Supplemental Information Document

Authors: Raj A*, LoCastro E, Kuceyeski A, Tosun D, Relkin N, Weiner M and the Alzheimer's Disease Neuroimaging Initiative (ADNI)

Supplemental Experimental Procedures

1. Extraction of discordant regions, i.e. those that have much higher end-of-study atrophy than would be predicted from baseline atrophy

We investigated which brain regions displayed the most discordant atrophy in our MRI-derived volumetric series from ADNI cohort. These are the regions that fall within the "off-diagonal" areas in Figure 3, as described in the main body of the paper. In order to determine these regions, we back-projected the points shown in the scatter plots of Figure 3, and calculated how often each brain region fell within this off-diagonal zone. For the purpose of this exercise, the off-diagonal zone was defined as any point whose end atrophy was more than 3 times its baseline atrophy. The resulting data was converted to a histogram, as a ratio, such that a region mapped to 0 on this histogram would never end up in the discordant zone, whereas a region mapped to 1 would always end up in the discordant zone. These histogram ratios were then plotted on the whole brain using our standard glass-brain rendering, such that each region is depicted by a sphere. The size of the sphere is proportional to this discordant histogram ratio, and its color represents the lobe to which the region belongs. This data is shown in Figure SI-2. Only the MCI-converter group is being shown here.

2. Characterization of robustness via Monte Carlo simulation of additive noise in reference connectome

The issue of robustness of the model against noise in the connectome is critical, since the model is based on the connectome. We investigate this here, as follows. Starting from the observed

connectome data, we added increasing amounts of noise to each connection, drawn from an independent and identically distributed (i.i.d.) random noise of increasing variance. Then, for each variance level, Monte Carlo simulations with 100 trials were run and the R statistic of the Pearson correlation between the subjects' atrophy and the (noise-corrupted) model computed. The mean R statistic over 100 trials was reported at each noise variance level. The results of this analysis are contained in Figure SI-7. The predictions appear reasonably tolerant to moderate levels of connectome noise.

3. Characterization of robustness via Bootstrap analysis of variability among subjects

The issue of robustness against natural variability between subjects is also critical. While we showed above that the choice of connectome and its noise do not unduly affect the performance of our predictors, the issue of variability in patient data remains. How sensitive are the presented results to the choice of particular subjects used in our study? How internally consistent are the atrophy maps of each disease group? The most principled way we know of exploring the effect of variability in available samples is **Bootstrap resampling test**, a well-known statistical resampling method (Severiano et al. 2011), which we are proposing in response to these concerns. Bootstrap sampling involves drawing a large number of samples from the existing set of samples, but doing this independently for each sample, and allowing replacements – i.e. a single sample from the set may occur multiple times in any of the bootstrap samples. By repeatedly resampling from the available data a subset of subjects, and gathering the outcome variable in each case, this technique builds a distribution of the likely variability in the outcome variable due to sampling errors that reflect the inter-subject variability. From well-known principles, this histogram is then thought to reflect the natural variability expected from this sample.

We performed a thorough bootstrap test to address this issue, as follows: We characterized the effect of inter-subject variability in the measured data from the ADNI cohort, by resampling the patients in each diagnostic group. For each sample we recomputed both MRI-derived atrophy and PET-derived metabolism maps, applied the network diffusion model to it and computed the correlation of model prediction with measured atrophy t-statistic of the entire ADNI cohort. 1000 bootstrap samples with replacement of each disease group were evaluated for resulting changes in the outcome variable, in this case the correlation R of the ND model against ADNI t-statistics. The histogram of this statistic was estimated using the non-parametric kernel smoothing density estimation method (Bowman et al. 1997). These distributions are plotted in Figure SI-8, representing a total of 6 bootstrap experiments: 3 disease groups (MCI-N, MCI-C, AD) X 2 modalities (MRI, FDG-PET). The distributions confirm that the variability is indeed small, thus demonstrating that inter-subject variability in the ADNI cohort did not cause undesirable variability in the predictor. This analysis (see Figure SI-8) revealed that our results are highly robust to inter-subject variability and well within the range expected from sampling errors. We also obtained 95% confidence intervals around each reported R statistic, which is listed in Table 3 and also indicated in Figure SI-8 by vertical lines. In addition, there was almost zero bias between the reported statistics and the mean of 1000 bootstrap samples.

4. Detailed Methods and Acknowledgments pertaining to ADNI

ADNI-related Methods:

Data used in the preparation of this article were obtained from the Alzheimer's 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 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 lessen the time and cost of clinical trials.

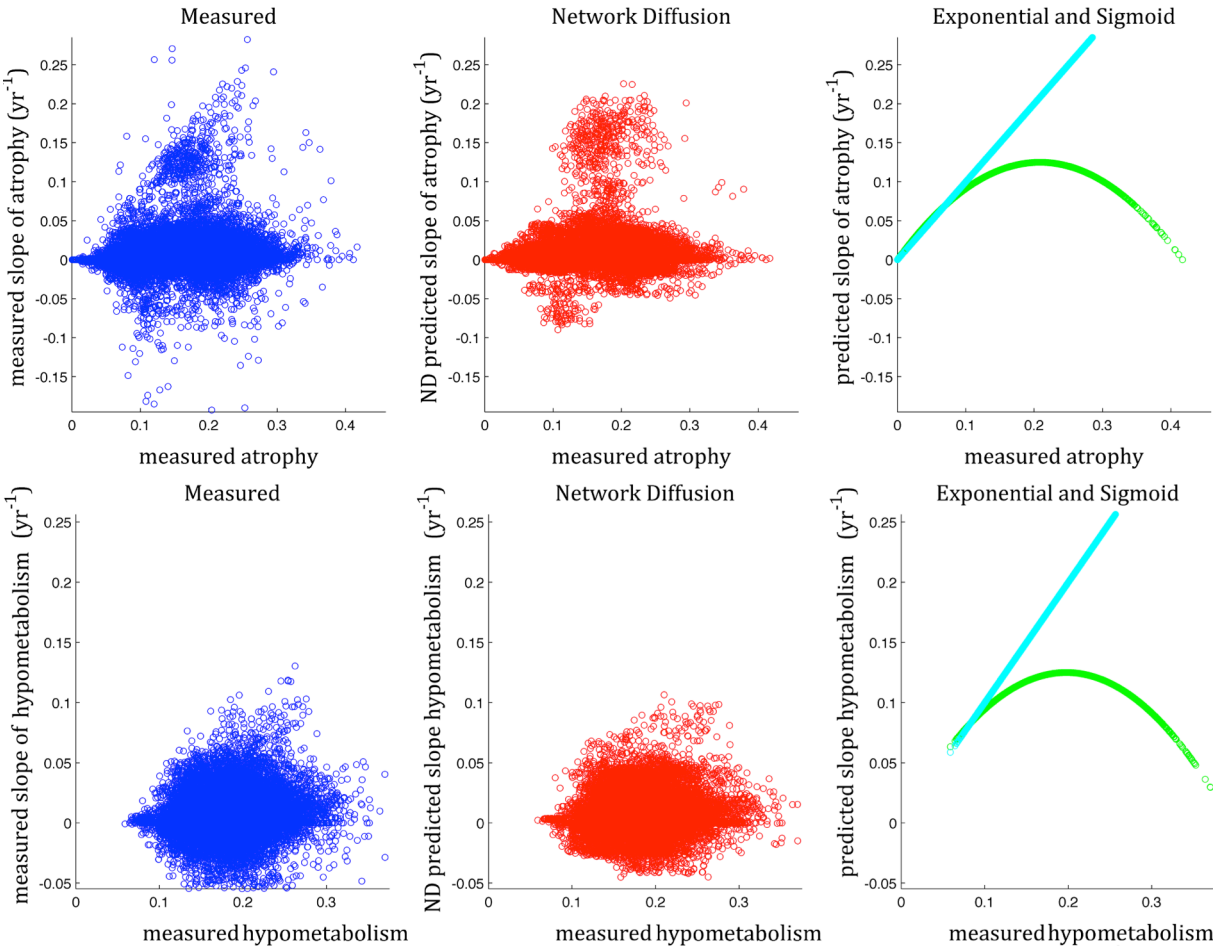
The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California – San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 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 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.

ADNI 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: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its

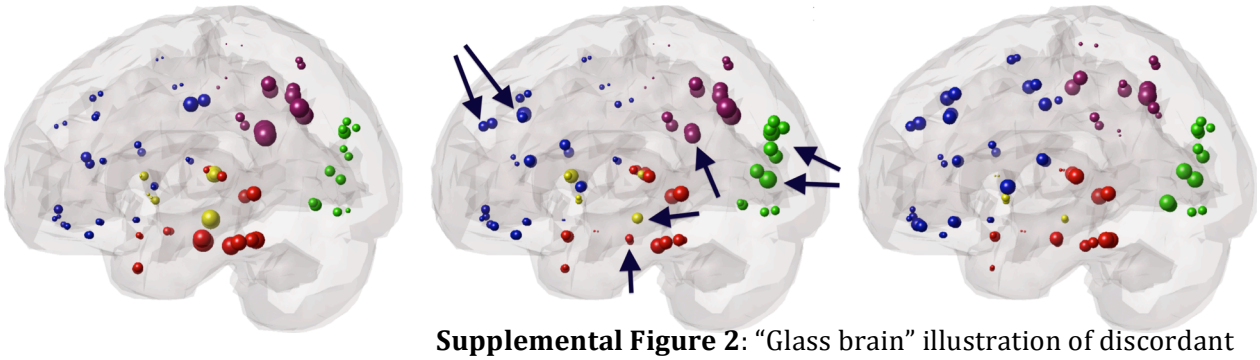
affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; ; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; 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.

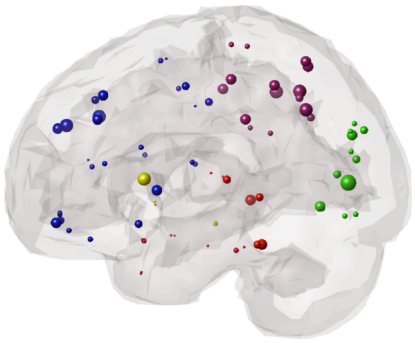
Supplemental Figures



Supplemental Figure 1: **The relationship between regional atrophy/hypometabolism and their local rate of change of the entire ADNI patient cohort**. The top panel pertains to MRI-derived atrophy, and bottom panel to FDG-PET-derived hypometabolism. Left: measured data, middle: ND model predictions, right: predictions from both exponential and sigmoid models. The ND model is able to capture the essential atrophy-slope relationship in all its complexity. **Related to Figure 1.**

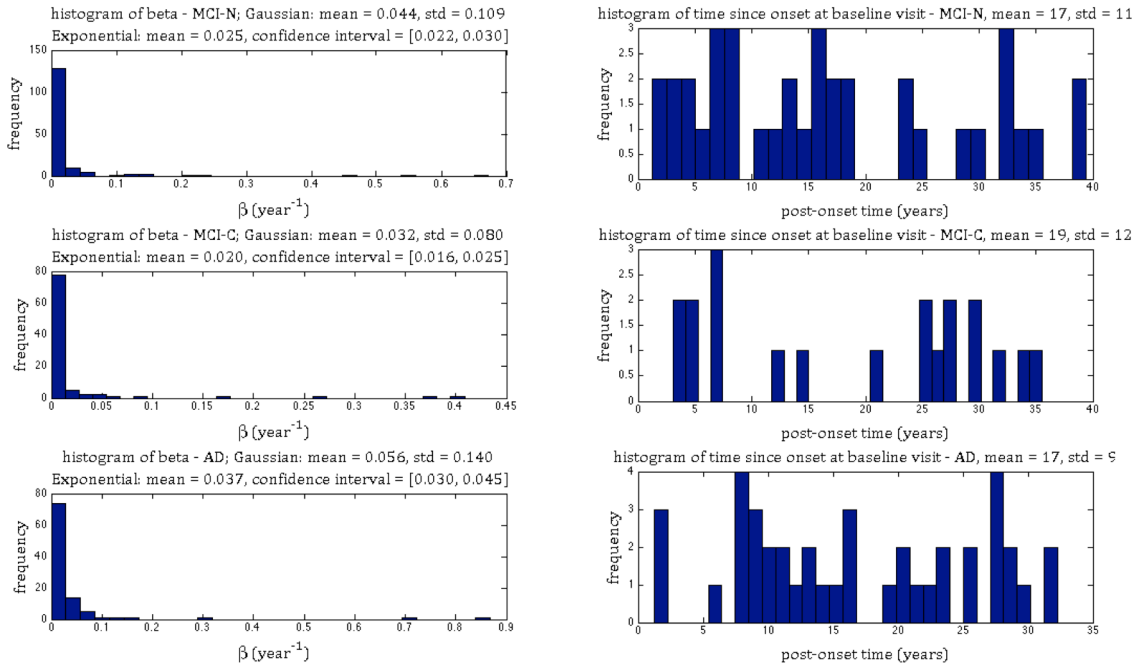
Baseline metabolism, change of metabolism and network diffusion predicted change



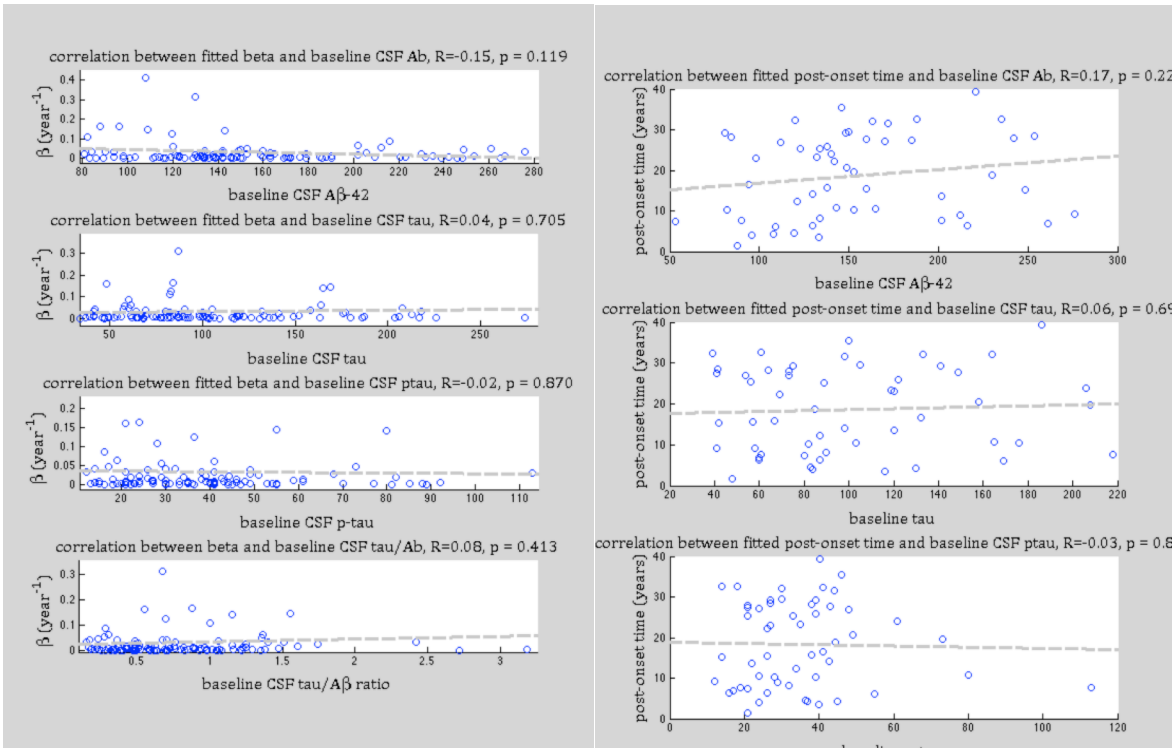


Supplemental Figure 2: "Glass brain" illustration of discordant regions. **Top**: The relationship between regional disease pattern and its rate of change. The spheres are proportional to group t-statistics over all ADNI patients: baseline FDG-derived regional hypometabolism (left), its rate of change evaluated at each region (middle), and the model predicted rate of change (right). Spheres are color-coded by lobe - frontal=blue, parietal=purple, occipital=green, temporal=red, subcortical=yellow. The baseline pattern is not a good predictor of change in frontal and occipital regions. Black arrows point to specific structures that are discordant, including mesial temporal, frontal and occipital structures, where the network diffusion model appears to be a better predictor of change of atrophy.

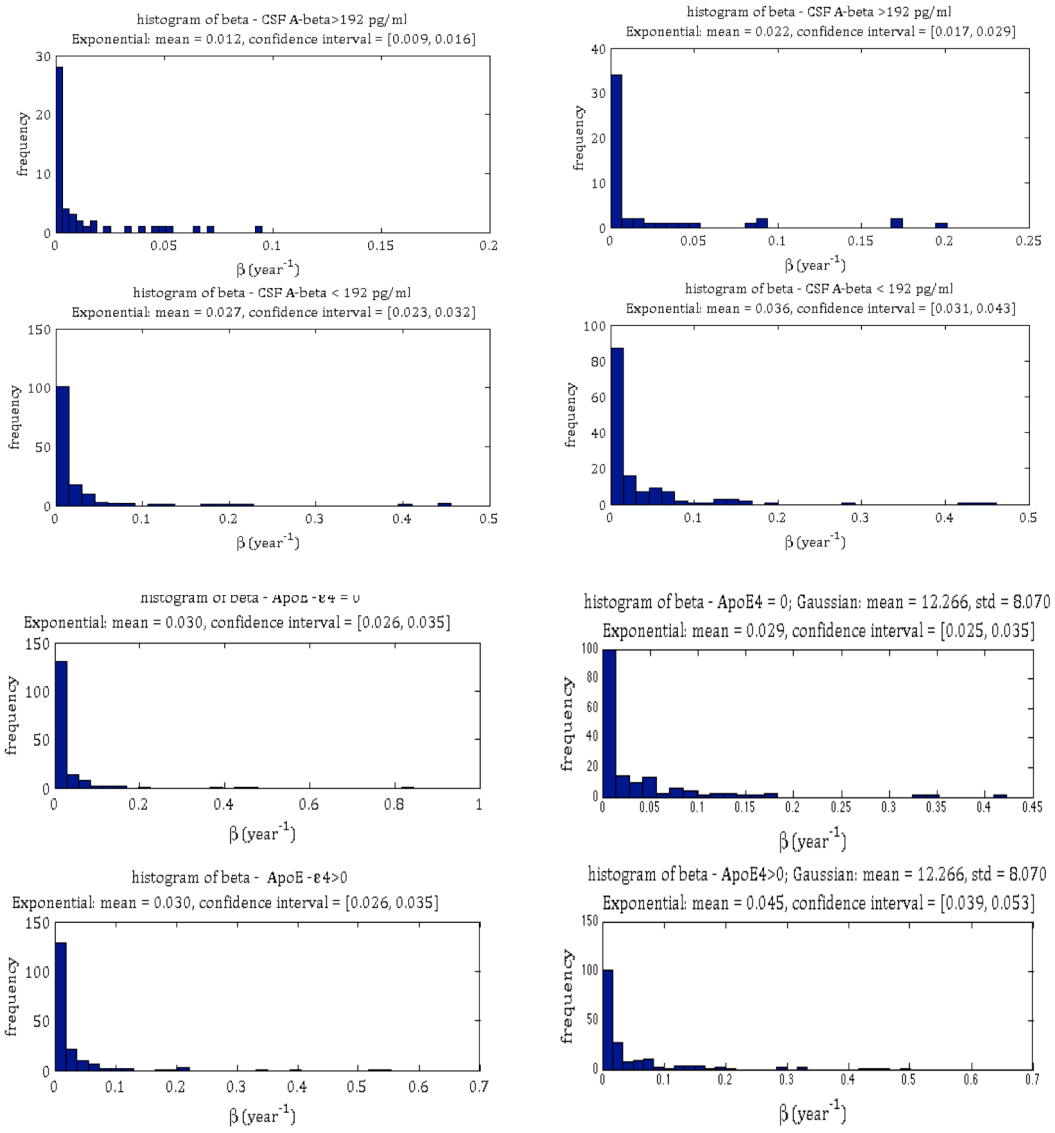
Bottom: The most discordant regions with respect to longitudinal atrophy in the ADNI MCI-converter cohort. Sphere size is proportional to the frequency at which each region appears within the discordant zone. The most common discordant regions are found in the frontal, parietal and occipital lobes. **Related to Figures 1,2.**



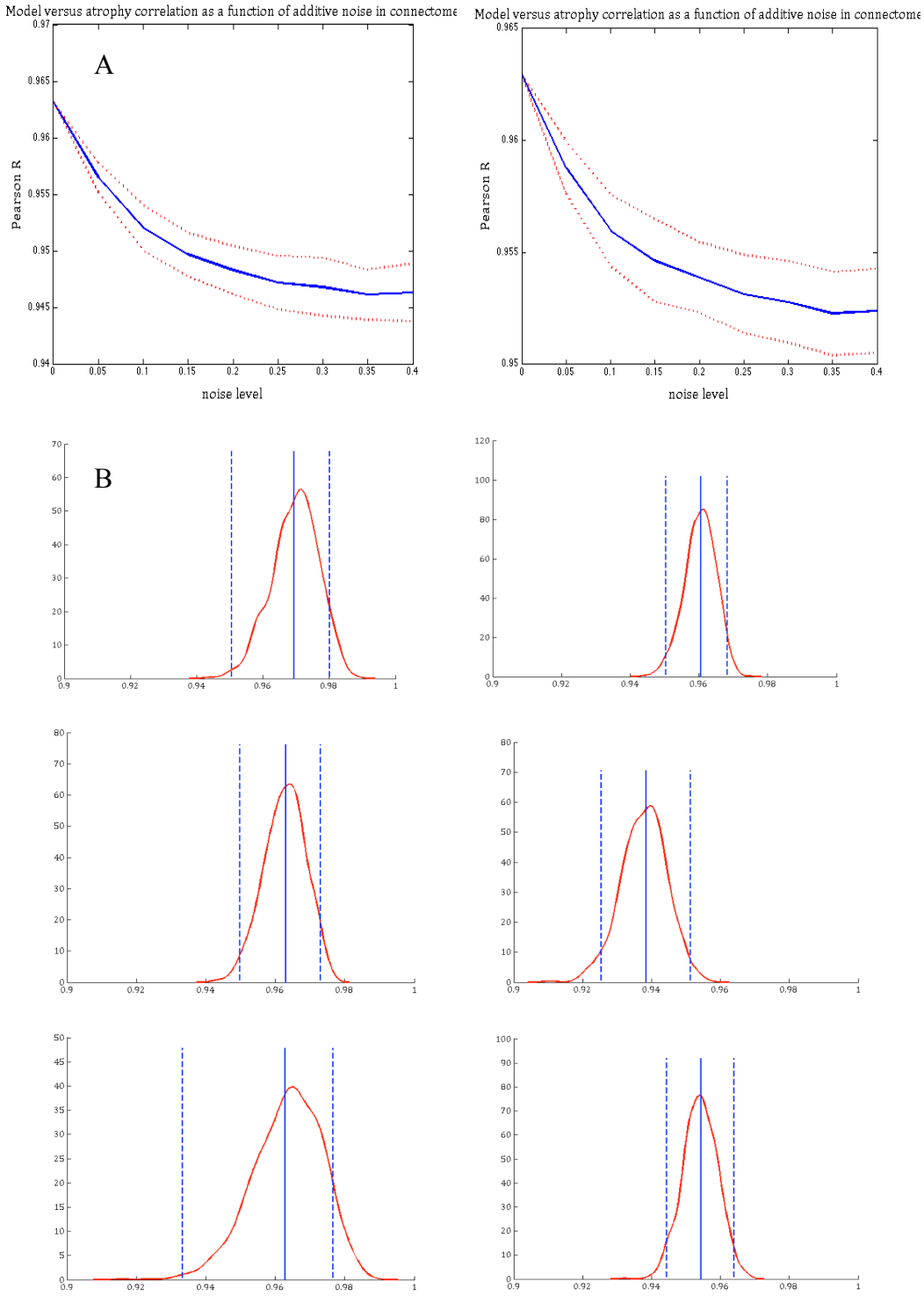
Supplemental Figure 3: Histograms of subject-wise fitting of β (left) and post-onset time at baseline visit (right) on MRI atrophy data of the ADNI cohort. The 3 diagnostic groups are shown, from top to bottom: MCI non-converter, MCI converter, AD. The distribution of β appears to fit an exponential distribution in each case, but with a different mean parameter λ, as summarized in Table 3. Both Gaussian and exponential fitting indicate that β is significantly different between groups. The distribution of the post-onset time does not fit any recognizable distribution (except perhaps the uniform distribution), nor is its mean value different between groups. Similar results were found for FDG-PET-derived hypometabolism data, and they are summarized in Table 3 but not shown here. **Related to Experimental Procedures and Table 3.**



Supplemental Figure 4: Correlation between baseline CSF biomarkers and subject-wise fitted estimates of β (left) and post-onset time at baseline visit (right) of the ADNI cohort. From top to bottom: $A\beta - 42$, tau, p-tau, and $A\beta - 42/$ tau ratio. No significant result can be discerned in any of these plots, nor from the Pearson correlation R or p values indicated alongside. There is a mild negative association between baseline CSF $A\beta - 42$ and β, and a similar mild association between $A\beta - 42$ and post-onset time; both associations are in the expected direction but are not significant. Similar results were found for FDG-PET-derived hypometabolism data, and they are not shown here. **Related to Experimental Procedures and Results.**



Supplemental Figure 5: subject-wise fitted estimates of β after dichotomizing patients using CSF biomarkers and APOE status. **Top**: Histogram of fitted β after dichotomizing baseline CSF $A\beta - 42$ into pathologic (< 192 pg/ml) and non-pathologic (> 192 pg/ml). MRI-derived fitting is shown on the left and FDG-PET-derived fitting on the right. The distribution of β appears to fit an exponential distribution in each case, but with a different mean parameter λ, as summarized in Table 3. Since the 95% confidence interval of λ from the low and high $A\beta - 42$ groups do not overlap, we conclude that β, a marker of the rate of progression, is significantly different between them.
Bottom: Histogram of fitted β after dichotomizing subjects by APOE-ϵ4 allele status: non-carriers and homozygous and heterozygous carriers. The distribution of β appears to fit an exponential distribution in each case, but a significant difference in the fitted exponential mean parameter λ is noted only in the FDG case. **Related to Experimental Procedures, Results and Table 3.**



Supplemental Figure 6: Robustness analysis. **A**: The effect of additive noise in reference connectome on the performance of the predictive model. **Left**: Pearson's R statistic of correlation between measured atrophy pattern at end of study and the ND model evaluated on baseline atrophy pattern from all MCI-converter subjects in the ADNI-I database. Additive noise of increasing variance was added to the reference connectome and the model was repeatedly computed on this noise-corrupted connectome. **Right**: R statistic between measured atrophy

pattern of all AD subjects in ADNI-I database, against ND model prediction using baseline atrophy and noise-corrupted connectome. The mean R over N=100 independent trials is shown at each noise level by the blue curve, and the +/- 1 standard deviation is denoted by dotted red curves. At these levels of additive noise, which are mild to moderate, the performance of the model degrades gracefully and by only a small amount.
B: Bootstrap analysis of variability within patient groups: MCI-nonconverters (top), MCI-converters (middle) and AD (bottom). Non-parametric probability density estimates of the main outcome variable, the R of correlation between predicted and measured end atrophy are shown for each group, along with mean and 95% confidence intervals (vertical lines). Distribution of R with respect to MRI-derived atrophy is shown on the left and from FDG-PET-derived metabolism on the right. In each case the distribution is tight around the mean, with very little bias compared to the mean R statistic reported in Table 2 and Figure 2. The 95% confidence intervals from the above analysis are also reported on Table 3. **Related to Figure 2, Table 2, Results, Experimental Procedures and Supplementary Experimental Procedures.**