

Evaluating the Predictive Power of Multivariate Tensor-based Morphometry in Alzheimers Disease Progression via Convex Fused Sparse Group Lasso

Sinchai Tsao^a Niharika Gajawelli^b Jiayu Zhou^c Jie Shi^c
Jieping Ye^c Yalin Wang^c Natasha Lepore^b

^aUniversity of Washington, Seattle, Washington, USA

^bChildren’s Hospital Los Angeles, Los Angeles, California, USA

^cArizona State University, Phoenix, Arizona, USA

ABSTRACT

Prediction of Alzheimers disease (AD) progression based on baseline measures allows us to understand disease progression and has implications in decisions concerning treatment strategy. To this end we combine a predictive multi-task machine learning method¹ with novel MR-based multivariate morphometric surface map of the hippocampus² to predict future cognitive scores of patients. Previous work by Zhou et al.¹ has shown that a multi-task learning framework that performs prediction of all future time points (or tasks) simultaneously can be used to encode both sparsity as well as temporal smoothness. They showed that this can be used in predicting cognitive outcomes of Alzheimers Disease Neuroimaging Initiative (ADNI) subjects based on FreeSurfer-based baseline MRI features, MMSE score demographic information and ApoE status. Whilst volumetric information may hold generalized information on brain status, we hypothesized that hippocampus specific information may be more useful in predictive modeling of AD. To this end, we applied Shi et al.²s recently developed multivariate tensor-based (mTBM) parametric surface analysis method to extract features from the hippocampal surface. We show that by combining the power of the multi-task framework with the sensitivity of mTBM features of the hippocampus surface, we are able to significantly improve predictive performance of ADAS cognitive scores 6, 12, 24, 36 and 48 months from baseline.

Keywords: Alzheimers Disease, Disease Progression, Multi-task learning, fused Lasso, ADAS-Cog, Tensor-based Morphometry, Hippocampus, Feature Selection

1. INTRODUCTION

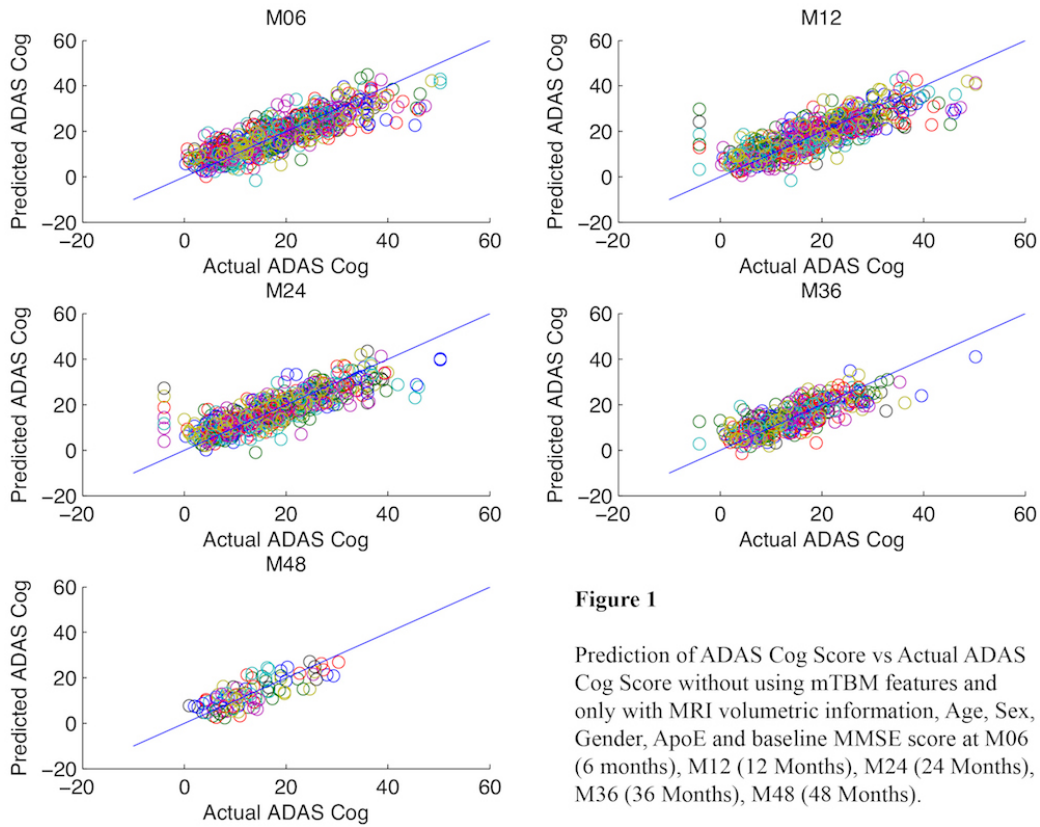
Recent work in psychological testing,³ genetic studies,⁴ magnetic resonance (MR) imaging,⁵ positron emission tomography (PET) imaging,⁶ cerebral spinal fluid (CSF) measurements,⁷ cardiovascular status⁸ and others have yielded tremendous amounts of diagnostic data for diagnosing and staging dementias, especially Alzheimers disease (AD). Moreover, many of these studies now also include longitudinal information.^{3,9} This has lead to a problem often referred to as the curse of dimensionality, where the size (number of dimensions) of the dataset makes it difficult to do various numerical analysis on the data. This in turn makes it increasingly difficult to draw consistent conclusions from the dataset. Statistical analysis together with clinical disease models have helped with determine how the different sets of diagnostic information interacts with one another but they require a large number of ad hoc assumptions and therefore does not lend itself well to large scale Medical Imaging-based features. These problems become even more important when trying to use machine learning techniques because at some point the predictive power of the model ceases to increase even though we’re adding more information or dimensions. The question is then about how to select the ”correct” features to maximize predictive power. This paper leverages existing sparsifying machine learning techniques with temporal priors,¹ built specifically for progressive disease models, such as AD, together with multivariate tensor-based morphometric (mTBM)

Further author information: (Send correspondence to S.T.)

S.T.: E-mail: mail@sinchaitso.com

N.L.: E-mail: lepore@usc.edu

features¹⁰ of the Hippocampus to try and predict AD progression up to 48 months from the baseline MRI measurement. The goal is to evaluate the predictive power of mTBM against those of cortical thickness and other FreeSurfer-based features, demographic information (sex and age) as well as genetic information (ApoE- ϵ 4 Copies).



2. METHODS

2.1 ADNI Data

Data used in the preparation of this article were obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 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 Alzheimers 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 lessen the time and cost of clinical trials.

The Principal Investigator of this initiative is Michael W. Weiner, MD, 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 subjects but ADNI has been followed by ADNI-GO and ADNI-2. To date these three protocols have recruited over 1500 adults, ages 55 to 90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with

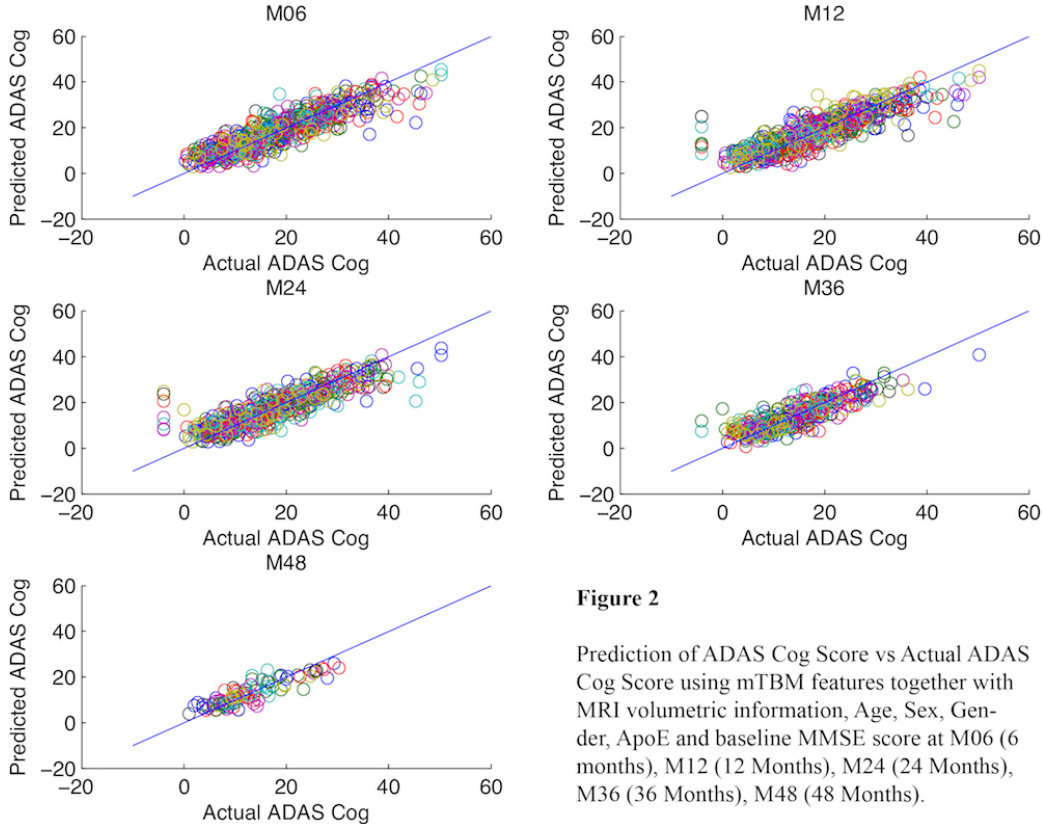


Figure 2
 Prediction of ADAS Cog Score vs Actual ADAS Cog Score using mTBM features together with MRI volumetric information, Age, Sex, Gender, ApoE and baseline MMSE score at M06 (6 months), M12 (12 Months), M24 (24 Months), M36 (36 Months), M48 (48 Months).

early AD. The follow up duration of each group is specified in the protocols for ADNI-1, ADNI-2 and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see www.adni-info.org.

For our experiment we used 616 subjects for M06, 606 for M12, 533 for M24, 364 for M36 and 97 for M48. 90% of the data was used for training and 10% used for testing. The reported results are for 20 different selection splits of training and testing. More information about the demographics and patient selection is available in Zhou et al 2013.¹

2.2 convex Fused Sparse Group Lasso (cFSGL)

Zhou et al 2013¹ has proposed a powerful multi-task learning technique that incorporates sparsity as well as temporal smoothing for modeling a progressive disease model. In their formulation, each task can be thought of a single forward predictor from baseline measurement to a measurement at a certain future time point. In their case, they used the ADNI dataset and predicted ADAS cognitive scores 6 months after baseline (M06), 12 months after baseline (M12), 24 months after baseline (M24), 36 months after baseline (M36) and 48 months after baseline (M48). In our study we aim to use the same ADNI dataset but also incorporate mTBM hippocampus features and compare it to features used in their study. We also attempt to combine the different feature sets to try to evaluate the predictive power of each set of features.

The proposed cFSGL can be considered a multi-task regression problem with t time points and from n subjects each with d features, where $\{x_1, x_2, \dots, x_n\}$ represents each of the d input features for each subject at baseline (i.e. $x_i \in \mathbb{R}^d$). Similarly, $\{y_1, y_2, \dots, y_N\}$ represents the target cognitive scores for each subject at N time points (i.e. $y_i \in \mathbb{R}^N$). For a single subject (n) each task can be seen as a projection of MR / demographic / genetic baseline measurements at $t = 0$ represented at x_n to a future cognitive score measurement at time $t = t_1$ (e.g. at 48 months) given by $y_n(t_1)$. We can extend this formulation to a multi-task one by performing

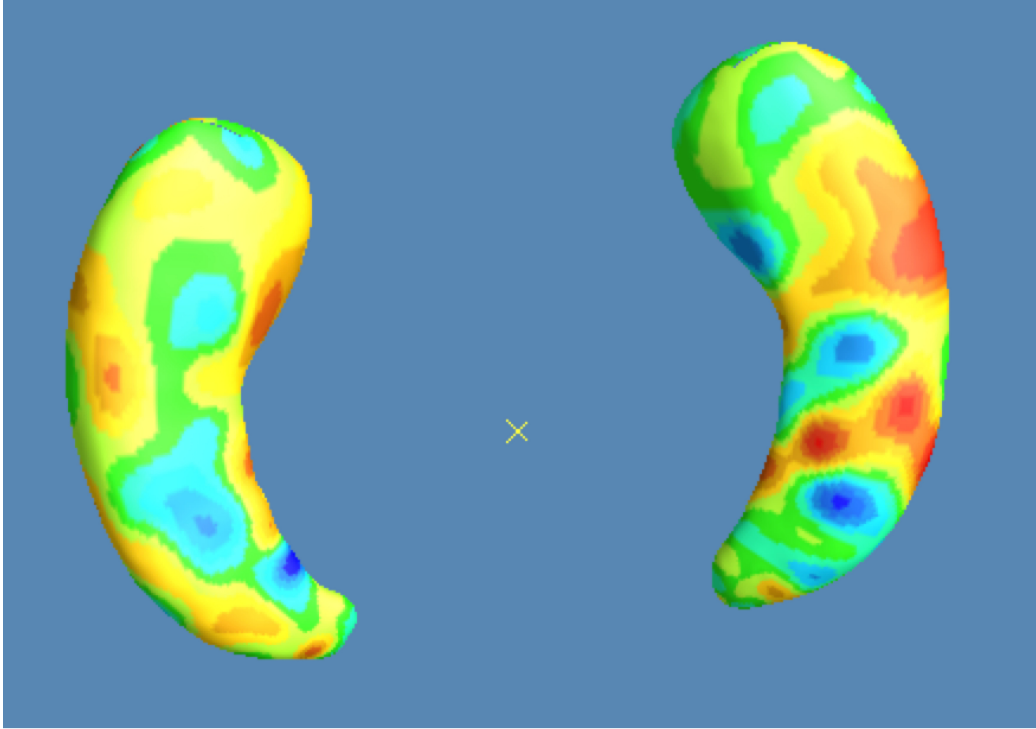


Figure 3
Average Weights for mTBM Feature 1 used for Prediction of Disease Progression

projections of all time points simultaneously. In other words, each set of baseline measurements at $t = 0$ given by x_n is projected to a vector (\mathbb{R}^N with N time points) given by y_1 . The entire mapping can be summarized as a linear operation using matrices X and Y . X and Y is formed by arranging the patient feature space row-wise, each row being x_n or y_N , and yields a $\mathbb{R}^{n \times d}$ X matrix and a $\mathbb{R}^{n \times N}$ Y matrix. Since this is a linear model, a set of weights W ($\mathbb{R}^{d \times N}$) is trained to map x_n to y_n or X to Y . To achieve a set of weights that encodes both sparsity and temporal smoothness. The following cost function is minimized during training.

$$\min_W \|XW - Y\|_F^2 + \lambda_1 \|W\|_1 + \lambda_2 \|RW^T\|_1 + \lambda_3 \|W\|_{2,1} \quad (1)$$

where $\|W\|_1$ is the L1-norm or lasso penalty that encodes for sparsity, $\|W\|_{2,1} = \sum_{i=1}^d \sqrt{\sum_{j=1}^t W_{ij}^2}$ is the group Lasso penalty that encodes for temporal grouping of features, $\|RW^T\|_1$ is the fused lasso penalty, $R = H^T$ and $H \in \mathbb{R}^{t \times (t-1)}$ where $H_{ij} = 1$ if $i = j$, $H_{ij} = -1$ if $i = j+1$, and $H_{ij} = 0$ otherwise that encodes for temporal smoothness.

2.3 Multivariate Tensor-based Morphometry (mTBM) features

After automatically segmenting hippocampus with FSL¹¹ from brain MR images, we build parametric meshes to model hippocampal shapes. High-order correspondences between hippocampal surfaces were enforced across subjects with a novel inverse consistent surface fluid registration method. Multivariate statistics consisting of multivariate tensor-based morphometry (mTBM) and radial distance were computed for surface deformation analysis.²

Table 1

Comparison of the model performance in predicting ADAS Cognitive Score with and without mTBM features. The base set of features used were MRI volumetric information, Sex, Gender, Age, ApoE and baseline MMSE score.

	w/o mTBM	with mTBM
nMSE	0.345 ± 0.075	0.249 ± 0.039
wR	0.828 ± 0.036	0.873 ± 0.022
M06 rMSE	5.259 ± 0.872	4.534 ± 0.883
M12 rMSE	5.653 ± 1.143	4.989 ± 1.134
M24 rMSE	5.532 ± 1.029	4.885 ± 1.094
M36 rMSE	4.777 ± 0.833	4.055 ± 1.024
M48 rMSE	4.367 ± 1.179	3.164 ± 1.091

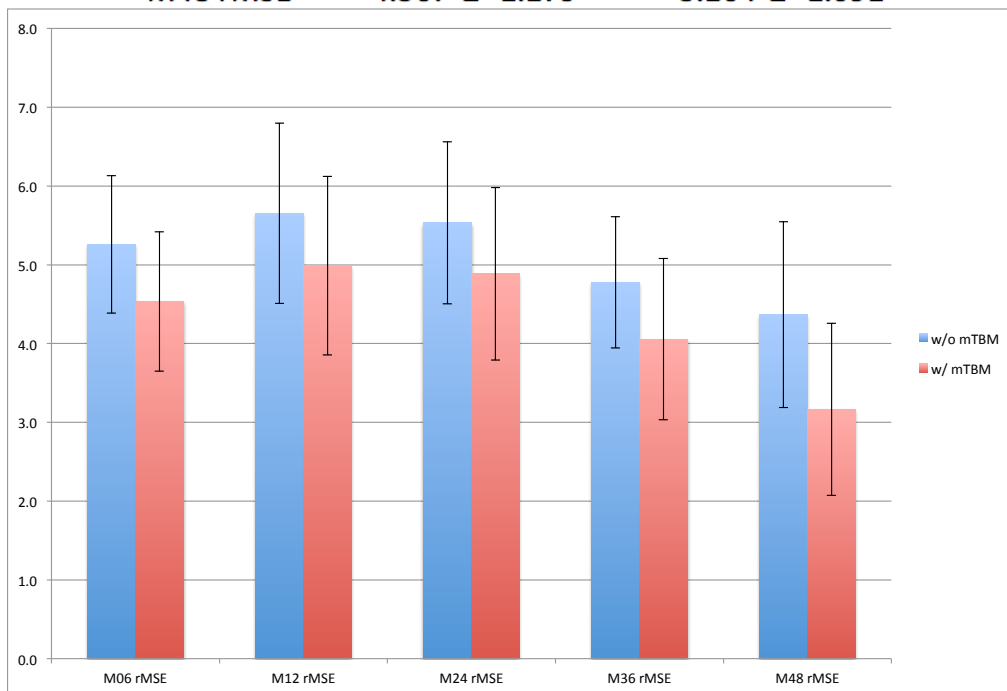


Figure 4: Bar Chart of the rMSE of predictions with and without mTBM features by time point

3. RESULTS

Predictions using mTBM significantly outperform prediction without using mTBM as shown in Figures 1 and 2. Quantitative measures such as nMSE, wR and rMSE show across the board improvements as shown in Table 1 and Figure 4. Average weights for one of the mTBM features across the 20 trials is shown in Figure 3.

4. DISCUSSION AND CONCLUSIONS

By merging fused multi-task learning that encodes temporal smoothing¹ together with AD sensitive mTBM maps of the parametric hippocampus surface², we were able to get significant gains in future ADAS cognitive score prediction. We believe that these results are some of the highest performing predictions based on baseline data only and is consistent with our survey of other comparable studies.¹ Other factors not addressed in this work is the effect of percentage of data used for training and testing. Previous work¹ has shown that although there would be a decrease in performance measured with a smaller training set, the trends and relative performance remains comparable. We have also treated the parametric surface data, patient demographics and MRI volumetric information as one continuous information vector. It would be interesting to see if adding neighborhood information based on the location on the parametric surface would give us smoother and more realistic weights on the parametric surface and perhaps even better or more consistent results.

The current study also serves as a illustration of how machine learning methods can be used with whole parametric surfaces or even volumetric volumes such as in fMRI studies. However, as the number of voxels and vertex points increase, we again run into problems with the curse of dimensionality. To counter such problems, sparsifying penalties such as in cFSGL can be employed. However, without a reasonable starting weight, finding a reasonable solution that has the required sparsity can get computational intensive. One solution that we intend to explore is the use of stability selection in seeding the initial weights for the algorithm in a hierarchical approach to learning. We believe that this a reasonable way of leveraging prior information whilst allowing the algorithm to impose explore ensure temporal smoothness and sparsity.

As this is a model of a epidemiological system, we cannot ignore the investigator’s selection of reasonable features. Moreover, the performance of the system is as interesting as the weights that yield the predictions. Our future work includes work in understanding the behavior of the weights across the parametric surface space as well as in time. Previous work has shown that stability selection may be a good fit for analyzing the feature weights on the model.

5. FUTURE WORK

Future work including stability analysis of the weights may yield more information about the relationship between the deformation of hippocampal subfields and other clinical indicators during AD progression.

6. ACKNOWLEDGEMENTS

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

REFERENCES

- [1] Zhou, J., Liu, J., Narayan, V. A., and Ye, J., “Modeling Disease Progression via Multi-task Learning,” *NeuroImage* (2013).
- [2] Shi, J., Thompson, P. M., Gutman, B., and Wang, Y., “Surface fluid registration of conformal representation: Application to detect disease burden and genetic influence on hippocampus,” *NeuroImage* (2013).
- [3] Caselli, R. J., Locke, D. E. C., Dueck, A. C., Knopman, D. S., Woodruff, B. K., Hoffman-Snyder, C., Rademakers, R., Fleisher, A. S., and Reiman, E. M., “The neuropsychology of normal aging and preclinical alzheimer’s disease,” *Alzheimers Dement* (Mar 2013).
- [4] Elias-Sonnenschein, L. S., Helisalmi, S., Natunen, T., Hall, A., Paaajanen, T., Herukka, S.-K., Laitinen, M., Remes, A. M., Koivisto, A. M., Mattila, K. M., Lehtimäki, T., Verhey, F. R. J., Visser, P. J., Soininen, H., and Hiltunen, M., “Genetic loci associated with alzheimer’s disease and cerebrospinal fluid biomarkers in a finnish case-control cohort,” *PLoS One* **8**(4), e59676 (2013).
- [5] Teipel, S. J., Grothe, M., Lista, S., Toschi, N., Garaci, F. G., and Hampel, H., “Relevance of magnetic resonance imaging for early detection and diagnosis of alzheimer disease,” *Med Clin North Am* **97**, 399–424 (May 2013).
- [6] Becker, G. A., Ichise, M., Barthel, H., Luthardt, J., Patt, M., Seese, A., Schultze-Mosgau, M., Rohde, B., Gertz, H.-J., Reiningner, C., and Sabri, O., “Pet quantification of 18f-florbetaben binding to -amyloid deposits in human brains,” *J Nucl Med* (Mar 2013).
- [7] Blennow, K. and Zetterberg, H., “The application of cerebrospinal fluid biomarkers in early diagnosis of alzheimer disease,” *Med Clin North Am* **97**, 369–76 (May 2013).
- [8] Hajjar, I., Brown, L., Mack, W. J., and Chui, H., “Alzheimer pathology and angiotension receptor blockers,” *JAMA Neurol* **70**, 414 (Mar 2013).
- [9] Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C. R., Jagust, W., Trojanowski, J. Q., Toga, A. W., and Beckett, L., “Ways toward an early diagnosis in alzheimer’s disease: the alzheimer’s disease neuroimaging initiative (adni),” *Alzheimers Dement* **1**, 55–66 (Jul 2005).
- [10] Wang, Y., Yuan, L., Shi, J., Greve, A., Ye, J., Toga, A. W., Reiss, A. L., and Thompson, P. M., “Applying tensor-based morphometry to parametric surfaces can improve mri-based disease diagnosis,” *Neuroimage* **74**, 209–30 (Jul 2013).
- [11] Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., and Smith, S. M., “Fsl,” *Neuroimage* **62**, 782–90 (Aug 2012).