

# Semi-Supervised Multimodal Relevance Vector Regression Improves Cognitive Performance Estimation from Imaging and Biological Biomarkers

Bo Cheng · Daoqiang Zhang · Songcan Chen · Daniel I. Kaufer · Dinggang Shen · the Alzheimer's Disease Neuroimaging Initiative

Published online: 16 March 2013  
© Springer Science+Business Media New York 2013

**Abstract** Accurate estimation of cognitive scores for patients can help track the progress of neurological diseases. In this paper, we present a novel semi-supervised multimodal relevance vector regression (SM-RVR) method for predicting clinical scores of neurological diseases from multimodal imaging and biological biomarker, to help evaluate pathological stage and predict progression of diseases, e.g., Alzheimer's diseases (AD). Unlike most existing methods, we predict clinical scores from multimodal (imaging and biological) biomarkers, including MRI, FDG-PET, and CSF. Considering that the clinical scores of mild cognitive impairment (MCI) subjects are often less stable compared to those of AD and normal control (NC) subjects due to the heterogeneity of MCI, we use only the multimodal data of MCI subjects, but no

corresponding clinical scores, to train a semi-supervised model for enhancing the estimation of clinical scores for AD and NC subjects. We also develop a new strategy for selecting the most informative MCI subjects. We evaluate the performance of our approach on 202 subjects with all three modalities of data (MRI, FDG-PET and CSF) from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The experimental results show that our SM-RVR method achieves a root-mean-square error (RMSE) of 1.91 and a correlation coefficient (CORR) of 0.80 for estimating the MMSE scores, and also a RMSE of 4.45 and a CORR of 0.78 for estimating the ADAS-Cog scores, demonstrating very promising performances in AD studies.

**Keywords** Alzheimer's disease (AD) · Mild cognitive impairment (MCI) · Semi-supervised learning · Relevance vector regression (RVR) · Multimodality

---

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([www.loni.ucla.edu/ADNI](http://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. A complete listing of ADNI investigators can be found at: [www.loni.ucla.edu/ADNI/Collaboration/ADNI\\_Authorship\\_list.pdf](http://www.loni.ucla.edu/ADNI/Collaboration/ADNI_Authorship_list.pdf).

B. Cheng · D. Zhang · S. Chen  
Department of Computer Science and Engineering,  
Nanjing University of Aeronautics and Astronautics,  
Nanjing 210016, China

B. Cheng · D. Zhang (✉) · D. Shen (✉) ·  
the Alzheimer's Disease Neuroimaging Initiative  
Department of Radiology and BRIC, University of North  
Carolina, Chapel Hill, NC 27599, USA  
e-mail: dqzhang@nuaa.edu.cn  
e-mail: dgshen@med.unc.edu

D. I. Kaufer  
Department of Neurology, University of North Carolina,  
Chapel Hill, NC 27599, USA

## Introduction

Many pattern classification methods have been proposed for the diagnosis of Alzheimer's disease (AD) or its prodromal stage, mild cognitive impairment (MCI). However, most existing pattern classification methods aim only for the binary classification, i.e., identifying whether a subject is diseased or healthy from the imaging data. Recently, many researchers have started investigating the application of pattern regression methods to estimating the continuous clinical scores from brain images (Fan et al. 2010; Stonnington et al. 2010; Wang et al. 2010). Compared to pattern classification, pattern regression methods can help assess pathological stage and predict future progression of neurological diseases. It is well known that many diseases result in a continuous spectrum of

structural and functional changes. For example, AD pathology is known to progress gradually over many years, sometimes starting decades before a final clinical stage (Wang et al. 2010). Many studies on the combination of neuropsychological and neuroimaging data show that the preclinical AD is associated with both cognitive and imaging changes (Caselli et al. 2009, 2007; Reiman et al. 1996, 2009; Twamley et al. 2006). Therefore, pattern regression methods can potentially be used to estimate the continuous clinical scores from neuropsychological and neuroimaging data for helping assess the stage of AD disease or predict clinical outcome.

Most existing methods for estimating cognitive test scores use only a single modality of data, e.g., magnetic resonance imaging (MRI) (Fan et al. 2010; Stonnington et al. 2010; Wang et al. 2010). However, biomarkers from different modalities can provide complementary information for AD diagnosis, which is demonstrated in the recent works on multimodal AD diagnosis (Bouwman et al. 2007b; Chetelat et al. 2005; Fan et al. 2008b; Fellgiebel et al. 2007; Geroldi et al. 2006; Hinrichs et al. 2009a, b; Vemuri et al. 2009; Visser et al. 2002; Walhovd et al. 2010a, b; Zhang et al. 2011). For example, MRI can measure spatial patterns of atrophy, and thus can be used to define surrogate markers of the underlying neurodegenerative AD pathology (Fan et al. 2008a). Fluorodeoxyglucose positron emission tomography (FDG-PET) may detect early neocortical dysfunction before atrophy appears (De Santi et al. 2001). Specifically, some recent studies have reported the reduction of glucose metabolism in parietal, posterior cingulate, and temporal brain regions with FDG-PET in AD patients (Diehl et al. 2004; Drzezga et al. 2003). Also, before the appearance of atrophy and the reduction of glucose metabolism, the biological cerebrospinal fluid (CSF) markers (such as concentrations of amyloid  $\beta$  ( $A\beta_{42}$ )) can appear at brain regions of AD patients (Bouwman et al. 2007b; de Leon et al. 2007; Fjell et al. 2010). Therefore, these multimodal data can be potentially used jointly for estimation of clinical cognitive test scores.

On the other hand, supervised relevance vector machine regression (RVR) method (Tipping 2001) has been adopted to estimate the continuous clinical scores from brain images (Fan et al. 2010; Stonnington et al. 2010; Wang et al. 2010). Moreover, Duchesne et al. (2009) demonstrated the relationship between the baseline MRI features and the decline in the Mini-Mental State Exam (MMSE) after 1 year using linear regression modeling for subjects with MCI (Duchesne et al. 2009). In addition, the Alzheimer's Disease Assessment Scale-Cognitive subtest (ADAS-Cog) is the gold standard in AD drug trials for cognitive assessment (Rosen et al. 1984). In all aforementioned studies, the supervised regression models are employed, which often require many labeled data for training in order to achieve good performance. However, the amount of labeled brain image data is usually limited in

clinical practice. Also, due to the heterogeneity of MCI, the clinical scores obtained in cognitive tests, including MMSE and ADAS-Cog (Inoue et al. 2011), are less stable than those for AD and NC subjects. To partially alleviate this issue, instead of using the clinical scores of MCI subjects in this study, we propose to use only their multimodal data to help train a semi-supervised regression model.

To the best of our knowledge, no previous studies have investigated the use of MCI subjects as unlabeled data for constructing a semi-supervised *regression* model to estimate the cognitive performance from multimodal imaging and non-imaging data. We do however, notice that some investigators have used MCI subjects (also as unlabeled data) for constructing semi-supervised *classification* models and then using them for classifying AD patients from normal controls (NC) (Filipovych et al. 2011; Zhang and Shen 2011). However, these semi-supervised classification models can only deal with discrete category (usually binary, i.e., diseased or normal), but cannot estimate continuous clinical scores for disease progression as discussed in this paper. We propose a semi-supervised multimodal relevance vector regression (SM-RVR) method to estimate clinical scores from both imaging and biological biomarkers (i.e., MRI, FDG-PET, and CSF) for new subjects. Moreover, we develop a new strategy to select the most informative MCI subjects for training our semi-supervised regression model (SM-RVR).

In our proposed SM-RVR method, we first employ a  $k$ -nearest neighbor ( $k$ -NN) regression model to estimate the clinical scores for the MCI training subjects. Then, we use a supervised multimodal RVR (M-RVR) to select the most informative MCI subjects by iteratively testing the capability of each MCI training subject in enhancing the estimation of clinical scores for the new testing subjects. Finally, we train an M-RVR model using all AD and NC training subjects and also those selected informative MCI subjects. We have evaluated the performance of our method on 202 subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database and achieved very promising results as compared to the state-of-the-art methods.

## Method

The data used in the preparation of this article were obtained from the ADNI database ([www.loni.ucla.edu/ADNI](http://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 in developing new treatments and monitor their effectiveness, as well as to lessen the time and cost of clinical trials.

ADNI is the result of efforts by many co-investigators from a broad range of academic institutions and private corporations. Subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 adults, aged 55 to 90, 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 (see [www.adni-info.org](http://www.adni-info.org) for up-to-date information). The research protocol was approved by each local institutional review board and written informed consent was obtained from each participant.

## Subjects

The ADNI general eligibility criteria are described at [www.adni-info.org](http://www.adni-info.org). Briefly, subjects are between 55 and 90 years of age and have a study partner able to provide an independent evaluation of functioning. Specific psychoactive medications will be excluded. General inclusion/exclusion criteria are as follows: 1) healthy subjects: Mini-Mental State Examination (MMSE) scores between 24 and 30, a Clinical Dementia Rating (CDR) of 0, non-depressed, non MCI, and non-demented; 2) MCI subjects: MMSE scores between 24 and 30, a memory complaint, having objective memory loss measured by education adjusted scores on Wechsler Memory Scale Logical Memory II, a CDR of 0.5, absence of significant levels of impairment in other cognitive domains, essentially preserved activities of daily living, and an absence of dementia; and 3) Mild AD: MMSE scores between 20 and 26, CDR of 0.5 or 1.0, and meets the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) criteria for probable AD.

In this paper, all ADNI subjects with all corresponding baseline MRI, FDG-PET and CSF data were selected, leading to a total of 202 subjects, including 51 AD patients, 99 MCI patients (including 43 MCI converters (MCI-C) and 56 MCI non-converters (MCI-NC)), and 52 normal controls. Table 1 lists the demographics of all these subjects.

## MRI, PET and CSF

All MRI data were acquired from 1.5 T scanners. Data were collected by a variety of scanners with protocols individualized for each scanner. Raw Digital Imaging and Communications in

**Table 1** Clinical and demographic information for 202 subjects (mean  $\pm$  std)

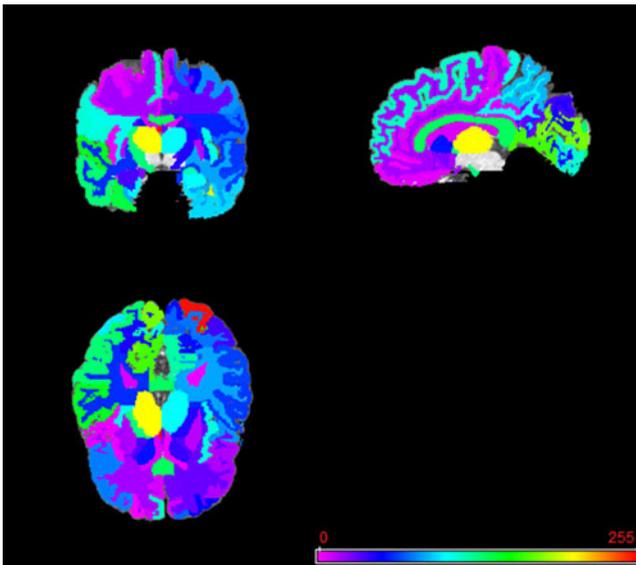
	AD	MCI	NC
No. of subjects	51	99	52
Age	75.2 $\pm$ 7.4	75.3 $\pm$ 7.0	75.3 $\pm$ 5.2
Education	14.7 $\pm$ 3.6	15.9 $\pm$ 2.9	15.8 $\pm$ 3.2
MMSE	23.8 $\pm$ 1.9	27.1 $\pm$ 1.7	29.0 $\pm$ 1.2
ADAS-Cog	18.3 $\pm$ 6.0	11.4 $\pm$ 4.4	7.4 $\pm$ 3.2

AD Alzheimer's disease, MCI mild cognitive impairment, NC normal control, MMSE mini-mental state examination, ADAS-Cog Alzheimer's disease assessment scale-cognitive subscale

Medicine (DICOM) MRI scans were reviewed for quality. Spatial distortions caused by gradient nonlinearity and B1 field inhomogeneity were automatically corrected. PET images were acquired 30–60 min post-injection, averaged, spatially aligned, interpolated to a standard voxel size, intensity normalized, and smoothed to a common resolution of 8-mm full width at half maximum. CSF was collected in the morning after an overnight fast using a 20- or 24-gauge spinal needle, frozen within 1 h of collection, and transported on dry ice to the ADNI Biomarker Core laboratory at the University of Pennsylvania Medical Center. In this paper, CSF A $\beta$ 42, CSF t-tau, and CSF p-tau are used as features.

## Image Analysis

To obtain effective image features for score estimation, we use the specific application tool for image pre-processing similar to (Zhang et al. 2011). First, for all images, we corrected the intensity inhomogeneity using the N3 algorithm (Sled et al. 1998). Next, for all structural MR images, we used both brain surface extractor (BSE) (Shattuck et al. 2001) and brain extraction tool (BET) (Smith 2002) to perform skull-stripping. Then, we further removed cerebellum material. After skull-stripping and cerebellum removal, for each structural MR image, we used the FSL package (Zhang et al. 2001) to segment it into three different tissues: gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). After tissue segmentation, we use atlas warping to partition each subject into 93 ROIs via a template shown in Fig. 1. Then, for each of 93 ROIs, we computed the volume of GM tissue in that ROI as a feature. For the PET image, we used a rigid transformation to align it onto its respective MR image of the same subject, and then computed the average intensity of each ROI in the PET image as a feature. By performing this series of image pre-processing steps, for each subject we acquire 93 features from the MRI image, and another 93 features from the PET image. In addition, we used 3 features from the CSF biomarkers, as mentioned above.



**Fig. 1** A template with 93 manually-labeled ROI regions used for automated labeling of MRI and PET images

### Semi-Supervised Multimodal RVR (SM-RVR)

In this section, we first extended the standard relevance vector regression (RVR) method to the multimodal RVR (M-RVR), and then employed the  $k$ -NN regression and M-RVR for selecting the most informative MCI subjects. Finally, we trained an M-RVR model using all AD and NC training subjects, and the selected more-informative MCI subjects to estimate the clinical scores of new subjects.

### Multimodal Relevance Vector Regression (M-RVR)

**Relevance Vector Regression (RVR)** We will first briefly review the standard RVR algorithm. The main idea of RVR is summarized as follows. Specifically, RVR is a sparse kernel method formulated in a Bayesian framework (Tipping 2001). Given a training set with its corresponding target values, such as  $\{\mathbf{x}_i, t_i\}_{i=1}^N$ , RVR aims to find out the relationship between the input feature vector  $\mathbf{x}$  and its corresponding target value  $t$ :

$$t = f(\mathbf{x}, \mathbf{w}) + \varepsilon_i \quad (1)$$

Where  $\varepsilon_i$  is the measurement noise (assumed independent and following a zero-mean Gaussian distribution), and  $f(\mathbf{x}, \mathbf{w})$  is a linear combination of basis kernel functions  $k(\mathbf{x}, \mathbf{x}_i)$  with the following form:

$$f(\mathbf{x}, \mathbf{w}) = \sum_{i=1}^N w_i k(\mathbf{x}, \mathbf{x}_i) + w_0 \quad (2)$$

Where  $\mathbf{w} = (w_0, w_1, \dots, w_N)^T$  is a weight vector. According to (Tipping 2001), we can obtain a sparse kernel regression model based on the weight vector  $\mathbf{w}$ .

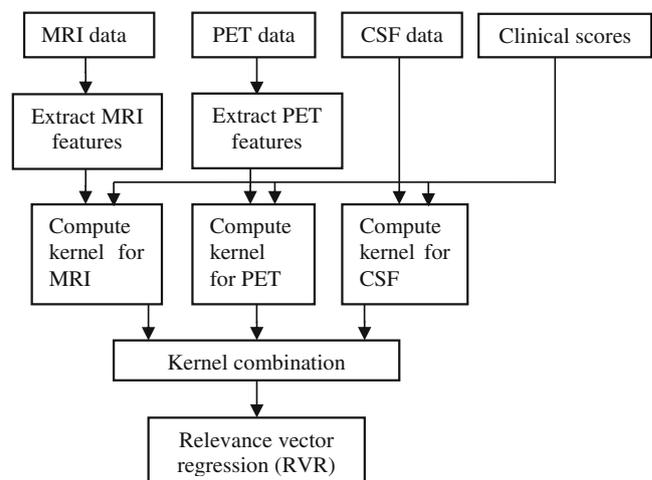
**Multimodal RVR (M-RVR)** Now we can extend RVR to multimodal RVR (M-RVR) for multimodal regression by defining a new integrated kernel function for comparison of two multimodal data  $\mathbf{x}$  and  $\mathbf{x}_i$  as below:

$$k(\mathbf{x}, \mathbf{x}_i) = \sum_m c_m k^{(m)}(\mathbf{x}^{(m)}, \mathbf{x}_i^{(m)}) \quad (3)$$

Where  $\mathbf{x}^{(m)}$  and  $\mathbf{x}_i^{(m)}$  denote the  $m$ -th modality of multimodal data  $\mathbf{x}$  and  $\mathbf{x}_i$ , respectively,  $k^{(m)}$  denotes the kernel function over the  $m$ -th modality, and  $c_m$  denotes the weight on the  $m$ -th modality. This new integrated multiple-kernel can be embedded into the conventional single-kernel RVR, and thus solved by the programs developed for the conventional single-kernel RVR (Tipping 2001). To find the optimal values for the weights  $c_m$ , in this paper, we constrained them so that  $\sum_m c_m = 1$  and then adopt a *coarse-grid search* through cross-validation on the training samples, which has been shown effective in our previous work for multi-kernel support vector machine (SVM) (Zhang and Shen 2011; Zhang et al. 2012, 2011). Compared with most existing multi-kernel learning methods (Rakotomamonjy et al. 2008; Wang et al. 2008; Xu et al. 2010), the main advantage of our method is that it can be conveniently embedded into the conventional RVR and thus solved using the conventional RVR solvers, e.g., Sparse Bayesian toolbox (Tipping 2001). In this paper, we used M-RVR to fuse data from three different modalities, i.e., MRI, PET and CSF. Figure 2 shows the flowchart of M-RVR for multimodal regression using MRI, PET and CSF data.

### Semi-Supervised M-RVR (SM-RVR)

As discussed before, due to the heterogeneity of MCI, the clinical scores of MCI subjects are usually less stable than those of the AD and NC subjects. Therefore, the direct use

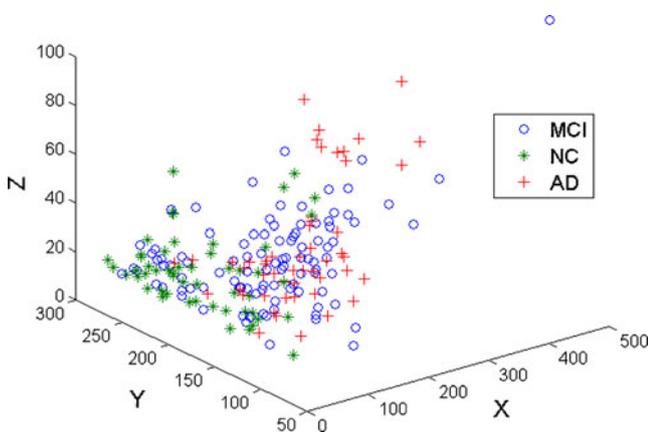


**Fig. 2** The flowchart of M-RVR

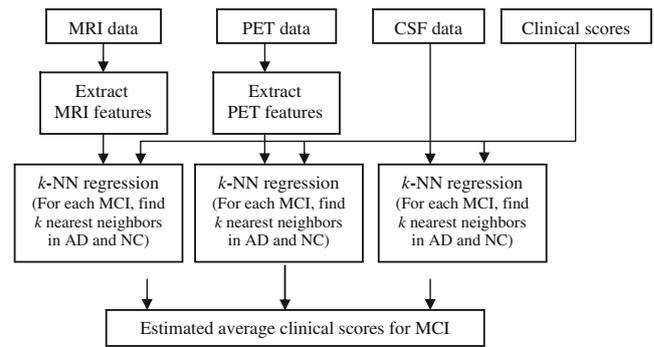
of MCI subjects (i.e., including their corresponding clinical scores as targets) together with the AD and NC subjects for training the regression model could lead to low performance. To address this issue, we propose to first employ the  $k$ -NN regression method for estimating clinical scores of MCI subjects by using clinical scores of the AD and NC subjects, which are more stable than those of MCI subjects. Then, we suggest using M-RVR to select the most informative MCI subjects for subsequent data regression by iteratively testing the capability of each MCI training subject in enhancing the estimation of clinical scores for the new testing subjects. These two steps are described below one by one.

*Estimating Clinical Scores of MCI Subjects* Figure 3 plots the distributions of AD, NC, and MCI subjects with CSF features. As we can see from Fig. 3, there exists a large overlap between MCI subjects and AD/NC subjects. This overlap implies that in the CSF feature space the MCI subjects are similar to some AD or NC subjects. Similar phenomena can also be observed in MRI and PET feature spaces, which to some extent explains why AD and NC subjects can be used for estimating the clinical scores of MCI subjects. The flowchart of estimating clinical scores of MCI subjects is shown in Fig. 4. Specifically, for each modality of MRI, PET, and CSF, we used a respective  $k$ -NN (i.e., find the  $k$ -nearest neighbors among AD and NC training subjects) to estimate clinical scores of each MCI subject, and then computed the average of estimated scores from all three modalities as the final clinical score for that MCI subject.

*M-RVR Based Recursive Sample Selection, to Select the Most Informative MCI Subjects* After getting the estimated clinical score for each MCI subject, one can simply merge those ‘re-labeled’ MCI subjects with AD and NC training



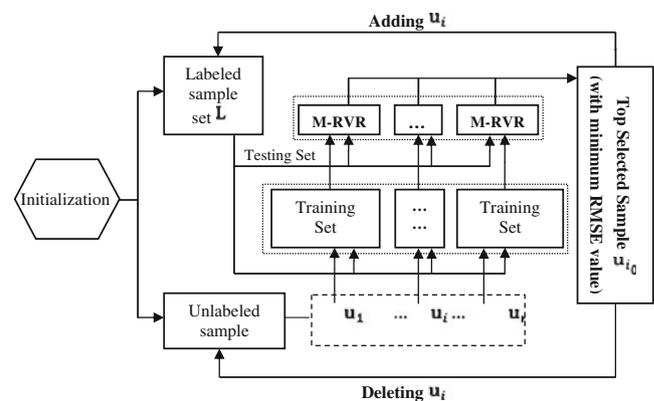
**Fig. 3** Distributions of AD, NC and MCI subjects with CSF features. X is CSF  $A\beta_{42}$ , Y is CSF  $t$ -tau, and Z is CSF  $p$ -tau



**Fig. 4** The flowchart of estimating clinical scores for MCI subjects

subjects to train a regression model. However, to further improve the performance, we propose to first select the most informative MCI subjects and then use them together with the AD and NC training subjects to train the regression model. The flowchart of selecting the most informative MCI subjects is shown in Fig. 5. Specifically, we denote  $L$  as the labeled set which initially contains only AD and NC training subjects, and denote  $U = [u_1, \dots, u_i, \dots, u_r]$  as the unlabeled set which initially contains all MCI subjects with their estimated clinical scores (from AD and NC training subjects using  $k$ -NN) as targeted values. Then, we individually combine each sample in  $U$  with all samples in  $L$  to train an M-RVR regression model (using  $L$  as corresponding testing set) and record the corresponding root-mean-square error (RMSE). Finally, the sample  $u_{i_0}$  with the minimum RMSE value (i.e., most informative) in  $U$  is selected and added into  $L$ , and further deleted from  $U$ . The above procedure was repeated for  $T$  iterations to select the most informative subset of MCI subjects.

Finally, we trained the M-RVR model using the AD and NC training subjects and also those above-selected most informative MCI subjects (with their estimated clinical scores as targets) to estimate the clinical scores for the new test subjects.



**Fig. 5** The flowchart of M-RVR based recursive sample selection

## Validation

To evaluate the performance of our proposed method, we adopted the popular root-mean-square error (RMSE) and the correlation coefficient (CORR) as performance measures. In our experiments, we used a 10-fold cross-validation strategy to divide the labeled sample set (i.e., AD and NC subjects) into a training set and a testing set. Specifically, the whole labeled sample set was equally partitioned into 10 subsets. Each time, the subject samples within one subset were successively selected as the testing samples and all remaining subject samples in the other 9 subsets were used for training. This process was repeated for 10 times independently to avoid any bias introduced by randomly partitioning the dataset in the cross-validation. The RVM regression learner was implemented using Sparse Bayesian toolbox,<sup>1</sup> with the Gaussian kernel. The iteration number  $T$  ( $1 \leq T \leq 99$ ), the Gaussian kernel width  $\sigma \in \{2, 4, 8, 16, 32, 64, 128, 256, 512\}$ , and the number of the nearest neighbors  $k$  ( $1 \leq k \leq 20$ ) were learned based on the training samples by another cross validation. Similarly, the weights in the M-RVR were also learned based on the training samples, through a grid search using the range from 0 to 1 at a step size of 0.1. Also, for each modality feature in the labeled samples and the unlabeled samples, the same feature normalization scheme as used in (Zhang et al. 2011) was adopted here.

## Results

### Comparison on Different Combinations of MRI, PET and CSF Modalities

We first evaluated the performance of SM-RVR on different combinations of MRI, PET and CSF modalities (with totally 7 different combinations). Table 2 shows the performance measures of our SM-RVR method for different combinations. Figures 6 and 7 give the corresponding scatter plots of the actual clinical scores vs. the estimated clinical scores for the estimation of MMSE and ADAS-Cog scores, respectively. As we can see from Table 2 and Figs. 6 and 7, the combination of MRI, PET, and CSF can consistently achieve better results than both single-modality and two-modalities cases. Specifically, SM-RVR using all three modalities can achieve a RMSE of 1.92 and a CORR of 0.80 for MMSE scores, and a RMSE of 4.45 and a CORR of 0.78 for ADAS-Cog scores as shown in Table 2. Table 2 also indicates that the use of two modalities can also improve the regression performance compared with the single-modality cases, although they are inferior to the use of all three modalities. These results validate the advantage of multimodal regression over the conventional single-modal regression in estimation of clinical scores.

<sup>1</sup> <http://www.miketipping.com/index.php?page=rvm>

**Table 2** Regression performance of SM-RVR with respect to different combination of MRI, PET and CSF modalities

Modality	MMSE		ADAS-Cog	
	RMSE	CORR	RMSE	CORR
MRI	2.17	0.73	5.16	0.70
CSF	2.45	0.60	5.62	0.64
PET	2.46	0.62	5.04	0.71
MRI + CSF	2.03	0.77	4.98	0.74
MRI + PET	2.09	0.76	4.73	0.76
PET + CSF	2.38	0.66	4.89	0.73
MRI + CSF + PET	1.92	0.80	4.45	0.78

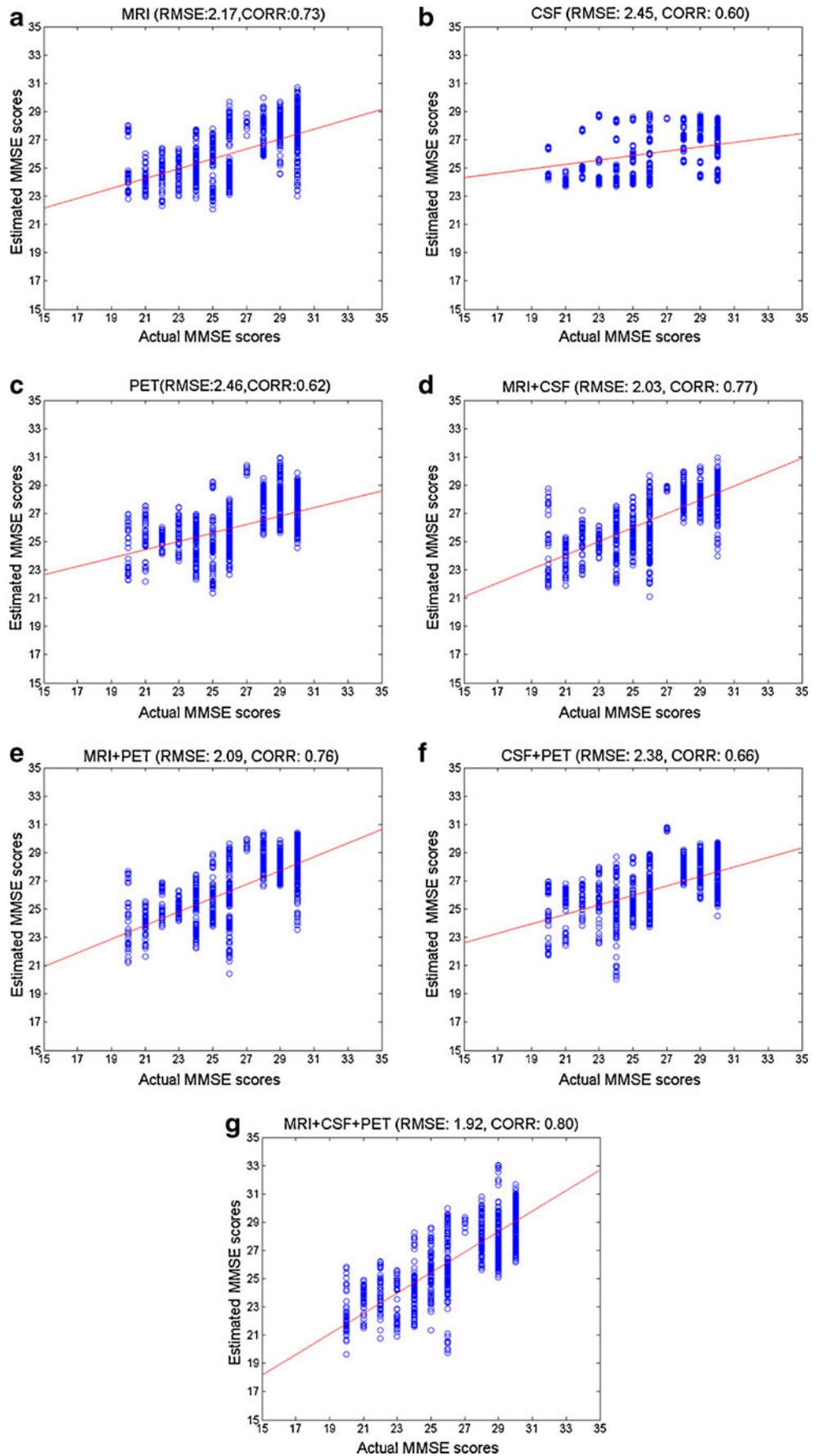
MMSE mini-mental state examination, ADAS-Cog Alzheimer's disease assessment scale-cognitive subscale, RMSE root mean square error, CORR correlation coefficient

Furthermore, to investigate the effect of different combining weights of MRI, PET and CSF ( $C_{MRI}$ ,  $C_{PET}$  and  $C_{CSF}$ ) in the estimation of MMSE and ADAS-Cog scores by our SM-RVR method, we tested all of their possible values, ranging from 0 to 1 at a step size of 0.1, under the constraint of  $C_{MRI} + C_{PET} + C_{CSF} = 1$ . Figures 8 and 9 give MMSE and ADAS-Cog estimation results with respect to different combining weights of MRI, PET and CSF for our SM-RVR method, respectively. Figures 8(a) and 9(a) show the RMSE with respect to different combining weights of MRI, PET and CSF for MMSE and ADAS-Cog estimations, respectively, while Figs. 8(b) and 9(b) show the correlation coefficient (CORR) with respect to different combining weights of MRI, PET and CSF for MMSE and ADAS-Cog estimations, respectively. Note that in each plot of Figs. 8 and 9, only the squares in the upper triangular part have valid values because of the constraint  $C_{CSF} + C_{MRI} + C_{PET} = 1$ . Also, in each plot, the three vertices of the upper triangle, i.e., the top left, top right, and bottom left squares, denote the individual-modality based regression results, respectively. Moreover, for each plot, the three edges of the upper triangle (excluding the three vertices of the upper triangle) denote two-modalities based regression results using MRI+CSF ( $C_{PET}=0$ ), MRI+PET ( $C_{CSF}=0$ ), and CSF+PET ( $C_{MRI}=0$ ). From Figs. 8(a) and 9(a), most of RMSE values within squares of the upper triangle are smaller than those on the three vertices and edges. Also, Figs. 8(b) and 9(b) show that most of CORR values within squares of the upper triangle are larger than those on the three vertices and edges. These results further validate that combining three modalities can achieve better regression results than combining only two modalities or using only one modality.

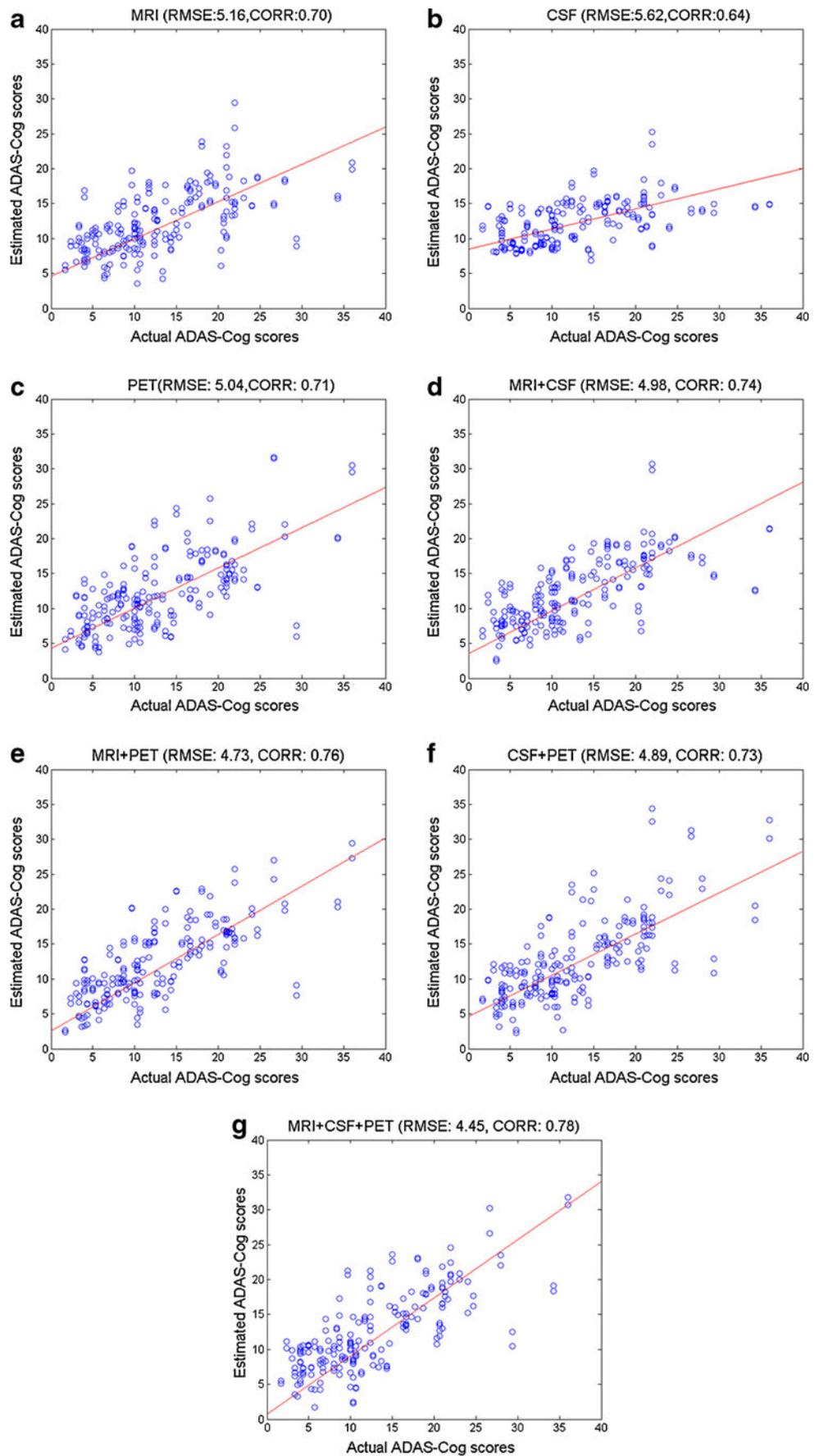
### Comparison with Supervised M-RVR and Other Variants

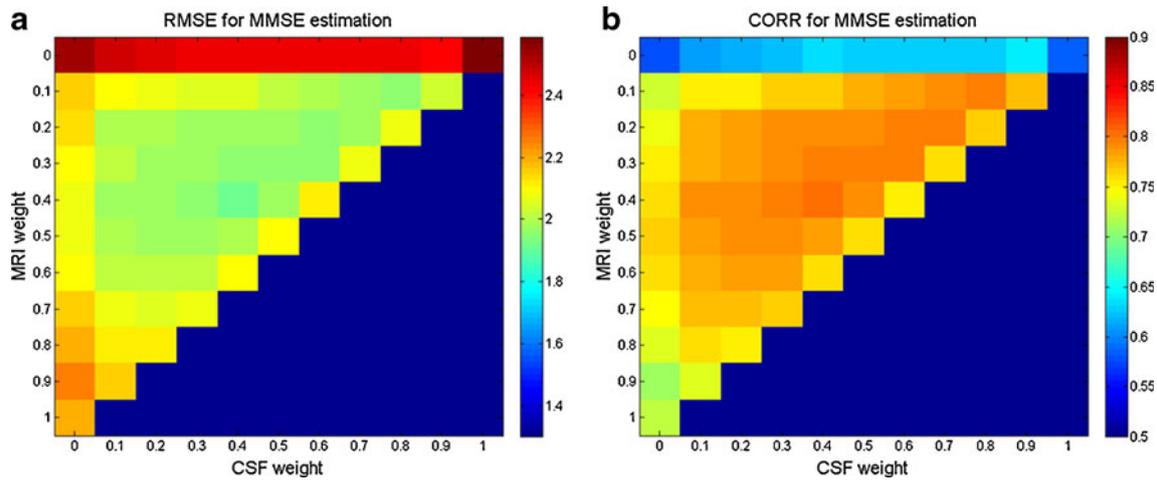
To investigate the efficacy of using the unlabeled MCI samples for helping regression, we first compare SM-RVR with

**Fig. 6** Scatter plots of actual MMSE scores vs. the estimated MMSE scores for seven different combinations of modalities. **a** MRI, **b** CSF, **c** PET, **d** MRI + CSF, **e** MRI + PET, **f** CSF + PET, and **g** MRI + CSF + PET



**Fig. 7** Scatter plots of actual ADAS-Cog scores vs. the estimated ADAS-Cog scores for seven different combinations of modalities. **a** MRI, **b** CSF, **c** PET, **d** MRI + CSF, **e** MRI + PET, **f** CSF + PET, and **g** MRI + CSF + PET



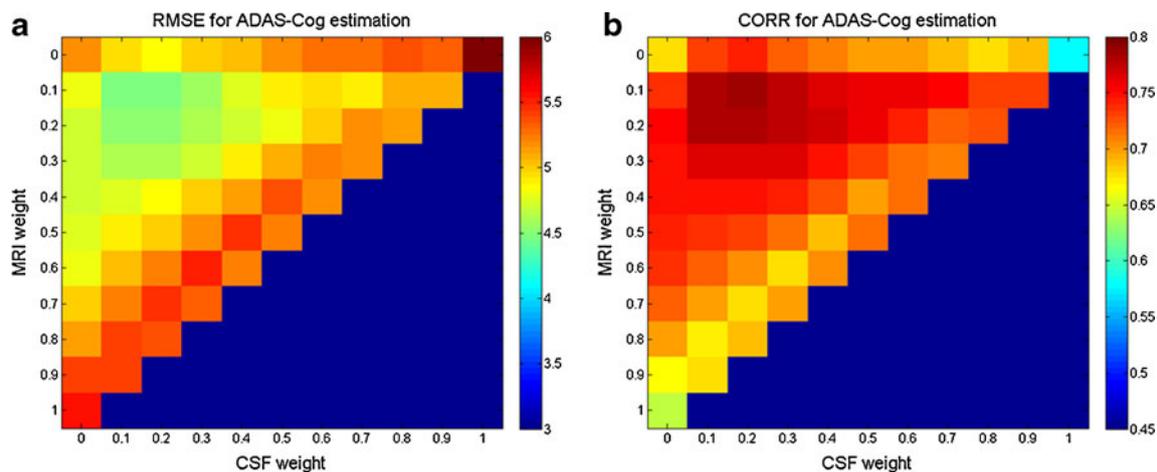


**Fig. 8** MMSE estimation results with respect to different combining weights of MRI, PET and CSF for SM-RVR. CSF weight denotes as  $C_{CSF}$  and MRI weight denotes as  $C_{MRI}$ . Note that PET weight  $C_{PET}$

is not shown, since it can be determined as  $C_{PET} = 1 - C_{CSF} - C_{MRI}$ . **a** RMSE for MMSE Estimation, **b** CORR for MMSE Estimation

the supervised M-RVR (M-RVR for short). In our experiment, we implemented two versions of M-RVR, i.e., 1) M-RVR (AD+NC) which trains a supervised model using only the AD and NC subjects, and 2) M-RVR (AD+NC+MCI) which trains a supervised model using not only the AD and NC subjects, but also the extra MCI subjects (with their corresponding clinical scores). Here, for fair comparison, all methods (including both versions of M-RVR and SM-RVR) are validated on the same testing data involving only the AD and NC subjects, since the corresponding clinical scores of the MCI subjects are less stable due to the heterogeneity of MCI. Notice that although both M-RVR and SM-RVR involve MCI subjects in training models, the former uses both data and the corresponding clinical scores of MCI subjects (together with AD and NC subjects) for training a supervised model, while the latter uses only the data of MCI subjects as unlabeled data

(together with the labeled AD and NC subjects) for training a semi-supervised model. The corresponding results are shown in Table 3. As shown in Table 3, SM-RVR consistently outperforms the two different versions of M-RVR on each performance measure. These findings validate the efficacy of our SM-RVR method in using MCI subjects as unlabeled samples in a semi-supervised framework for improving regression predictive utility. Furthermore, in Fig. 10, we plotted the curves of regression performance measures (i.e., RMSE and CORR) with respect to the use of different number of unlabeled MCI samples for SM-RVR. It shows that the regression performance of SM-RVR steadily improves at first as the number of unlabeled MCI samples increases, (and is significantly better than that of M-RVR in most cases), but it decreases after reaching a certain value. This indicates the importance of using only the selected informative MCI



**Fig. 9** ADAS-Cog estimation results with respect to different combining weights of MRI, PET and CSF for SM-RVR. CSF weight denotes as  $C_{CSF}$  and MRI weight denotes as  $C_{MRI}$ . Note

that PET weight  $C_{PET}$  is not shown, since it can be determined as  $C_{PET} = 1 - C_{CSF} - C_{MRI}$ . **a** RMSE for ADAS-Cog Estimation, **b** CORR for ADAS-Cog Estimation

**Table 3** Comparison of regression performance of SM-RVR and M-RVR

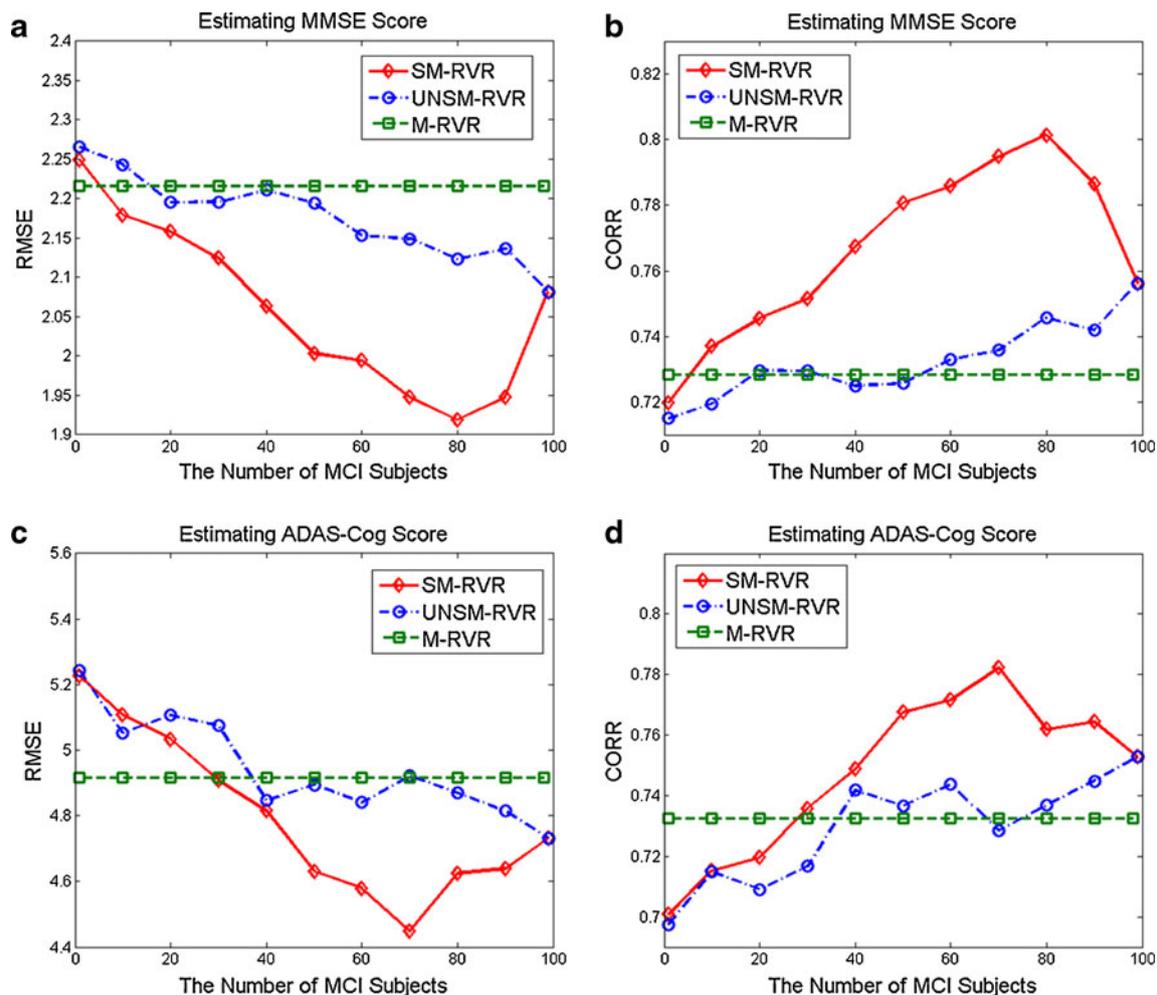
Methods	MMSE		ADAS-Cog	
	RMSE	CORR	RMSE	CORR
M-RVR (AD + NC)	2.22	0.73	4.92	0.73
M-RVR (AD + NC + MCI)	2.25	0.75	5.03	0.75
SM-RVR (AD + NC + MCI)	1.92	0.80	4.45	0.78

MMSE mini-mental state examination, ADAS-Cog Alzheimer's disease assessment scale-cognitive subscale, RMSE root mean square error, CORR correlation coefficient

subjects, not all MCI subjects, as unlabeled samples for estimation of clinical scores.

To further validate the efficacy of selecting the most informative MCI subjects for clinical score estimation, we also implemented a variant of SM-RVR, i.e., SM-RVR without selecting the most informative MCI subjects (denoted here as UNSM-RVR). Specifically, in the UNSM-

RVR method, MCI subjects (with their corresponding estimated clinical scores) are directly combined with AD and NC training subjects to train the M-RVR regression model. In Fig. 10, we also plotted the curves of regression performance measures (i.e., RMSE and CORR) with respect to the use of different number of unlabeled MCI samples when using SM-RVR and UNSM-RVR methods. Here, for UNSM-RVR, we randomly choose a part of MCI samples and directly merged them with the current training set (i.e., training AD and NC subjects) to train the regression model. From Fig. 10, we can observe that both curves show a rising trend in performance with the increase of MCI subjects in most cases, but SM-RVR is significantly and consistently better than UNSM-RVR in estimating both MMSE and ADAS-Cog scores. This implies that the use of selected MCI subjects as unlabeled samples is superior to using the selected MCI subjects as labeled samples, and also better than the use of all available MCI subjects in clinical score estimation.



**Fig. 10** Comparison of SM-RVR with UNSM-RVR and M-RVR, with respect to the use of different number of unlabeled samples (MCI subjects). **a** RMSE in estimating MMSE, **b** CORR in estimating MMSE, **c** RMSE in estimating ADAS-Cog, and **d** CORR in estimating ADAS-Cog

## Discussion

In this paper, we have proposed a new semi-supervised multimodal regression method, called SM-RVR, to estimate standard clinical cognitive test scores (including MMSE and ADAS-Cog) from multimodal brain imaging and non-imaging data. The experimental results on 202 subjects (with all acquired MRI, FDG-PET and CSF data) from ADNI database show that our method can consistently and substantially improve the regression performance, compared with the individual modality-based regression methods or the conventional supervised multimodal regression methods. Our experiment's results validate our assumption that the clinical scores of MCI subjects are less stable than those of AD and NC subjects due to the heterogeneity of MCI, and thus provide little help in further improving the performance of a supervised model trained using AD and NC subjects. On the other hand, the multimodal data of MCI subjects may contain more useful information than their corresponding clinical scores, and can lead to great performance improvement if used for training a semi-supervised model by treating the MCI subjects as unlabeled data and the AD and NC subjects as labeled data. As for clinical applications, our method can be used to help assess the pathological stage and further predict the future progression of AD or other neurological diseases.

### Semi-Supervised Learning

Semi-supervised learning is a recently developed machine learning technique, which learns from both labeled and unlabeled samples (Adams 2009; Belkin et al. 2004; Belkin and Niyogi 2004; Belkin et al. 2006; Ding and Zhao 2010; Ni et al. 2012). In the semi-supervised learning, one key issue is how to estimate labels for the new samples by using the relatedness between labeled and unlabeled samples. The existing semi-supervised learning methods deal with this issue in the following ways: (1) using a manifold hypothesis and the graph Laplacian for delivering the relatedness between labeled and unlabeled samples (Belkin et al. 2004; Belkin and Niyogi 2004; Belkin et al. 2006); (2) learning a propagable graph for semi-supervised classification and regression (Ni et al. 2012); (3) transductive learning and varifold Laplacian framework based on the hypothesis of weaker varifold structure (Bruzzone et al. 2006; Ding and Zhao 2010), etc.

Recently, some researchers have used semi-supervised classification methods to diagnose AD or MCI with brain images (Filipovych et al. 2011; Zhang and Shen 2011). In (Filipovych et al. 2011), MCI subjects are used as unlabeled samples, and the transductive SVM is adopted for MCI conversion classification. In (Zhang and Shen 2011), MCI subjects are also used as unlabeled samples while the multimodal LapRLS method is adopted for AD vs. NC classification. Both studies suggest that the semi-supervised methods can achieve better performance

than the corresponding supervised methods by using unlabeled data (MCI subjects in both cases). However, pattern classification methods can only assign a discrete category to each subject, and cannot deal with continuous scores in disease progression. Accordingly, in this paper, we proposed to use the semi-supervised *regression* model for estimating continuous clinical cognitive test scores from multimodal data. Experimental results demonstrate that our method achieved better regression performance than the conventional supervised regression methods.

### Multimodal Regression

Many recent studies have indicated that different modalities contain complementary information for discrimination of AD, and thus many works on combining different modalities of biomarkers have been reported for multimodal classification (Bouwman et al. 2007a; Chetelat et al. 2005; Fan et al. 2008b; Fellgiebel et al. 2007; Geroldi et al. 2006; Vemuri et al. 2009; Visser et al. 2002; Walhovd et al. 2010a). In those methods, features from all different modalities are typically concatenated into a longer feature vector for the purpose of multimodal classification. More recently, the multiple-kernel method has also been used for multimodal data fusion and classification, and achieves better performance than the feature concatenation method (Hinrichs et al. 2011; Hinrichs et al. 2009b; Zhang et al. 2011).

In this paper, we also compare M-RVR and SM-RVR with the corresponding feature concatenation method, with results shown in Table 4. As we can see from Table 4, both multi-kernel based methods (including SM-RVR and M-RVR) achieve better performance than the corresponding feature concatenation methods. For example, SM-RVR achieves a RMSE of 1.92 and a CORR of 0.80 for estimating MMSE scores, and a RMSE of 4.45 and a CORR of 0.78 for estimating ADAS-Cog scores, respectively, while the corresponding feature concatenation methods (namely

**Table 4** Comparison of regression performance of M-RVR vs. MConcat, and SM-RVR vs. SMConcat

Methods	MMSE		ADAS-Cog	
	RMSE	CORR	RMSE	CORR
MConcat	2.17	0.52	5.04	0.57
M-RVR	2.14	0.54	4.69	0.64
SMConcat	2.07	0.75	4.72	0.75
SM-RVR	1.92	0.80	4.45	0.78

*MMSE* mini-mental state examination, *ADAS-Cog* Alzheimer's disease assessment scale-cognitive subscale, *RMSE* root mean square error, *CORR* correlation coefficient, *MConcat* feature concatenation for supervised multimodal RVR, *SMConcat* feature concatenation for semi-supervised multimodal RVR

SMConcat) achieve only RMSE of 2.07 and 4.72, and CORR of 0.75 and 0.75 for estimating MMSE and ADAS-Cog scores, respectively.

We also perform experiments to compare RVR with another popular regression method, i.e., support vector regression (SVR), in multimodal regression. To do this, we replace RVR with SVR in both M-RVR and SM-RVR, and denote the corresponding methods as M-SVR and SM-SVR, respectively. Table 5 gives the comparisons between these different methods. It indicates that both RVR based methods (including SM-RVR and M-RVR) achieve better performance than the corresponding SVR based methods (including SM-SVR and M-SVR) on high-dimensional pattern regression.

### Estimating Clinical Scores

Several recent works in AD research have studied estimating the clinical scores from brain images (Fan et al. 2010; Stonnington et al. 2010; Wang et al. 2010). It is worth noting that most existing works on estimating clinical scores are based on the supervised regression models using a single-modal data. In our previous work (Zhang et al. 2012), we developed a multimodal multi-task learning model for jointly estimating multiple clinical scores from multimodal imaging data, showing improved performance over single-modal based methods. However, the model proposed in (Zhang et al. 2012) is still a supervised regression model and thus cannot effectively integrate MCI subjects (with less stable clinical scores) for helping build the regression model.

For a clear comparison, in Table 6 we listed several state-of-the-art methods on estimating clinical scores, along with their respective performances. Specifically, in (Wang et al. 2010), a Bagging RVR was adopted to estimate the MMSE score from the baseline MRI data, and the best result of RMSE of 3.29 and the correlation coefficient of 0.76 were achieved

on 23 AD, 74 MCI and 22 NC. In (Fan et al. 2010), a recursive feature elimination strategy for feature selection and RVR were adopted to estimate MMSE and ADAS-Cog scores from structural MRI data, and the best results of RMSE of 2.12 and 5.03 and the CORR of 0.57 and 0.52 were achieved on 52 AD, 148 MCI and 64 NC in estimating MMSE and ADAS-Cog scores, respectively. In (Stonnington et al. 2010), a RVR was adopted to estimate MMSE and ADAS-Cog scores from MRI data, and the best results of CORR of 0.7 and 0.57 were achieved on dataset 1 (73 AD, 91 NC), dataset 2 (113 AD, 351 MCI, 122 NC) and dataset 3 (39 AD, 92MCI, 32NC) in estimating MMSE and ADAS-Cog scores, respectively. In contrast, our proposed SM-RVR can achieve the best result of RMSE of 1.92 and the CORR of 0.80 for the MMSE score, and RMSE of 4.45 and the CORR of 0.78 for the ADAS-Cog score. These results validate the advantage of our method, compared with the conventional supervised and individual-modality-based regression methods.

### Predicting the Future Conversion of MCI Subjects

In addition to estimating the clinical scores, the proposed SM-RVR method can also be used for predicting the future conversion of MCI subjects, from the multimodal imaging and non-imaging biomarkers. Although the full investigation on this issue is beyond the main scope of the current paper, we examine the classification performance of our SM-RVR method using the same dataset with 51 AD, 52 NC, and 99 MCI (including 43 MCI converters (MCI-C) and 56 MCI non-converters (MCI-NC)). For training SM-RVR model for classification between MCI-C and MCI-NC, we adopt the 10-fold cross-validation strategy to divide the training and testing set for 99 MCI subjects, and use the training MCI subjects as unlabeled data together with all AD and NC subjects as labeled data to train a semi-supervised model. For comparison, we also report the classification results of M-RVR using the same training data as SM-RVR (i.e., training set of MCI subjects and all AD and NC subjects). However, different from SM-RVR, M-RVR trains a supervised model by treating the AD and MCI-C as one class and the NC and MCI-NC as the other class. On the dataset mentioned, our SM-RVR method achieves a classification accuracy of 69.4 % (with 72.4 % sensitivity and 67.8 % specificity), which is comparable to some recently reported results for MCI conversion prediction in the literature. In contrast, M-RVR only achieves a classification accuracy of 66.4 % (with 69.3 % sensitivity and 63.7 % specificity). Also, the M-RVR with only using the AD and NC as training set for MCI conversion prediction achieves a classification accuracy of only 58.4 % (with 64.1 % sensitivity and 54.1 % specificity). This again validates the advantage of our SM-RVR method over the conventional methods in predicting the future conversion of MCI to AD.

**Table 5** Comparison of regression performance of M-RVR vs. M-SVR, and SM-RVR vs. SM-SVR

Methods	MMSE		ADAS-Cog	
	RMSE	CORR	RMSE	CORR
M-SVR	2.16	0.53	4.87	0.62
M-RVR	2.14	0.54	4.69	0.64
SM-SVR	2.01	0.77	4.58	0.77
SM-RVR	1.92	0.80	4.45	0.78

MMSE mini-mental state examination, ADAS-Cog Alzheimer's disease assessment scale-cognitive subscale, RMSE root mean square error, CORR correlation coefficients, M-SVR supervised multimodal SVM regression, SM-SVR semi-supervised multimodal SVM regression, M-RVR supervised multimodal RVR, SM-RVR semi-supervised multimodal RVR

**Table 6** Comparison with the state-of-the-art methods

Method	Clinical score	Data set	Result
Wang et al. 2010	MMSE	23 AD, 74 MCI, 22 NC	<b>MMSE:</b> RMSE=3.29, CORR=0.76
Fan et al. 2010	MMSE, ADAS-Cog	52 AD, 148 MCI, 64 NC	<b>MMSE:</b> RMSE=2.12, CORR=0.57 <b>ADAS-Cog:</b> RMSE=5.03, CORR=0.52
Stonnington et al. 2010	MMSE, ADAS-Cog	<b>Set1</b> 73 AD, 91 NC	<b>Set1</b> <b>MMSE:</b> CORR=0.70
		<b>Set2</b> 113 AD, 351 MCI, 122 NC	<b>Set2</b> <b>MMSE:</b> CORR=0.48 <b>ADAS-Cog:</b> CORR=0.57
		<b>Set3</b> 39 AD, 92 MCI, 32 NC	<b>Set3</b> <b>MMSE:</b> CORR=0.51 <b>ADAS-Cog:</b> CORR=0.47
Zhang et al. 2012	MMSE, ADAS-Cog	45 AD, 91 MCI, 50 NC	<b>MMSE:</b> CORR=0.70 <b>ADAS-Cog:</b> CORR=0.74
Our proposed method	MMSE, ADAS-Cog	51 AD, 99 MCI, 52 NC	<b>MMSE:</b> RMSE=1.92, CORR=0.80 <b>ADAS-Cog:</b> RMSE=4.45, CORR=0.78

*RMSE* root mean square error, *CORR* correlation coefficients, *MMSE* mini-mental state examination, *ADAS-Cog* Alzheimer's disease assessment scale-cognitive subscale

### Clinical Implications

Current clinical practice in Alzheimer's disease (AD) relies of brain imaging and other diagnostic tests primarily to rule out other entities, such as vascular dementia. Although CSF amyloid and tau biomarker tests are commercially available, they are seldom used due primarily to the lack of an effective disease-modifying therapy. The ADNI consortium dataset provides a robust sample for investigating and validating clinically relevant biomarkers along the spectrum from MCI to AD. The wealth of data available from volumetric MRI brain scans, fluorodeoxyglucose PET imaging, and CSF amyloid and tau markers, together with a standard clinical assessment test battery are designed to facilitate the development of more useful clinic diagnostic tests and objective therapeutic outcome measures. The superior results achieved with SM-RVR using all three modalities are not surprising, as each of the three modalities reflects complementary and clinically relevant information. CSF amyloid and tau markers may be more useful in pre-clinical diagnosis and therapeutic monitoring of specific anti-amyloid and/or anti-tau therapeutic agents. Quantitative MRI measures of brain atrophy, particularly in the hippocampal regions, are more likely to reflect episodic memory deficits that are the *sine qua non* of AD and the amnesic form of MCI. However, variability across MCI subjects that do not conform to the "amnesic" subtype is likely to be accounted for better by the use of the SM-RVR multiple regression technique. Brain fluorodeoxyglucose PET imaging provides a measure of synaptic functioning and may be most useful as a predictor of clinical progression, as exemplified by

predicting conversion of MCI to AD. Depending on the specific application pertaining to diagnostic accuracy or therapeutic monitoring, SM-RVR methodology may be adapted to help streamline clinical therapeutic trials.

### Limitations

Our proposed method is limited in the following aspects. First, our proposed regression method separately estimates different clinical variables, without considering their inherent correlations and further exploiting the class labels to aid the accurate estimation of regression variables. Second, our proposed method is based on the multimodal data, i.e., MRI, PET, and CSF, and thus requires each subject to have the corresponding modality data, which limits the size of sample set that can be used for study. Besides MRI, PET, and CSF, there also exist other modalities of data, i.e., APOE, in ADNI database. However, due to the fact that not every subject has all available multimodal data, we are not able to investigate the contribution of those modality data in the current study, although in principle the inclusion of data from more modalities could further improve the regression performance.

### Conclusion

This paper proposes a novel semi-supervised multimodal regression method, namely SM-RVR, to estimate the clinical scores of subjects from both imaging and biological biomarkers, i.e., MRI, PET and CSF. Our method assumes

that the clinical scores obtained from MCI subjects may be less stable than those of AD and NC subjects due to the heterogeneity of MCI, and thus MCI subjects should be used as unlabeled data in a semi-supervised regression framework, rather than as labeled data in the conventional supervised regression framework. Furthermore, a scheme for selecting the most informative MCI subjects for helping train regression model is also developed. Experimental results on 202 baseline subjects from ADNI database show that our proposed method can achieve better performance than the state-of-the-art methods in estimating clinical scores based on pattern regression.

### Information Sharing Statement

The dataset used in this paper were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) which are available at <http://www.adni-info.org>. Source code and binary programs developed in this paper are available via our website (<http://bric.unc.edu/ideagroup/>). The RVM regression code is implemented using Sparse Bayesian toolbox, which is available at <http://www.miketipping.com/index.php?page=rvm>.

**Acknowledgments** Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: 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., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles.

This work was supported in part by NIH grants EB006733, EB008374, EB009634, and AG041721, by NSFC grants 61075010 and 61170151, by SRFDP grant 20123218110009, by Qing Lan Project, and also by The National Basic Research Program of China (973 Program) grant No. 2010CB732505.

### References

- Adams, N. (2009). Semi-supervised learning. *Journal of the Royal Statistical Society Series a-Statistics in Society*, 172, 530–530.
- Belkin, M., & Niyogi, P. (2004). Semi-supervised learning on Riemannian manifolds. *Machine Learning*, 56, 209–239.
- Belkin, M., Matveeva, I., & Niyogi, P. (2004). Regularization and semi-supervised learning on large graphs. *Learning Theory, Proceedings*, 3120, 624–638.
- Belkin, M., Niyogi, P., & Sindhvani, V. (2006). Manifold regularization: a geometric framework for learning from labeled and unlabeled examples. *Journal of Machine Learning Research*, 7, 2399–2434.
- Bouwman, F. H., Schoonenboom, S. N., van der Flier, W. M., van Elk, E. J., Kok, A., Barkhof, F., et al. (2007). CSF biomarkers and medial temporal lobe atrophy predict dementia in mild cognitive impairment. *Neurobiology of Aging*, 28, 1070–1074.
- Bouwman, F. H., van der Flier, W. M., Schoonenboom, N. S. M., van Elk, E. J., Kok, A., Rijmen, F., et al. (2007). Longitudinal changes of CSF biomarkers in memory clinic patients. *Neurology*, 69, 1006–1011.
- Bruzzzone, L., Chi, M. M., & Marconcini, M. (2006). A novel transductive SVM for semisupervised classification of remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 44, 3363–3373.
- Caselli, R. J., Reiman, E. M., Locke, D. E., Hutton, M. L., Hentz, J. G., Hoffman-Snyder, C., et al. (2007). Cognitive domain decline in healthy apolipoprotein E epsilon4 homozygotes before the diagnosis of mild cognitive impairment. *Archives of Neurology*, 64, 1306–1311.
- Caselli, R. J., Dueck, A. C., Osborne, D., Sabbagh, M. N., Connor, D. J., Ahern, G. L., et al. (2009). Longitudinal modeling of age-related memory decline and the APOE epsilon 4 effect. *The New England Journal of Medicine*, 361, 255–263.
- Chetelat, G., Eustache, F., Viader, F., De la Sayette, V., Pelerin, A., Mezenge, F., et al. (2005). FDG-PET measurement is more accurate than neuropsychological assessments to predict global cognitive deterioration in patients with mild cognitive impairment. *Neurocase*, 11, 14–25.
- de Leon, M. J., Mosconi, L., Li, J., De Santi, S., Yao, Y., Tsui, W. H., et al. (2007). Longitudinal CSF isoprostane and MRI atrophy in the progression to AD. *Journal of Neurology*, 254, 1666–1675.
- De Santi, S., de Leon, M. J., Rusinek, H., Convit, A., Tarshish, C. Y., Roche, A., et al. (2001). Hippocampal formation glucose metabolism and volume losses in MCI and AD. *Neurobiology of Aging*, 22, 529–539.
- Diehl, J., Grimmer, T., Drzezga, A., Riemenschneider, M., Forstl, H., & Kurz, A. (2004). Cerebral metabolic patterns at early stages of frontotemporal dementia and semantic dementia. A PET study. *Neurobiology of Aging*, 25, 1051–1056.
- Ding, L., & Zhao, P. B. (2010). Semi-supervised learning with varifold Laplacians. *Neurocomputing*, 73, 1580–1586.
- Drzezga, A., Lautenschlager, N., Siebner, H., Riemenschneider, M., Willoch, F., Minoshima, S., et al. (2003). Cerebral metabolic changes accompanying conversion of mild cognitive impairment into Alzheimer's disease: a PET follow-up study. *European Journal of Nuclear Medicine and Molecular Imaging*, 30, 1104–1113.
- Duchesne, S., Caroli, A., Geroldi, C., Collins, D. L., & Frisoni, G. B. (2009). Relating one-year cognitive change in mild cognitive impairment to baseline MRI features. *NeuroImage*, 47, 1363–1370.
- Fan, Y., Batmanghelich, N., Clark, C. M., Davatzikos, C., & Initia, A. D. N. (2008). 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. (2008). Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study. *NeuroImage*, 41, 277–285.
- Fan, Y., Kaufner, D., Shen, D., (2010). Joint estimation of multiple clinical variables of neurological diseases from imaging patterns. *In Proceedings of 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI 2010)*, 852–855.

- Fellgiebel, A., Scheurich, A., Bartenstein, P., & Muller, M. J. (2007). FDG-PET and CSF phospho-tau for prediction of cognitive decline in mild cognitive impairment. *Psychiatry Research-Neuroimaging*, *155*, 167–171.
- Filipovych, R., Davatzikos, C., & Initia, A. D. N. (2011). Semi-supervised pattern classification of medical images: application to mild cognitive impairment (MCI). *NeuroImage*, *55*, 1109–1119.
- Fjell, A. M., Walhovd, K. B., Fennema-Notestine, C., McEvoy, L. K., Hagler, D. J., Holland, D., et al. (2010). CSF biomarkers in prediction of cerebral and clinical change in mild cognitive impairment and Alzheimer's disease. *Journal of Neuroscience*, *30*, 2088–2101.
- Geroldi, C., Rossi, R., Calvagna, C., Testa, C., Bresciani, L., Binetti, G., et al. (2006). Medial temporal atrophy but not memory deficit predicts progression to dementia in patients with mild cognitive impairment. *Journal of Neurology, Neurosurgery, and Psychiatry*, *77*, 1219–1222.
- Hinrichs, C., Singh, V., Mukherjee, L., Xu, G. F., Chung, M. K., & Johnson, S. C. (2009a). Spatially augmented LPboosting for AD classification with evaluations on the ADNI dataset. *NeuroImage*, *48*, 138–149.
- Hinrichs, C., Singh, V., Xu, G., & Johnson, S. (2009b). MKL for robust multi-modality AD classification. *Medical Image Computing and Computer Assisted Intervention*, *12*, 786–794.
- Hinrichs, C., Singh, V., Xu, G., & Johnson, S. C. (2011). Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population. *NeuroImage*, *55*, 574–589.
- Inoue, M., Jimbo, D., Taniguchi, M., & Urakami, K. (2011). Touch Panel-type Dementia Assessment Scale: a new computer-based rating scale for Alzheimer's disease. *Psychogeriatrics*, *11*, 28–33.
- Ni, B. B., Yan, S. C., & Kassim, A. A. (2012). Learning a propagable graph for semisupervised learning: classification and regression. *Ieee Transactions on Knowledge and Data Engineering*, *24*, 114–126.
- Rakotomamonjy, A., Bach, F. R., Canu, S., & Grandvalet, Y. (2008). SimpleMKL. *Journal of Machine Learning Research*, *9*, 2491–2521.
- Reiman, E. M., Caselli, R. J., Yun, L. S., Chen, K. W., Bandy, D., Minoshima, S., et al. (1996). Preclinical evidence of Alzheimer's disease in persons homozygous for the epsilon 4 allele for apolipoprotein E. *The New England Journal of Medicine*, *334*, 752–758.
- Reiman, E. M., Chen, K. W., Liu, X. F., Bandy, D., Yu, M. X., Lee, W. D., et al. (2009). Fibrillar amyloid-beta burden in cognitively normal people at 3 levels of genetic risk for Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 6820–6825.
- Rosen, W. G., Mohs, R. C., & Davis, K. L. (1984). A new rating scale for Alzheimer's disease. *The American Journal of Psychiatry*, *141*, 1356–1364.
- Shattuck, D. W., Sandor-Leahy, S. R., Schaper, K. A., Rottenberg, D. A., & Leahy, R. M. (2001). Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*, *13*, 856–876.
- Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging*, *17*, 87–97.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*, 143–155.
- Stonnington, C. M., Chu, C., Kloppel, S., Jack, C. R., Ashburner, J., Frackowiak, R. S. J., et al. (2010). Predicting clinical scores from magnetic resonance scans in Alzheimer's disease. *NeuroImage*, *51*, 1405–1413.
- Tipping, M. E. (2001). Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, *1*, 211–244.
- Twamley, E. W., Ropacki, S. A. L., & Bondi, M. W. (2006). Neuropsychological and neuroimaging changes in preclinical Alzheimer's disease. *Journal of the International Neuropsychological Society*, *12*, 707–735.
- Vemuri, P., Wiste, H. J., Weigand, S. D., Shaw, L. M., Trojanowski, J. Q., Weiner, M. W., et al. (2009). MRI and CSF biomarkers in normal, MCI, and AD subjects predicting future clinical change. *Neurology*, *73*, 294–301.
- Visser, P. J., Verhey, F. R. J., Hofman, P. A., Scheltens, P., & Jolles, J. (2002). Medial temporal lobe atrophy predicts Alzheimer's disease in patients with minor cognitive impairment. *Journal of Neurology, Neurosurgery, and Psychiatry*, *72*, 491–497.
- Walhovd, K. B., Fjell, A. M., Brewer, J., McEvoy, L. K., Fennema-Notestine, C., Hagler, D. J., Jr., et al. (2010). Combining MR imaging, positron-emission tomography, and CSF biomarkers in the diagnosis and prognosis of Alzheimer disease. *AJNR. American Journal of Neuroradiology*, *31*, 347–354.
- Walhovd, K. B., Fjell, A. M., Dale, A. M., McEvoy, L. K., Brewer, J., Karow, D. S., et al. (2010b). Multi-modal imaging predicts memory performance in normal aging and cognitive decline. *Neurobiology of Aging*, *31*, 1107–1121.
- Wang, Z., Chen, S. C., & Sun, T. K. (2008). MultiK-MHKS: a novel multiple kernel learning algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*, 348–353.
- Wang, Y., Fan, Y., Bhatt, P., & Davatzikos, C. (2010). High-dimensional pattern regression using machine learning: from medical images to continuous clinical variables. *NeuroImage*, *50*, 1519–1535.
- Xu, Z., Jin, R., Yang, H., King, I., Lyu, M. R. (2010). Simple and Efficient Multiple Kernel Learning by Group Lasso. *In Proceedings of the 27th Conference on Machine Learning (ICML 2010)*.
- Zhang, D., & Shen, D. (2011). Semi-supervised multimodal classification of Alzheimer's disease. *In Proceedings of 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI 2011)*, 1628–1631.
- Zhang, D., & Shen, D. (2012). Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. *NeuroImage*, *59*, 895–907.
- Zhang, Y. Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, *20*, 45–57.
- Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D., & Initia, A. D. N. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. *NeuroImage*, *55*, 856–867.