Constraining Disease Progression Models Using Subject Specific Connectivity Priors

Anvar Kurmukov[1,2], Yuji Zhao[3], Ayagoz Mussabaeva[1], and Boris Gutman[1,3]

[1] Institute for Information Transmission Problems
[2] National Research University Higher School of Economics
[3] Illinois Institute of Technology

Abstract. We propose a simple yet powerful extension for event-based progression disease model by exploiting the Network Diffusion Hypothesis. Our approach allows incorporating connectivity information derived from diffusion MRI data in the form of an informative prior on event ordering. This simple extension using a definition of transition probability based on network path length leads to improved reproducibility and discriminative power. We report experimental results on a subset of the Alzheimers Disease Neuroimaging Initiative data set (ADNI 2). Though trained solely on cross-sectional data, our model successfully assigns higher progression scores to patients converting to more severe stages of dementia.

Keywords: Connectomes · Disease Progression Model · Alzheimer's disease

1 Introduction

Imaging biomarkers of neurodegeneration have played an increasingly important role in clinical trials and disease stage assessment in recent years [1,2]. At the same time, as the trial design has grown increasingly complex, the very notion of "biomarker" has evolved. Classical notions of biomarker efficacy, such as the power under Normal assumptions [3,4], and classification accuracy [5] have given way to temporally aware models of disease [6,7]. This more recent approach to modeling disease, generally termed Disease Progression Modeling (DPM), assigns a time-dependent disease score (or stage) to each patient as a well as a canonical model of imaging and potentially non-imaging patient data as a function of this score. Unlike traditional classification approaches, this approach rests on the idea that different clinical and imaging features are discriminative at different stages of the disease: each marker has a specific finite time window during which it is affected by the illness. One of the earliest such models used with neuroimaging data is the Event-Based Model (EBM) [8]. Here, the disease score is treated as a discrete variable to be identified with a neurodegenerative "event" in each input phenotype, such as regional gray matter or the presence of misfolded proteins as measured with MRI or PET. The canonical event order is estimated by sampling from a Bayesian posterior formulation, generally using

specific parametric distribution assumptions for healthy and diseased subjects. Variations on EBM include discriminative EBM [9], and simultaneous staging and unsupervised subject subtype identification [10]. Models beyond EBM allow for an explicit continuous-time reparameterization of each subject, effectively modeling both the continuous canonical form of all phenotypes in concert as well as individual "neurological reserve" of each patient. The fully longitudinal DPM's (LDPM), first proposed in [11] as a parametric sigmoidal progression function, were later expanded for spatially dense imaging features with additional spatial priors [12]. The parametric progression form was further relaxed in a Gaussian process formulation in [13].

A noteworthy aspect of the above methods is the lack of informative priors on the order in which specific phenotypes undergo degeneration. In fact, such a prior is readily discernible from available MRI data and has been used elsewhere. Specifically, Raj et al. proposed the Network Diffusion Hypothesis, whereby the neurodegenerative process develops in a highly stereotyped manner, according to the brain's structural connectivity [14]. One recent work on DPM has indeed fused these two ideas [15]. However, even there only a mean "standard" connectivity is used for all subjects. Here for the first time, we propose a subject-specific network prior to constrain DPM. We develop the idea in the context of EBM and apply the model to the ADNI 2 dataset. Initial results indicate a better longitudinal generalization compared to standard EBM, and better predictive ability of the resulting progression score when applied to new subjects.

The remaining paper is structured as follows: Section 2 describes the EBM model and introduces the connectivity prior. We explain our experimental pipeline in Section 3. Section 4 summarizes the results of our experiments. Finally, in Section 5 we discuss some possible enhancements of our approach and conclude.

2 Event-Based Models and the Connectivity Prior

2.1 The Event-Based Model

The idea of modeling disease progression in the form of distinct ordered events goes back to [8]. The authors propose that disease development could be divided into stages. Every stage is defined by an event, i.e. the moment when some biomarker switches its state from normal to abnormal. Specifically, given a set of M biomarkers $X = \{x_1, \dots, x_M\}$ the EBM estimates order π (permutation of indexes $1 \dots M$) in which each biomarker becomes abnormal. The model then takes the following form:

$$p(k|X,\pi) \coloneqq \prod_{j=1}^{k} p(x_j|E_{\pi(j)}) \prod_{j=k+1}^{M} p(x_j|\neg E_{\pi(j)}). \tag{1}$$

This formula defines the probability $p(k|X,\pi)$ of being at progression stage k given a set of biomarkers X and an order π. Here, $p(x_j|E_{\pi(j)})$ and $p(x_j|\neg E_{\pi(j)})$ are likelihoods of a measurement x_j given that event $E_{\pi(j)}$ has or has not

occurred. Given N subjects, the total likelihood of observing the data $\mathbf{X} = \{X_1, \ldots, X_N\}$ is the following:

$$p(\mathbf{X}|\pi) = \prod_{i}^{N} \sum_{k=0}^{M} p(k)p(k|X_i, \pi), \tag{2}$$

where $p(k)$ is the probability of being at stage k. $p(k)$ is typically treated as an uninformative (uniform) prior. Once (2) is maximized and an optimal π is found, one could easily evaluate disease stage k for every patient using the expression in (1). To find optimal π, one needs to maximize the posterior distribution, i.e. find the optimal order π given $\mathbf{X}$:

$$p(\pi|\mathbf{X}) \propto p(\pi)p(\mathbf{X}|\pi), \tag{3}$$

here $p(\pi)$ denotes the prior probability of specific order, which is also typically set to be uniform.

In the present paper, we suggest to use connectivity information obtained from diffusion MRI, to get an informative, non-uniform $p(\pi)$. This allows us to use personal information more directly since both $p(x|E)$ and $p(x|\neg E)$ are estimated on groups of subjects, but $p(\pi)$ could be computed for every subject separately. The original EBM uses longitudinal data but treats each observation as a separate. Here, we fit our model using exclusively cross-sectional data, which allows as to remove the possible effect of overfitting since observations from the same subject are highly correlated. Though our proposition could be easily implemented in any EBM extension, here we decided to use the original EBM, which allows us to isolate the effect of using the connectivity prior.

2.2 Connectome Prior via Path Probability

We now introduce some additional notation. As before we denote by $X = \{x_1, \ldots, x_M\}$ the set of biomarkers, in our case the gray matter thickness for M cortical regions. By G we denote connectivity matrix with exactly M nodes: $\{v_1, \ldots v_M\}$. Each node v_j uniquely corresponds to a biomarker x_j. We use the subscript i to denote different subjects, so X_i means subject i, x_{ij} means j-th biomarker of i-th subject and the same for G_i and v_{ij}. By $p(v_a \to v_b)$ we denote the probability of transitioning from node v_a to node v_b.

We define the probability of a specific path $\pi_{a,b,c} = \{v_a \to v_b \to v_c\}$ as a product of probabilities of every individual step:

$$p(\pi_{a,b,c}) \coloneqq p(v_a \to v_b) \cdot p(v_b \to v_c). \tag{4}$$

next, we define the individual transition probability to be proportional to the shortest path between nodes:

$$p(v_a \to v_b) \propto e^{-\sigma(v_a, v_b)}, \tag{5}$$

where $\sigma(v_a, v_b)$ denotes the shortest path between two nodes, thus transitioning to closer nodes is more probable. We normalize an exponent in such a way that

all probabilities of transitioning from node v_a to all other nodes sum to 1, thus $p(v_a \to v_b) \neq p(v_b \to v_a)$. Specifically, for every connectome, we compute square $(M \times M)$ matrix of shortest paths S:

$$S_{a,b} = \sigma(v_a, v_b), \tag{6}$$

next we apply an exponent as in equation (5), finally, we divide each element in each row by the sum over this row to make individual values sum to 1:

$$p(v_a \to v_b) = \frac{e^{-S_{a,b}}}{\sum_{b=1}^{M} e^{-S_{a,b}}}. \tag{7}$$

The intuition of Network Diffusion Hypothesis is the following: if some region becomes abnormal, it will affect other regions that are structurally closer (in terms of connectivity) faster than regions that are structurally farther. And the structural closeness of two nodes is the length of the shortest path between them.

2.3 Optimizing π

To find the optimal order of events π we need to optimize (3):

$$\begin{aligned}
\pi^* &= \arg\max_{\pi} p(\pi) p(\mathbf{X}|\pi) \\
&= \arg\max_{\pi} p(\pi) \prod_{i}^{N} \sum_{k=0}^{M} p(k|X_i, \pi) = \arg\max_{\pi} \prod_{i}^{N} p(\pi)^{1/N} \sum_{k=0}^{M} p(k|X_i, \pi) \\
&= \arg\max_{\pi} \sum_{i}^{N} \log \left[p(\pi)^{1/N} \sum_{k=0}^{M} p(k|X_i, \pi) \right] \\
&= \arg\max_{\pi} \frac{1}{N} \sum_{i}^{N} \log p(\pi) + \sum_{i}^{N} \log \left[\sum_{k=0}^{M} p(k|X_i, \pi) \right],
\end{aligned} \tag{8}$$

One could compute $p(\pi)$ based on the average connectome ($\hat{(}G) = \frac{1}{N} \sum_i^N G_i$). However, we find that using $\sum_i^N \log p_i(\pi)$ instead of $\sum_i^N \log p(\pi)$ leads to much better results. In other words, every $p_i(\pi)$ is computed from the corresponding individual connectome; for every subject, the prior on the order of events π is different.

As the Bayesian formulation with the connectome prior is identical in form, the optimization of (8) can be done using the MCMC procedure exactly as in [8].

3 Experiments

3.1 Data

Our data consisted of 84 unique subjects from the ADNI 2 dataset: 32 Alzheimer's patients, including 8 women (mean age 76.8 ± 7.6) and 24 men (mean age

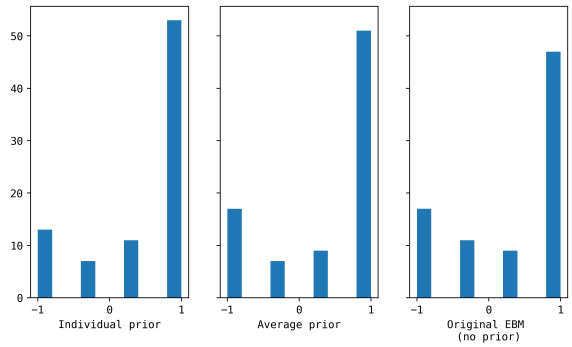

Fig. 1. Agreement results. Agreement between stages and visit order was measured using Kendall tau. Under the assumption that abnormality is irreversible, we measure Kendall tau between the vector of subject visits order and vector of subject disease progression scores. This figure summarizes the distribution of Kendall tau over all subjects.

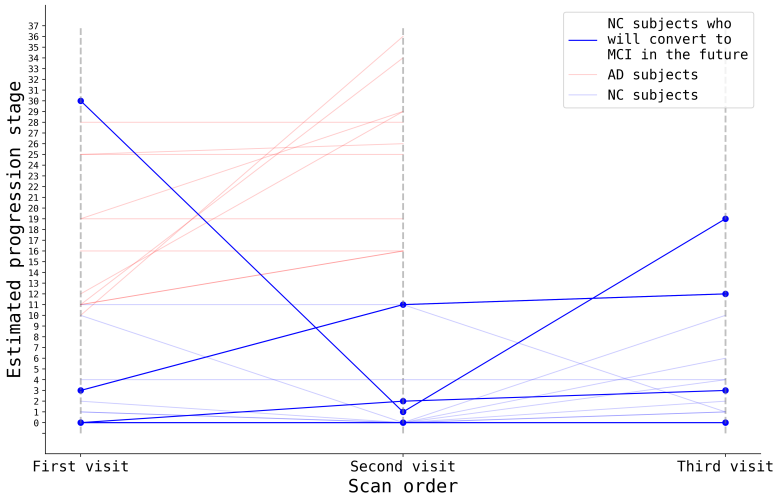

Fig. 2. Selected ADNI participants. Progression scores using order obtained with individual prior. Progression scores were computed using formula (1). Once we obtain optimal order we compute progression score for every subject at each time point.

75.9 ± 7.8) (AD), and 52 cognitively normal controls, including 26 women (mean age 72.0 ± 4.9) and 26 men (mean age 73.6 ± 6.4)(NC). We used anatomical MRI data from 3 visits for each NC subject, and 2 visits for AD subjects. Of

the control subjects, 4 are known future converters to MCI. Regional gray matter thickness was obtained using FreeSurfer 5.3, based on the Dessikan-Killiany atlas. We used only the baseline visit diffusion MRI to construct individual connectomes. Briefly, we used FSLs eddie correction and ANTs SyN for EPI artifact correction to T1 MRI. To extract streamlines, we used constrained spherical deconvolution (CSD) with a probabilistic tractography algorithm, as implemented in Dipy [16]. Finally, weighted connectivity matrices G have 0 on a main diagonal and the weights of edges are inversely proportional to the logarithm of the number of streamlines:

$$G_{a,b} = \begin{cases} \frac{1}{\log(1+w_{a,b})}, \text{if } w_{a,b} > 0 \\ 0, \text{if } w_{a,b} = 0 \end{cases} \tag{9}$$

where $G_{a,b}$ is the edge between nodes a and b; $w_{a,b}$ is the number of streamlines between corresponding regions. The idea behind this specific weighting scheme is the following: firstly, we need the edges to be inversely proportional to number of streamlines (so the weight on edges has notion of distance not similarity); secondly, we do not want to penalize weak connections too much, so we take logarithm; finally, for streamlines with weight 1 we want an edge in a resulting connectome, so we add 1.

3.2 Experimental pipeline

We compare three different versions of the EBM:

1. The original EBM [8].
2. EBM with connectivity prior obtained from average connectome.
3. EBM with individual connectivity priors.

All models were trained on a first time point (cross-sectional data) observation and tested on latter time points (longitudinal data). Recall that evaluation of subject stage using formula (1) does not require $p(\pi)$, but only subject feature vector $X_i = \{x_{i1}, \ldots, x_{iM}\}$ and optimal π^*. Positional uncertainty of the orders was measured over multiple (200) MCMC runs. As additional performance markers, we used ROC AUC in using the inferred disease stage to discriminate AD and NC subjects.

4 Results

The natural way to measure agreement between the predicted stage of a disease based on anatomical features and each subject's actual visit order is with an ordinal correlation. Kendall's τ is the simplest choice, which we use here. We display the distribution of τ over all subjects in figure 1. Mean τ for standard EBM was 0.34, for mean connectome prior 0.41, and for individual connectome prior 0.49. Predicted disease stage for the subjects, including the 4 converters, is displayed in figure 2.

Classification accuracy followed a similar progression, improving with the mean connectome prior, and improving further with the individual prior (figure 3). Mean (standard deviation) ROC AUC over 200 independent MCMC optimizations was 0.816 (0.008) for standard EBM, 0.83 (0.026) for EBM with mean connectome prior, and 0.88 (0.046) for EBM with individual connectome prior.

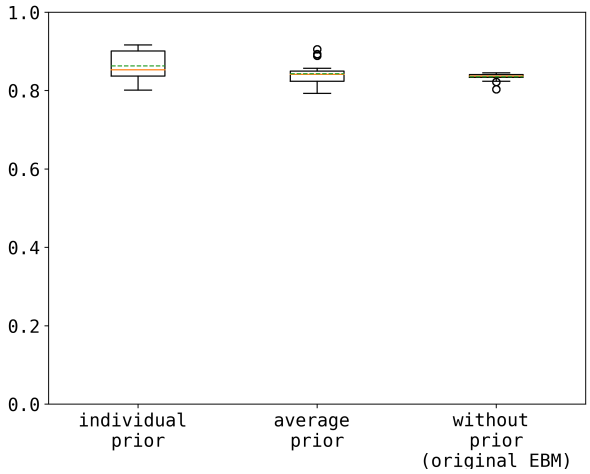

Fig. 3. Binary classification results. Classification was done based on subject stage. Performance was measured in terms of ROC AUC. Classification uncertainty was measured using 200 independent MCMC runs. All models were trained on cross-sectional data and performance was measured on longitudinal data (excluding training observations).

Finally, in figures 4-6 we display the region by order probability matrix as an indicator of the stability of the canonical order computation. Unsurprisingly, using additional individual prior information makes the canonical order significantly less stable. This suggests that the overall EBM model with a single canonical order may not be sufficient to capture true subject variability in the way disease affects different brain regions over time.

Code reproducing all the results is published online. [4]

5 Conclusion

We have presented a direct way to incorporate the Network Diffusion Hypothesis into an established disease progression model. The extension via an informative prior improves several aspects of biomarker performance, including classification

[4] https://github.com/kurmukovai/ebm-connectivity-prior

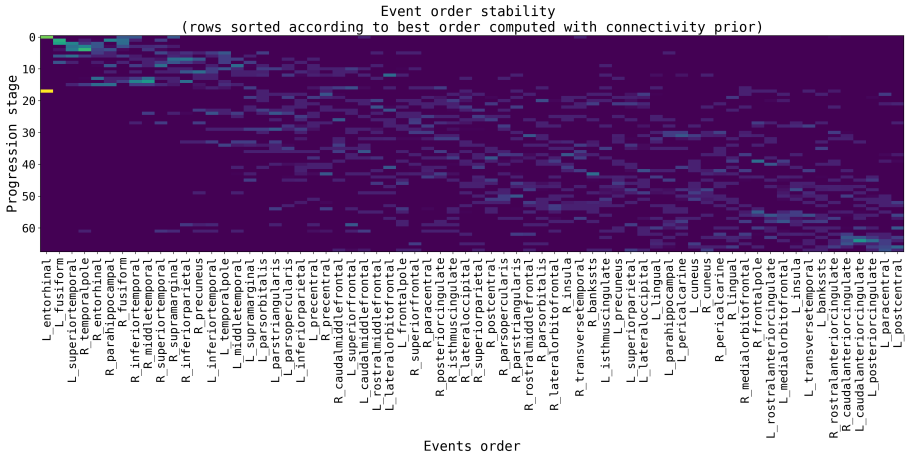

Fig. 4. Event order stability. Uncertainty is measured over multiple MCMC runs, rows are sorted according to best order (using individual prior).

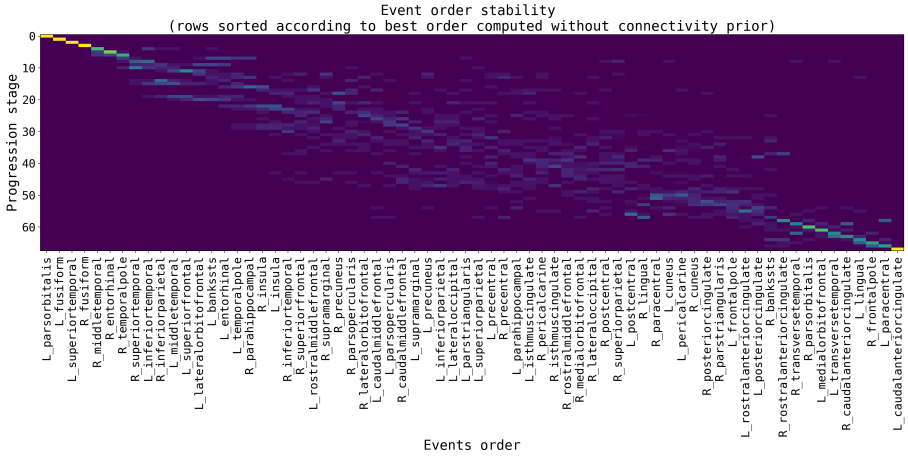

Fig. 5. Event order stability. Uncertainty is measured over multiple MCMC runs, rows are sorted according to best order (no prior).

accuracy, and conversion prediction. Importantly, the work highlight the need to develop more sophisticated models of disease progression that take into account

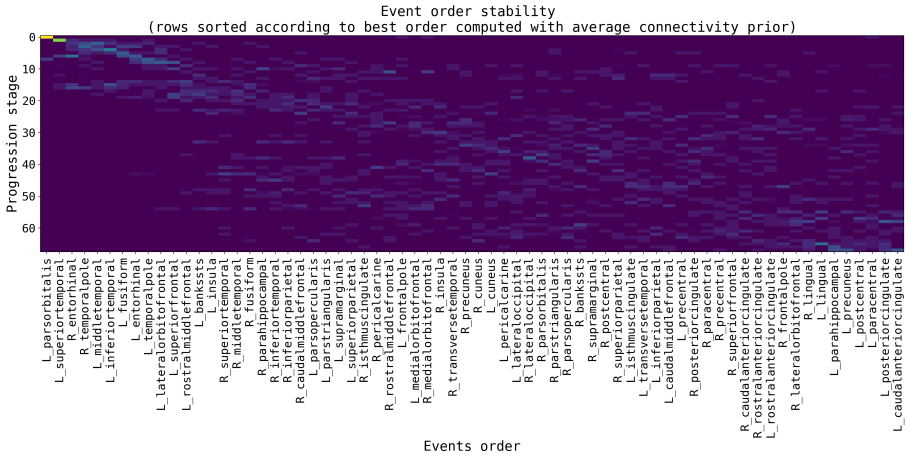

Fig. 6. Event order stability. Uncertainty is measured over multiple MCMC runs, rows are sorted according to best order (using average prior).

individual differences in brain connectivity and the resulting manner in which the disease and specific symptoms are likely to progress. This may include updating the longitudinal DPM models, for example by placing priors on sigmoidal progression parameters, as well as entirely new formulations that replace the notion of a canonical progression with a two-level stochastic process. Implications of the improved progression modeling include better subject stratification for clinical trials, lower drug development costs, and more accurate prediction of future cognitive decline.

6 Acknowledgments

Work by BG was supported by the Alzheimer's Association grant 2018-AARG-592081, Advanced Disconnectome Markers of Alzheimer's Disease. AK and AM were supported by the Russian Science Foundation under grant 17-11-01390.

References

1. Mrquez, F., and Yassa, M.A.: Neuroimaging Biomarkers for Alzheimers Disease, Molecular Neurodegeneration, 2019, 14, (1), pp. 21
2. J. R. Petrella, W. Hao, A. Rao, P. Murali Doraiswamy, and Alzheimers Disease Computational Modeling Initiative, Computational Causal Modeling of the Dynamic Biomarker Cascade in Alzheimers Disease.

3. L. A. Beckett, "Community-based studies of Alzheimer's disease: statistical challenges in design and analysis," Stat Med, vol. 19, pp. 1469-80, Jun 15-30 2000.
4. B. A. Gutman, Y. Wang, I. Yanovsky, X. Hua, A. W. Toga, C. R. Jack Jr, M. W. Weiner, P. M. Thompson, and Alzheimers Disease Neuroimaging Initiative, Empowering imaging biomarkers of Alzheimers disease, Neurobiol. Aging, vol. 36 Suppl 1, pp. S6980, Jan. 2015.
5. C. Gao, H. Sun, T. Wang, M. Tang, N. I. Bohnen, M. L. T. M. Mller, T. Herman, N. Giladi, A. Kalinin, C. Spino, W. Dauer, J. M. Hausdorff, and I. D. Dinov, Model-based and Model-free Machine Learning Techniques for Diagnostic Prediction and Classification of Clinical Outcomes in Parkinsons Disease, Sci. Rep., vol. 8, no. 1, p. 7129, May 2018.
6. C. R. Jack Jr, D. S. Knopman, W. J. Jagust, L. M. Shaw, P. S. Aisen, M. W. Weiner, R. C. Petersen, and J. Q. Trojanowski, Hypothetical model of dynamic biomarkers of the Alzheimers pathological cascade, Lancet Neurol., vol. 9, no. 1, pp. 119128, Jan. 2010.
7. Oxtoby, Neil P., and Daniel C. Alexander. "Imaging plus X: multimodal models of neurodegenerative disease." Current opinion in neurology 30.4 (2017): 371.
8. Fonteijn, Hubert M., et al. "An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease." NeuroImage 60.3 (2012): 1880-1889.
9. Venkatraghavan, Vikram, et al. "Disease progression timeline estimation for Alzheimer's disease using discriminative event based modeling." NeuroImage 186 (2019): 518-532.
10. Young, Alexandra L., et al. "Multiple orderings of events in disease progression." International Conference on Information Processing in Medical Imaging. Springer, Cham, 2015.
11. B. M. Jedynak, A. Lang, B. Liu, E. Katz, Y. Zhang, B. T. Wyman, D. Raunig, C. P. Jedynak, B. Caffo, J. L. Prince, and Alzheimers Disease Neuroimaging Initiative, A computational neurodegenerative disease progression score: method and results with the Alzheimers disease Neuroimaging Initiative cohort, Neuroimage, vol. 63, no. 3, pp. 14781486, Nov. 2012.
12. R. V. Marinescu, A. Eshaghi, M. Lorenzi, A. L. Young, N. P. Oxtoby, S. Garbarino, S. J. Crutch, D. C. Alexander, and Alzheimers Disease Neuroimaging Initiative, DIVE: A spatiotemporal progression model of brain pathology in neurodegenerative disorders, Neuroimage, vol. 192, pp. 166177, May 2019.
13. Lorenzi, M., Filippone, M., Frisoni, G.B., Alexander, D.C., and Ourselin, S.: Probabilistic disease progression modeling to characterize diagnostic uncertainty: Application to staging and prediction in Alzheimer's disease, NeuroImage, 2019, 190, pp. 56-68
14. A. Raj, A. Kuceyeski, and M. Weiner, A network diffusion model of disease progression in dementia, Neuron, vol. 73, no. 6, pp. 12041215, Mar. 2012.
15. Sara Garbarino and Marco Lorenzi, "Modeling and Inference of Spatio-Temporal Protein Dynamics Across Brain Networks," Information Processing in Medical Imaging, pp.57-69 Springer, Cham, 2019
16. E. Garyfallidis, M. Brett, B. Amirbekian, A. Rokem, S. van der Walt, M. Descoteaux, I. Nimmo-Smith, and Dipy Contributors, Dipy, a library for the analysis of diffusion MRI data, Front. Neuroinform., vol. 8, p. 8, Feb. 2014.