolbox with $C=1$, $\varepsilon=0.1$, and $\nu=0.5$, respectively.

Fourthly, to explore which type of preprocessing is best for age prediction, we varied three parameters during preprocessing: (a) affine (AF) vs. non-linear (NL) registration, (b) 4 mm (S4) vs. 8 mm (S8) FWHM smoothing kernel, and (c) 3 mm (R3), 4 mm (R4) vs. 8 mm (R8) for spatial resolution. Memory demands forbade spatial resolutions below 3 mm with this very large subject pool (Fig. 2, box ).

Fifthly, we analyzed the influence of the size of the training data set (i.e., the number of subjects), comparing the full training sample TRAIN1-3 (1/1) against half of the original training sample TRAIN1-3 (1/2) and against a quarter of the original training sample TRAIN1-3 (1/4) (Fig. 2, box ).

Finally, all the parameter variations examined before were integrated into one analysis to assess the proportional amount of influence of each parameter considered (Fig. 2, box ).

### *Application of the age estimation framework to data from the ADNI database*

To test the potential of this age estimation framework to provide clinically relevant predictions, the age of people with early AD and cognitively normal elderly control subjects was estimated. This test sample incorporated MRI data obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies, and non-profit organizations as a $60 million, 5-year public–private partnership. 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 Alzheimer's disease (AD). Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness as well as to lessen the time and cost of clinical trials. The Principle Investigator of this initiative is Michael W. Weiner, M.D., VA Medical Center and University of California-San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 adults, ages 55 to 90 years, to participate in the research—approximately 200 cognitively normal older individuals to be followed for 3 years, 400 people with MCI to be followed for 3 years, and 200 people with early AD to be followed for 2 years. For up-to-date information, see www.adni-info.org.

To compare the age estimations of people with early AD and cognitively normal elderly subjects, two groups were formed and analyzed using the age estimation framework. The AD group included $T_1$-weighted images of subjects who had a global Clinical Dementia Rating Scale (CDR; Morris, 1993) score of 1 at baseline ($n=102$; mean Mini-Mental State Examination (MMSE; Cockrell and Folstein, 1988) score = 22.87). Similarly, the group of healthy controls (NO) included $T_1$-weighted images of subjects who had a global CDR score of 0 at baseline ($n=232$; mean MMSE score = 29.10). Detailed characteristics of both groups can also be found in Table 1.

In order to get a meaningful comparative deviation score, the difference (or gap) between the estimated and the true age was computed. This deviation is termed *brain age gap estimation* (BrainAGE) score. The mean BrainAGE of the NO group should consequently be zero.

**Table 2**
Performance measures of the age estimation model for TEST1-3 and TEST4. Results indicate that the age of the healthy subjects in both test samples could be accurately estimated from MRI scans.

| | TEST1-3 | TEST4 | TEST1-3 + TEST4 |
|---|---|---|---|
| Mean absolute error (MAE) | 4.61 | 5.44 | 4.98 |
| Root mean squared error (RMSE) | 5.90 | 6.73 | 6.28 |
| Correlation ($r$) | 0.94 | 0.89 | 0.92 |
| Confidence interval (at overall mean age of 41 years) | ±10.7 | ±11.7 | ±11.5 |

## Results

### *Performance measures*

The age of healthy subjects in both test samples was accurately estimated from their MRI scans (see Table 2), with an overall

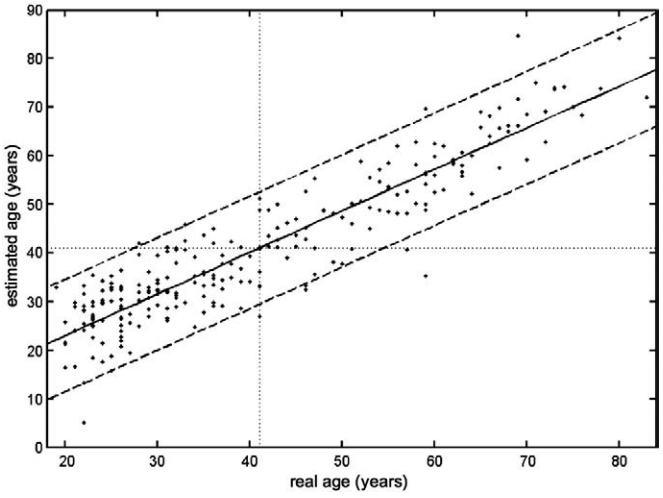


**Fig. 3.** Estimated age and real age are shown for the whole test sample (TEST1-3 + TEST4) with the confidence interval (dashed lines) at a real age of 41 years of ± 11.5 years. The overall correlation between estimated and real age is $r=0.92$, and the overall MAE = 4.98 years.

Please cite this article as: Franke, K., et al., Estimating the age of healthy subjects from $T_1$-weighted MRI scans using kernel methods: Exploring the influence of various parameters, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.01.005

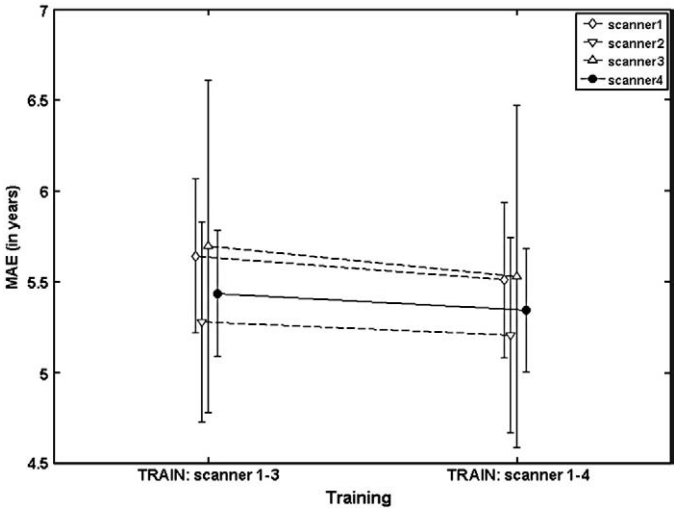


**Fig. 4.** To test for the effect of scanners on prediction accuracy, the IXI data set (scanners 1–3) as well as sample TEST4 (scanner 4) were randomly split into four groups. The first training run included 75% of the IXI data. For the modified training run, 75% of the TEST4 sample was added to the IXI training set. Age estimation was performed on the remaining data. Results were averaged over 20 trials and are shown for each scanner separately. Error bars depict the standard error of the mean (SEM).

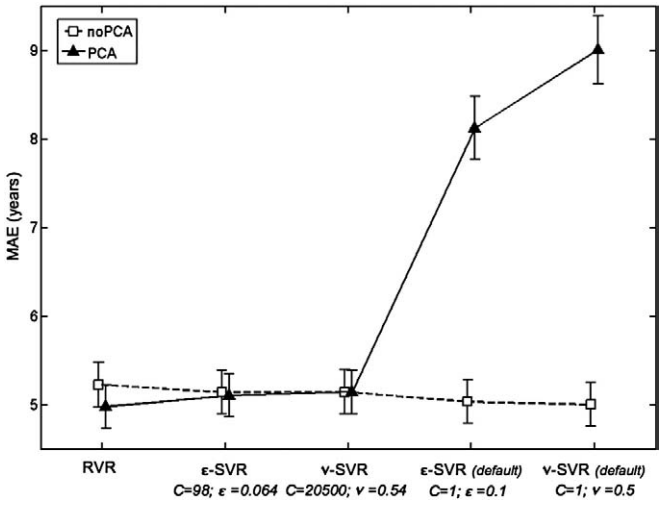


**Fig. 5.** Age estimation tended to be best when the dimensionality of the data was reduced via PCA (solid line) and RVR was used for model calculation. With the reduced data, the performance of $\varepsilon$-SVR and $\nu$-SVR was not stable but depended heavily on the choice of parameters. Error bars depict the SEM.

correlation of $r = 0.92$ and an MAE of just 5 years. The age prediction tended to be slightly more accurate in TEST1-3, which consisted of subjects scanned on the same three scanners as the subjects in the training sample, whereas the subjects in TEST4 had been scanned on a scanner that was not included in the training sample. The 95% confidence interval for the prediction of age was stable along the age range, with no broadening at old age (cf. age $= 20 \pm 11.6$ years, age $= 80 \pm 11.7$ years; see Fig. 3). Furthermore, a correlation of $r = -0.015$ between MAE and the true age indicated no systematical bias in the age estimations as a function of true ages.

The results did not depend on gender in terms of MAE (5.04 years for male, 4.92 years for female subjects) or correlation ($r = 0.92$ for both genders). Again, there was no correlation between estimation accuracy and true age for either gender (male: $r = 0.03$; female: $r = -0.05$).

The most important features in the MRI data that were used by the RVR for estimating the age are shown in the supplementary material (Fig. S2).

*Influence of different scanners*

As shown in the first analysis, estimating the age from MRI scans after training an RVR yields highly accurate predictions, even for completely new data from another scanner. To analyze the influence of scanners on the accuracy of age estimation, the analysis described in the Systematic analyses of different parameters influencing the age estimation model section was conducted, in which 75% of the subjects from either scanners 1–3 only or all scanners were used as the training group. After averaging the results from 20 trials, no difference in estimation accuracy was found between both training runs. When analyzing scanners separately, the accuracy of age prediction varied only slightly between individual scanners (see Fig. 4).

*Impact of regression methods and data reduction*

Because $\varepsilon$-SVR and $\nu$-SVR are kernel methods that are more common than RVR, it is desirable to investigate the differences between the performances of all three methods. Furthermore, dimensionality reduction via PCA may also influence the accuracy of age estimation.

As summarized in Table 3, age estimation tended to be more accurate when the dimensionality of the data was reduced to 410 principal components and RVR was used for model calculation (also see Fig. 5). On the other hand, especially when using principal components, the performance of $\varepsilon$-SVR and $\nu$-SVR was not stable but depended heavily on the choice of its parameters. While using sample-dependent parameters as proposed in Cherkassky and Ma (2004) and Chalimourda et al. (2004), the MAEs reached up to 5 years and thus were comparable to the MAE from the RVR model. Without using sample-dependent parameters or performing a grid search to find optimal parameters for $\varepsilon$-SVR and $\nu$-SVR, but instead using the default values (i.e., in *The Spider*: $C = 1$; $\varepsilon = 0.1$ and $\nu = 0.5$, respectively), the MAE for estimating the age with reduced data was substantially worse—scoring 8 and 9 years, respectively.

Taking a closer look at the number of principal components used in training and testing the age estimation model (using RVR), the accuracy continuously improved with an increasing number of principal components, with a convergence to the smallest MAE at about the first 350 principal components (Fig. 6). Severe overfitting was prevented due to the inherent characteristics of RVM.

**Table 3**
Results of training and testing the age estimation model utilizing different regression methods, each with and without dimension reduction via PCA. MAE (in years) is shown, with the best results in bold.

| Sample | RVR | | $\varepsilon$-SVR $C = 98$; $\varepsilon = 0.064$ | | $\nu$-SVR $C = 20500$; $\nu = 0.54$ | | $\varepsilon$-SVR (default) $C = 1$; $\varepsilon = 0.1$ | | $\nu$-SVR (default) $C = 1$; $\nu = 0.5$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | PCA | noPCA | PCA | noPCA | PCA | noPCA | PCA | noPCA | PCA | noPCA |
| TEST1-3 | **4.61** | 4.96 | 4.85 | 4.85 | 4.85 | 4.85 | 9.82 | 4.76 | 11.06 | 4.72 |
| TEST4 | 5.44 | 5.57 | 5.42 | 5.51 | 5.51 | 5.51 | 5.97 | 5.39 | 6.38 | **5.36** |
| TEST1-3 + TEST4 | **4.98** | 5.23 | 5.10 | 5.14 | 5.14 | 5.14 | 8.12 | 5.04 | 9.00 | 5.00 |

Please cite this article as: Franke, K., et al., Estimating the age of healthy subjects from $T_1$-weighted MRI scans using kernel methods: Exploring the influence of various parameters, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.01.005

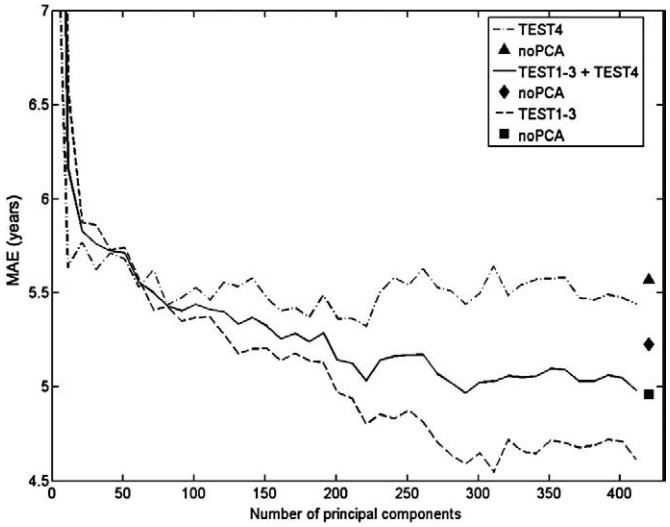


**Fig. 6.** The accuracy of the age estimation model (using RVR) continuously improves with an increasing number of principal components, observing a convergence to the smallest MAE at about the first 350 principal components. MAEs shown for each test sample separately as well as for both test samples together (solid line). Symbols represent MAEs resulting from training the age estimation model *without* data reduction, but utilizing the preprocessed MRI data (diamond for both test samples together).

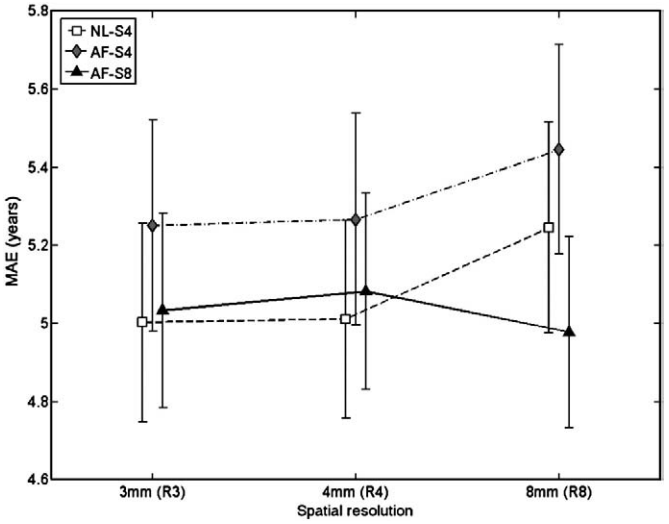


**Fig. 7.** Comparing the different kinds of registration (AF: affine *v*ersus NL: non-linear), different sizes of the smoothing kernel (S4: 4 mm vs. S8: 8 mm), and different spatial resolutions (R3: 3 mm, R4: 4 mm, R8: 8 mm), the MAE of age estimation changes only slightly, with the most accurate age estimation obtained for affine registration and a smoothing kernel of 8 mm (solid line). Error bars depict the SEM.

Furthermore, training and testing the age estimation model utilizing RVR or SVR was computationally fast, with a processing time for training and testing the reduced data of only a few seconds on MAC OS X, Version 10.4.11, Dual 2.5 GHz PowerPC G5 (Fig. S3).

*Comparison of variations in data preprocessing (affine vs. modulated, smoothing, and spatial resolution)*

With respect to preprocessing of the MRI data, we compared different kinds of registration (AF versus NL), different sizes of the smoothing kernel (S4 versus S8), and different spatial resolutions (R3, R4, and R8). The MAE of the age estimations ranged from 4.98 to 5.45 years, and the most accurate predictions occurred with affine registration and a smoothing kernel of 8 mm. The influence of spatial resolution was negligible (Table 4, Fig. 7).

*Influence of the size of training data*

Fig. 8 illustrates that the size of the training data set had a strong effect on the accuracy of age estimation. Whereas the full data set ($n = 410$ subjects) produced an MAE of less than 5 years, using only one half ($n = 205$) or a quarter ($n = 103$) of the training data set for training the age estimation model produced MAEs of 5.2 and 5.6 years, respectively.

**Table 4**
Results of analyses with respect to registration method (AF: affine versus NL: non-linear), size of the smoothing kernel (S4: 4 mm versus S8: 8 mm), and spatial resolution (R3: 3 mm, R4: 4 mm, R8: 8 mm). Results are shown in terms of MAE (in years), and the best results are marked in bold.

| Registration | NL | | | AF | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Smoothing kernel | S4 | | | S4 | | | S8 | | |
| Spatial resolution | R3 | R4 | R8 | R3 | R4 | R8 | R3 | R4 | R8 |
| TEST1-3 | 5.02 | 5.05 | 5.28 | 5.21 | 5.18 | 5.19 | 4.67 | 4.72 | **4.61** |
| TEST4 | 4.98 | **4.96** | 5.19 | 5.30 | 5.38 | 5.77 | 5.49 | 5.54 | 5.44 |
| TEST1-3 + TEST4 | 5.00 | 5.01 | 5.24 | 5.25 | 5.27 | 5.45 | 5.03 | 5.08 | **4.98** |

*Comparing the influence of the various parameters*

Merging the set of all adjustable parameters and methodologies, it can be seen in Fig. 9 that the accuracy of age estimation depended mostly on the number of subjects used for training. The method for preprocessing the $T_1$-weighted MRI images also showed a strong influence on the accuracy of age estimation, again favoring affine registration with a broad smoothing kernel. Furthermore, reducing the dimensionality of data via PCA also had a moderate effect on the MAE.

*Estimating the age of patients with early AD*

The age estimation framework was applied to $T_1$-weighted MRI images of the NO group and the AD group sampled from the ADNI database. The BrainAGE score was calculated for each subject. For the AD group, the mean BrainAGE score was 10 years, implying a

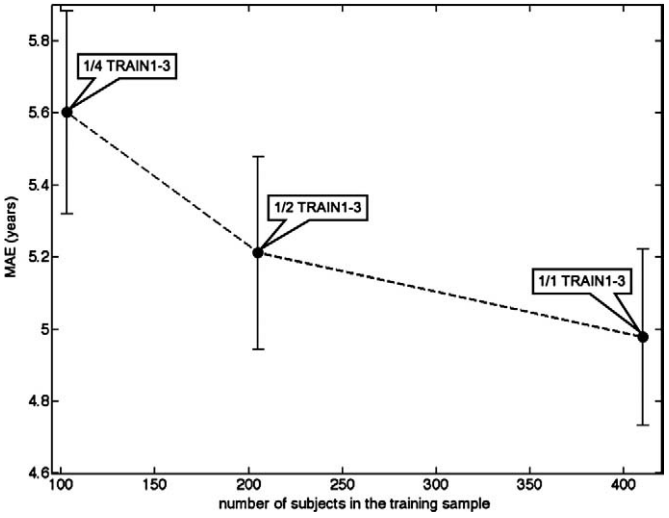


**Fig. 8.** Shown is the influence of the size of trainings data set. Whereas the full data (1/1 TRAIN1-3) set produced an MAE of less than 5 years, taking only one half (1/2 TRAIN1-3) or a quarter (1/4 TRAIN1-3) of the training data set for computing the age estimation model produced MAEs of 5.2 and 5.6 years, respectively. Error bars depict the SEM.

Please cite this article as: Franke, K., et al., Estimating the age of healthy subjects from $T_1$-weighted MRI scans using kernel methods: Exploring the influence of various parameters, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.01.005

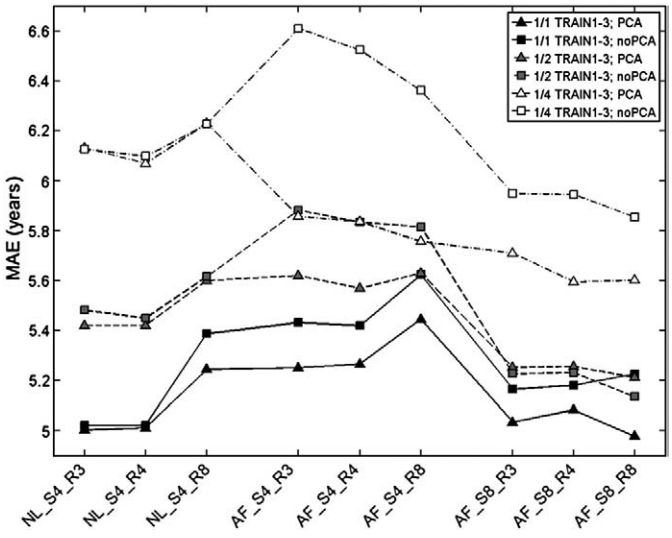


**Fig. 9.** Integrating the influences of the various parameters: the accuracy of age estimation essentially depends on the number of subjects used for training the age estimation model (solid line: full training set TRAIN1-3); the method for preprocessing the $T_1$-weighted MRI images also showed a strong influence on the accuracy of age estimation; and data reduction via PCA only had a moderate effect on the MAE (triangles).

systematically higher estimated than true age based on the MRI data (see Fig. 10). This deviation was highly significant ($p<0.001$; df = 332).

## Discussion

For estimating the age of healthy subjects from $T_1$-weighted MRI scans, we propose a framework that includes automatic preprocessing of the images, dimension reduction via PCA, training of an RVM for regression with a polynomial kernel of degree 1, and finally estimating the age of the subjects from the two test samples TEST1-3 and TEST4. This age estimating framework turns out to be a straightforward method to accurately and reliably estimate age with as little preprocessing and parameter optimization as possible. The additional challenge consisted of combining images from three different scanners for training and testing with an additional testing set from a fourth scanner not included during the training step.

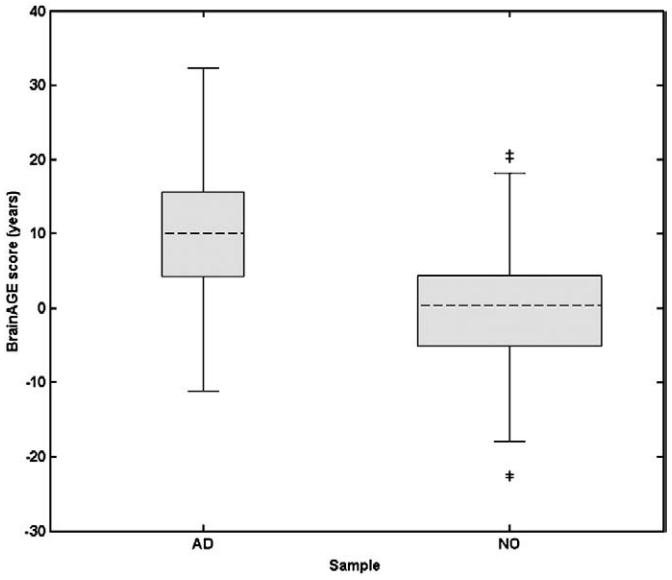


**Fig. 10.** Shown are box plots with BrainAGE scores (in years) for the two samples from the ADNI database (AD with CDR = 1, NO with CDR = 0). The gray boxes contain the values between the 25th and 75th percentiles of the samples, including the median (dashed line). Lines extending above and below each box symbolize data within 1.5 times the interquartile range (outliers are displayed with a +). The width of the boxes depends on the sample size.

Using MRI data from more than 650 healthy subjects aged between 19 and 86 and scanned on different scanners, the age estimation with RVR showed excellent performance for both test samples, with an overall MAE of only 5 years and a correlation of $r = 0.92$ between the estimated and the real age. Although the data in TEST4 were collected on a scanner that was not included in the training step, the performance measures for age estimation showed only minor differences to those of TEST1-3. We did not detect any systematical bias in the age estimation with older age or gender.

Including data from the fourth "unknown" scanner into the training sample did not improve the overall accuracy of age prediction. This could be due to the age range of the samples. TEST4 comprised data from subjects aged between 20 and 59 years, which were already frequently represented in the original training sample TRAIN1-3. On the other hand, adding data from healthy subjects with an age range of 60 to 90 would probably have had a stronger influence on the performance of RVR. Thus, with respect to combining data from different scanners, our results are in line with those of Klöppel et al. (2008b). They indicate that the effect of scanner is sufficiently different from that of the aging process that they could be separated by the regression method. These encouraging results suggest this framework as an accurate, scanner-independent, and efficient method for age estimation in healthy subjects.

In RVR, the type of kernel is the only parameter that has to be defined by the user. In contrast, in $\varepsilon$-SVR and $\nu$-SVR, another two parameters have to be chosen and can decrease the performance if they are not optimized for the specific sample. Age estimation with RVR tends to be slightly better with PCA than without. Furthermore, using the principal components for training and testing with RVR only needed a few seconds and thus is significantly faster than using the full original data set (see Fig. S3).

We decided to use PCA for data reduction because of several reasons: it is a rather simple and commonly used method, and a number of fast implementations exist that are compatible with large data sets. Furthermore, when testing other data reduction or feature selection methods (e.g., Recursive Feature Elimination; Guyon et al., 2002; Guyon and Elisseeff, 2003), we did not observe any improvement in accuracy of age estimation. Also, van der Maaten et al. (2007) reported that the results of their experiments on artificial and natural data sets indicate no clear improvement of non-linear techniques (for example, Isomap or Laplacian Eigenmaps and others) over traditional PCA.

The number of training samples was found to have the strongest influence on the accuracy of age prediction. Our results suggest that the preprocessing of the $T_1$-weighted MRI images can be done fairly rapidly by performing an affine registration only with a large smoothing kernel (e.g., 8 mm). Furthermore, given limited computing time and memory, a coarse spatial resolution (e.g., 8 mm) can be used without losing estimation accuracy. A dimensionality reduction of the data can be conducted using PCA, which tends to improve the accuracy and at the same time speeds up the computing of the RVR model and estimating the age values of the test subjects.

Finally, our age estimation framework has the potential to provide clinically relevant information. With a mean BrainAGE score of +10 years, the subjects with early AD showed signs of accelerated brain aging.

In conclusion, our age estimation framework could potentially help to recognize or indicate faster brain atrophy before the onset of clinical symptoms, thus contributing to an early diagnosis of neurodegenerative diseases and facilitate early treatment or a preventative intervention. Depending on the availability of subject

Please cite this article as: Franke, K., et al., Estimating the age of healthy subjects from $T_1$-weighted MRI scans using kernel methods: Exploring the influence of various parameters, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.01.005

data, future explorations could include applying this framework to other neurodegenerative diseases, evaluating the therapeutic effect of drugs or other treatment modalities, and to predict either the severity of symptoms or the possible rate of cognitive decline.

## Acknowledgments

We are grateful to Dr. Rachel Yotter and Dr. Daniel Mietchen for their comments on the manuscript. We would also like to thank our anonymous reviewers for helpful comments. This work was supported in part by BMBF grants 01EV0709 and 01GW0740.

The clinical data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu\ADNI). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public–private partnership. The Foundation for the National Institutes of Health (www.fnih.org) coordinates the private sector participation of the $60 million ADNI public–private partnership that was begun by the National Institute on Aging (NIA) and supported by the National Institutes of Health. To date, more than $27 million has been provided to the Foundation for NIH by Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Merck and Co., Inc., Novartis AG, Pfizer Inc., F. Hoffmann-La Roche, Schering-Plough, Synarc Inc., and Wyeth as well as non-profit partners, the Alzheimer's Association and the Institute for the Study of Aging.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2010.01.005.

## References

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. NeuroImage 38, 95–113.
Ashburner, J., 2009. Computational anatomy with the SPM software. Magn. Reson. Imaging 27, 1163–1174.
Ashburner, J., Friston, K.J., 2005. Unified segmentation. NeuroImage 26, 839–851.
Ashburner, J., Csernansky, J.G., Davatzikos, C., Fox, N.C., Frisoni, G.B., Thompson, P.M., 2003. Computer-assisted imaging to assess brain structure in healthy and diseased brains. Lancet Neurol. 2, 79–88.
Bennett, K.P., Campbell, C., 2003. Support vector machines: hype or hallelujah? SIGKDD Explorations 2, 1–13.
Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer, New York, NY.
Chalimourda, A., Schölkopf, B., Smola, A.J., 2004. Experimentally optimal nu in support vector regression for different noise models and parameter settings. Neural Netw. 17, 127–141.
Cherkassky, V., Ma, Y., 2004. Practical selection of SVM parameters and noise estimation for SVM regression. Neural Netw. 17, 113–126.
Cockrell, J.R., Folstein, M.F., 1988. Mini-Mental State Examination (MMSE). Psychopharmacol. Bull. 24, 689–692.
Cuadra, M.B., Cammoun, L., Butz, T., Cuisenaire, O., Thiran, J.-P., 2005. Comparison and validation of tissue modelization and statistical classification methods in T1-weighted MR brain images. IEEE Trans. Med. Imaging 24, 1548–1565.
Davatzikos, C., Shen, D., Gur, R.C., Wu, X., Liu, D., Fan, Y., Hughett, P., Turetsky, B.I., Gur, R.E., 2005. Whole-brain morphometric study of schizophrenia revealing a spatially complex set of focal abnormalities. Arch. Gen. Psychiatry 62, 1218–1227.
Davatzikos, C., Fan, Y., Wu, X., Shen, D., Resnick, S.M., 2008a. Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. Neurobiol. Aging 29, 514–523.
Davatzikos, C., Resnick, S.M., Wu, X., Parmpi, P., Clark, C.M., 2008b. Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI. NeuroImage 41, 1220–1227.
Davatzikos, C., Xu, F., An, Y., Fan, Y., Resnick, S.M., 2009. Longitudinal progression of Alzheimer's-like patterns of atrophy in normal older adults: the SPARE-AD index. Brain 132, 2026–2035.
Driscoll, I., Davatzikos, C., An, Y., Wu, X., Shen, D., Kraut, M., Resnick, S.M., 2009. Longitudinal pattern of regional brain volume change differentiates normal aging from MCI. Neurology 72, 1906–1913.
Duchesnay, E., Cachia, A., Roche, A., Rivière, D., Cointepas, Y., Papadopoulos-Orfanos, D., Zilbovicius, M., Martinot, J.-L., Régis, J., Mangin, J.-F., 2007. Classification based on cortical folding patterns. IEEE Trans. Med. Imaging 26, 553–565.
Fan, Y., Batmanghelich, N., Clark, C.M., Davatzikos, C., 2008a. Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. NeuroImage 39, 1731–1743.
Fan, Y., Resnick, S.M., Wu, X., Davatzikos, C., 2008b. Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study. NeuroImage 41, 277–285.
Faul, A., Tipping, M., 2002. Analysis of sparse Bayesian learning. Adv. Neural Inf. Process. Syst. (NIPS) 1, 383–390.
Fotenos, A.F., Mintun, M.A., Snyder, A.Z., Morris, J.C., Buckner, R.L., 2008. Brain volume decline in aging: evidence for a relation between socioeconomic status, preclinical Alzheimer disease, and reserve. Arch. Neurol. 65, 113–120.
Gaser, C., 2009. Partial volume segmentation with adaptive maximum a posteriori (MAP) approach. NeuroImage 47, S121.
Gaser, C., Volz, H.P., Kiebel, S., Riehemann, S., Sauer, H., 1999. Detecting structural changes in whole brain based on nonlinear deformations-application to schizophrenia research. NeuroImage 10, 107–113.
Ghosh, S., Mujumdar, P., 2008. Statistical downscaling of GCM simulations to streamflow using relevance vector machine. Adv. Water Resour. 31, 132–146.
Good, C.D., Johnsrude, I.S., Ashburner, J., Henson, R.N., Friston, K.J., Frackowiak, R.S., 2001. A voxel-based morphometric study of ageing in 465 normal adult human brains. NeuroImage 14, 21–36.
Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. J. Mach. Learn. Res. 3, 1157–1182.
Guyon, I., Weston, J., Barnhill, S., Vapnik, V.N., 2002. Gene selection for cancer classification using support vector machines. Mach. Learn. 46, 389–422.
Holmes, D.J., 1990. The robustness of the usual correction for restriction in range due to explicit selection. Psychometrika 55, 19–32.
Kirkpatrick, B., Messias, E., Harvey, P.D., Fernandez-Egea, E., Bowie, C.R., 2008. Is schizophrenia a syndrome of accelerated aging? Schizophr. Bull. 34, 1024–1032.
Klöppel, S., Stonnington, C.M., Barnes, J., Chen, F., Chu, C., Good, C.D., Mader, I., Mitchell, L.A., Patel, A.C., Roberts, C.C., Fox, N.C., Jack, C.R., Ashburner, J., Frackowiak, R.S.J., 2008a. Accuracy of dementia diagnosis—a direct comparison between radiologists and a computerized method. Brain 131, 2969–2974.
Klöppel, S., Stonnington, C.M., Chu, C., Draganski, B., Scahill, R.I., Rohrer, J.D., Fox, N.C., Jack, C.R., Ashburner, J., Frackowiak, R.S.J., 2008b. Automatic classification of MR scans in Alzheimer's disease. Brain 131, 681–689.
Klöppel, S., Chu, C., Tan, G.C., Draganski, B., Johnson, H., Paulsen, J.S., Kienzle, W., Tabrizi, S.J., Ashburner, J., Frackowiak, R.S.J., 2009. Automatic detection of preclinical neurodegeneration: presymptomatic Huntington disease. Neurology 72, 426–431.
Lao, Z., Shen, D., Xue, Z., Karacali, B., Resnick, S.M., Davatzikos, C., 2004. Morphological classification of brains via high-dimensional shape transformations and machine learning methods. NeuroImage 21, 46–57.
Liu, Y., Teverovskiy, L., Carmichael, O., Kikinis, R., Shenton, M., Carter, C., Stenger, V., Davis, S., Aizenstein, H.J., Becker, J.T., Lopez, O., Meltzer, C., 2004. Discriminative MR image feature analysis for automatic schizophrenia and Alzheimer's disease classification. Learn. Theory: Proceed. 3216, 393–401.
Meda, S.A., Giuliani, N.R., Calhoun, V.D., Jagannathan, K., Schretlen, D.J., Pulver, A., Cascella, N., Keshavan, M., Kates, W., Buchanan, R., Sharma, T., Pearlson, G.D., 2008. A large scale ($N=400$) investigation of gray matter differences in schizophrenia using optimized voxel-based morphometry. Schizophr. Res. 101, 95–105.
Morris, J.C., 1993. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 43, 2412–2414.
Neeb, H., Zilles, K., Shah, N.J., 2006. Fully-automated detection of cerebral water content changes: study of age-and gender-related H2O patterns with quantitative MRI. NeuroImage 29, 910–922.
Ohnishi, T., Matsuda, H., Tabira, T., Asada, T., Uno, M., 2001. Changes in brain morphology in Alzheimer disease and normal aging: is Alzheimer disease an exaggerated aging process? AJNR Am. J. Neuroradiol. 22, 1680–1685.
Pfefferbaum, A., Mathalon, D.H., Sullivan, E.V., Rawles, J.M., Zipursky, R.B., Lim, K.O., 1994. A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. Arch. Neurol. 51, 874–887.
Rajapakse, J.C., Giedd, J.N., Rapoport, J.L., 1997. Statistical approach to segmentation of single-channel cerebral MR images. IEEE Trans. Med. Imaging 16, 176–186.
Resnick, S.M., Pham, D.L., Kraut, M.A., Zonderman, A.B., Davatzikos, C., 2003. Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. J. Neurosci. 23, 3295–3301.
Schölkopf, B., Smola, A., 2002. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, Cambridge, MA.
Sluimer, J.D., Van Der Flier, W.M., Karas, G.B., Van Schijndel, R., Barnes, J., Boyes, R.G., Cover, K.S., Olabarriaga, S.D., Fox, N.C., Scheltens, P., Vrenken, H., Barkhof, F., 2009. Accelerating regional atrophy rates in the progression from normal aging to Alzheimer's disease. Eur. Radiol. doi:10.1007/s00330
SPM 8, 2009. Wellcome Trust Centre for Neuroimaging. Institute of Neurology, UCL, London, UK. http://www.fil.ion.ucl.ac.uk/spm/.
Spulber, G., Niskanen, E., Macdonald, S., Smilovici, O., Chen, K., Reiman, E.M., Jauhiainen, A.M., Hallikainen, M., Tervo, S., Wahlund, L.O., Vanninen, R., Kivipelto, M., Soininen, H., 2008. Whole brain atrophy rate predicts progression from MCI to Alzheimer's disease. Neurobiol. Aging doi:10.1016/j.neurobiolaging.2008.1008.1018.
Teipel, S., Born, C., Ewers, M., Bokde, A., Reiser, M., Möller, H., Hampel, H., 2007. Multivariate deformation-based analysis of brain atrophy to predict Alzheimer's disease in mild cognitive impairment. NeuroImage 38, 13–24.
Terribilli, D., Schaufelberger, M.S., Duran, F.L.S., Zanetti, M.V., Curiati, P.K., Menezes, P.R., Scazufca, M., Amaro, E., Leite, C.C., Busatto, G.F., 2009. Age-related gray matter

volume changes in the brain during non-elderly adulthood. Neurobiol. Aging doi:10.1016/j.neurobiolaging.2009.1002.1008.
Tipping, M.E., 2000. The relevance vector machine. In: Solla, S.A., Leen, T.K., Müller, K.-R. (Eds.), Advances in Neural Information Processing Systems 12. MIT Press, pp. 652–658.
Tipping, M.E., 2001. Sparse Bayesian learning and the relevance vector machine. J. Mach. Learn. Res. 1, 211–244.
Tipping, M., Faul, A., 2003. Fast marginal likelihood maximization for sparse Bayesian models. Proc. Ninth Int. Workshop Artificial Intell. Stat. 3–6.
Toga, A.W., Thompson, P.M., Sowell, E.R., 2006. Mapping brain maturation. Trends Neurosci. 29, 148–159.
Tohka, J., Zijdenbos, A., Evans, A., 2004. Fast and robust parameter estimation for statistical partial volume models in brain MRI. NeuroImage 23, 84–97.
van der Maaten, L.J.P., 2007. An Introduction to Dimensionality Reduction Using Matlab. Technical Report MICC 07-07. Maastricht University, Maastricht, The Netherlands.
van der Maaten, L.J.P., 2008. Matlab Toolbox for Dimensionality Reduction. http://ticc.uvt.nl/~lvdrmaaten/Laurens_van_der_Maaten/Matlab_Toolbox_for_Dimensionality_Reduction.html.
van der Maaten, L.J.P., Postma, E.O., van den Herik, H.J., 2007. Dimensionality reduction: a comparative review. http://ict.ewi.tudelft.nl/~lvandermaaten/Publications_files/JMLR_Paper.pdf.
Vemuri, P., Gunter, J.L., Senjem, M.L., Whitwell, J.L., Kantarci, K., Knopman, D.S., Boeve, B.F., Petersen, R.C., Jack, C.R., 2008. Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies. NeuroImage 39, 1186–1197.
Vemuri, P., Wiste, H.J., Weigand, S.D., Shaw, L.M., Trojanowski, J.Q., Weiner, M.W., Knopman, D.S., Petersen, R.C., Jack, C.R., Initiative, A.s.D.N., 2009a. MRI and CSF biomarkers in normal, MCI, and AD subjects: predicting future clinical change. Neurology 73, 294–301.
Vemuri, P., Wiste, H.J., Weigand, S.D., Shaw, L.M., Trojanowski, J.Q., Weiner, M.W., Knopman, D.S., Petersen, R.C., Jack, C.R., Initiative, A.s.D.N., 2009b. MRI and CSF biomarkers in normal, MCI, and AD subjects: diagnostic discrimination and cognitive correlations. Neurology 73, 287–293.
Wang, P.-N., Liu, H.-C., Lirng, J.-F., Lin, K.-N., Wu, Z.-A., 2009. Accelerated hippocampal atrophy rates in stable and progressive amnestic mild cognitive impairment. Psychiatry Res. 171, 221–231.
Weston, J., Elisseeff, A., BakIr, G., Sinz, F., 2006. The Spider. http://www.kyb.mpg.de/bs/people/spider/main.html.
Zheng, Y., Neo, S., Chua, T., Tian, Q., 2008. Probabilistic optimized ranking for multimedia semantic concept detection via RVM. Proc. 2008 Int. Conference Content-based Image Video Retrieval 161–168.

Please cite this article as: Franke, K., et al., Estimating the age of healthy subjects from $T_1$-weighted MRI scans using kernel methods: Exploring the influence of various parameters, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.01.005