

Medical Image Generation and Analysis using Bayesian Generative Models

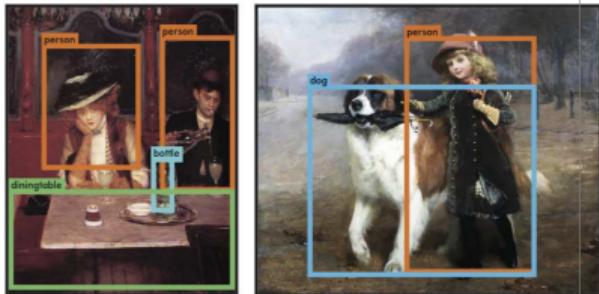
Răzvan V. Marinescu

Massachusetts Institute of Technology



Machine Learning algorithms have achieved impressive milestones

Object detection (YOLO)



Text-to-Image Generation (DALL-E)



prompt: "an armchair in the shape of an avocado"

Image Generation (StyleGAN2)



Text generation (GPT-3)

Title: United Methodists Agree to Historic Split
Subtitle: Those who oppose gay marriage will form their own denomination
Article: After two days of intense debate, the United Methodist Church has agreed to a historic split - one that is expected to end in the creation of a new denomination, one that will be "theologically and socially conservative," according to The Washington Post. The majority of delegates attending the church's annual General Conference in May voted to strengthen a ban on the ordination of LGBTQ clergy and to write new rules that will "discipline" clergy who officiate at same-sex weddings. But those who opposed these measures have a new plan: They say they will form a separate denomination by 2020, calling their church the Christian Methodist denomination.

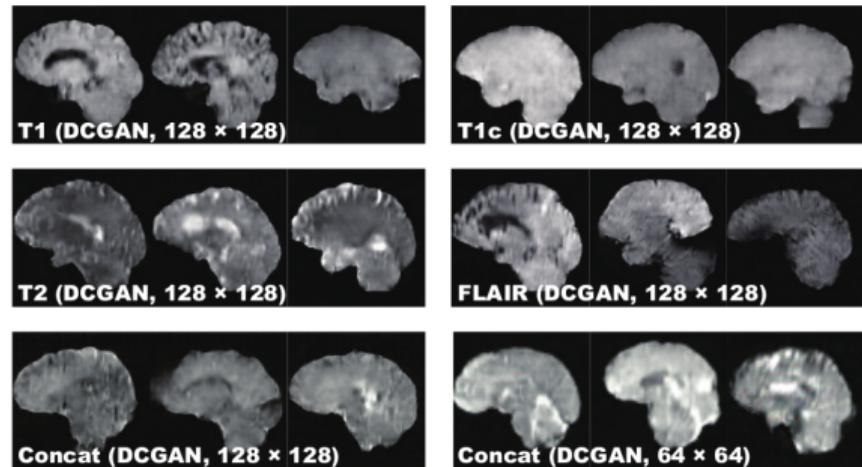
However, such milestones have not been translated to medical applications

► Prediction of clinical variables not always working:

- No algorithm/33, could predict cognitive scores in Alzheimer's
(TADPOLE Challenge, Marinescu 2020)

- Generated images are crude, not high-resolution, mostly 2D

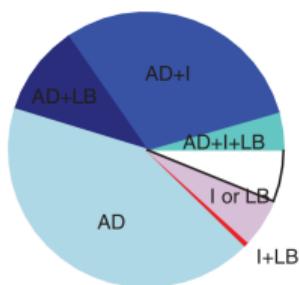
Brain MRI generation (Han, 2018)



Why are Machine Learning models not working on medical applications?

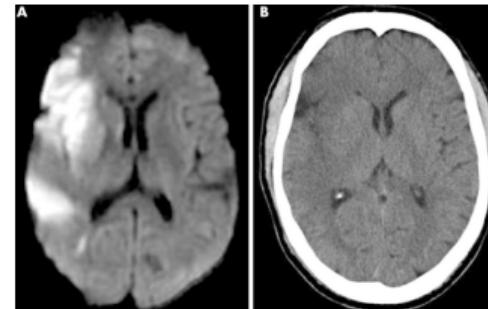
Lack of good labels

- Alzheimer's diagnosis accuracy just 42%



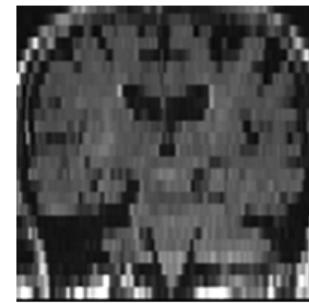
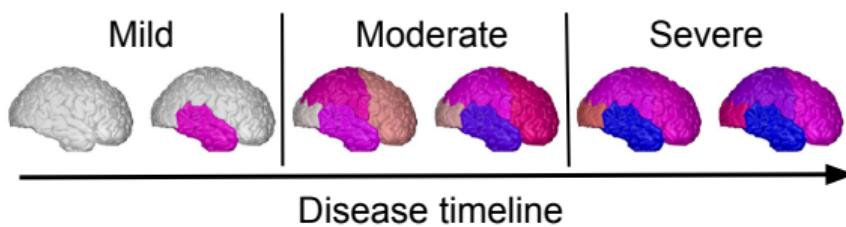
Lack of good input data/signal

- Limited contrast



- Low-resolution

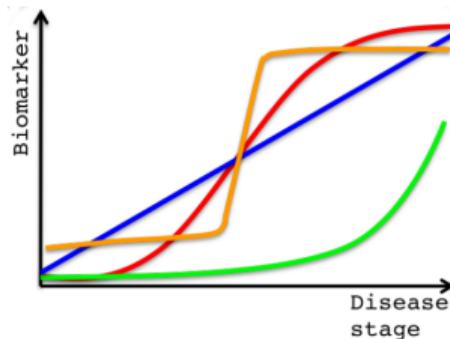
- Labels are categorical instead of continuous



What can we do?

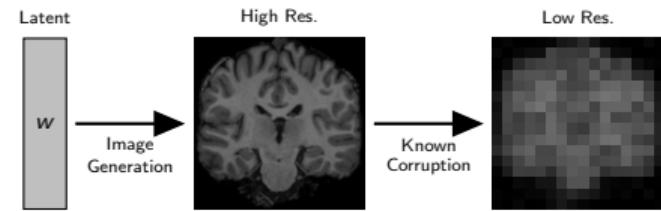
Lack of good labels

Solution: Unsupervised Learning of Continuous Dynamics
= Disease Progression Modelling



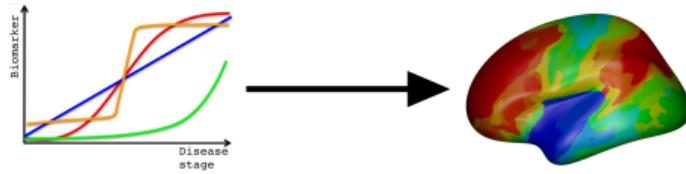
Lack of good input data/signal

Solution: Image Reconstruction using Deep Generative Models

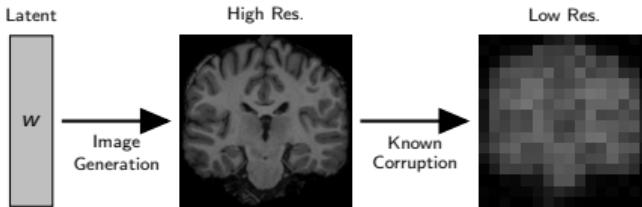


1. Disease progression modelling of Alzheimer's disease

1.1 Towards unsupervised clustering of biomarker trajectories



2. Image Reconstruction using Deep Generative Models



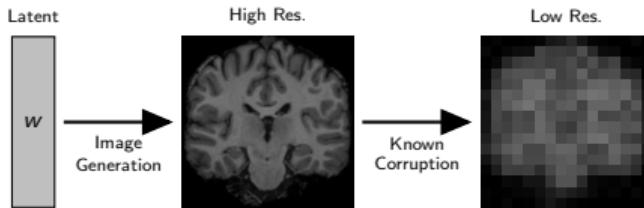
3. Future work

1. Disease progression modelling of Alzheimer's disease

1.1 Towards unsupervised clustering of biomarker trajectories



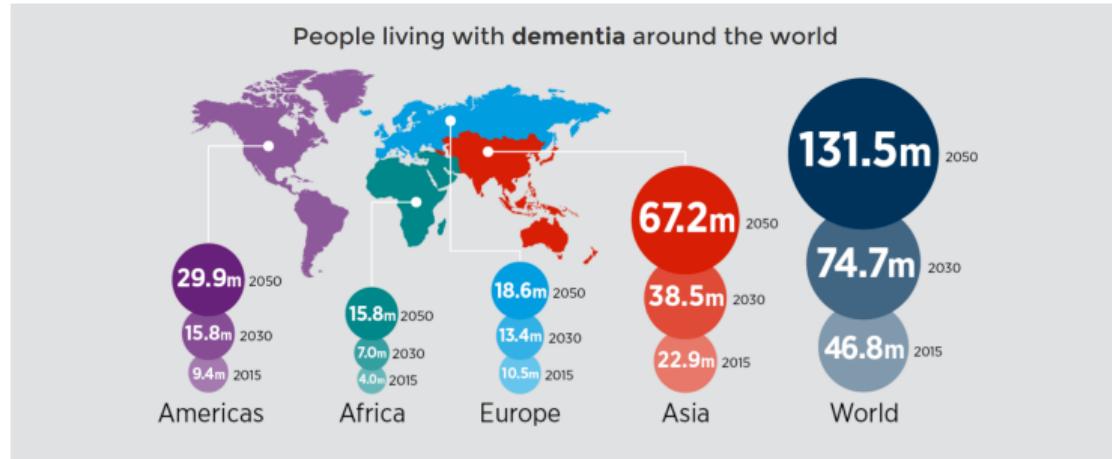
2. Image Reconstruction using Deep Generative Models



3. Future work

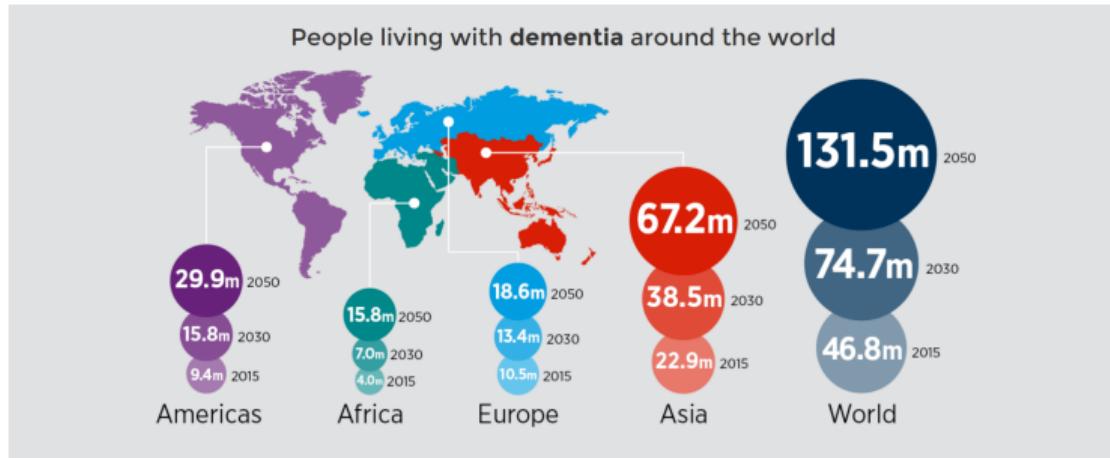
Alzheimer's Disease is a Devastating Disease

- 46 million people affected worldwide



Alzheimer's Disease is a Devastating Disease

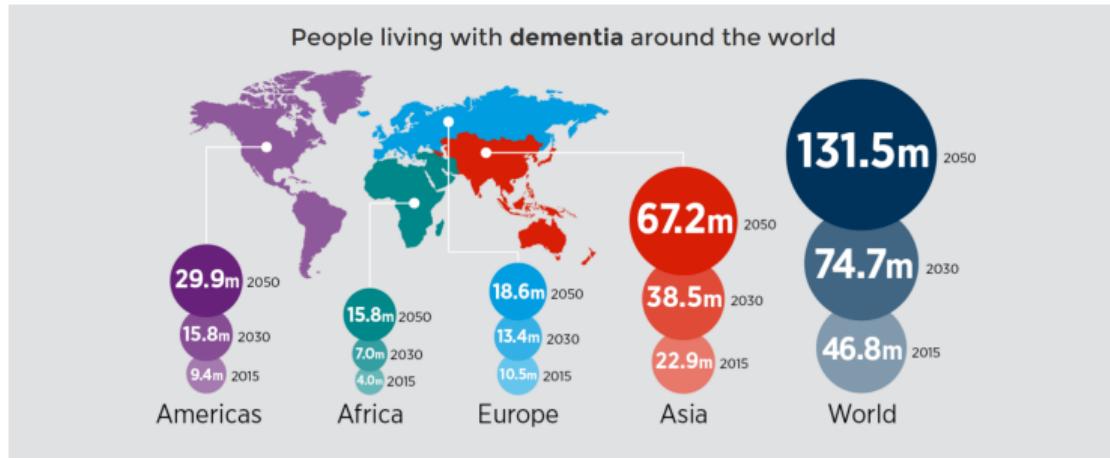
- ▶ 46 million people affected worldwide



- ▶ No treatments available that stop or slow down cognitive decline
- ▶ Q: Why did clinical trials fail? A: Treatments were not administered early enough

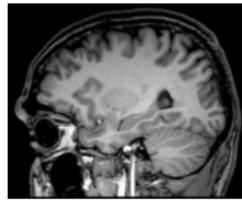
Alzheimer's Disease is a Devastating Disease

- ▶ 46 million people affected worldwide

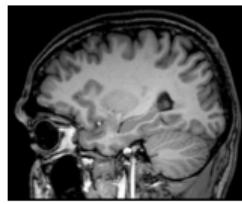


- ▶ No treatments available that stop or slow down cognitive decline
- ▶ Q: Why did clinical trials fail? A: Treatments were not administered early enough
- ▶ Q: How can we then identify subjects **early** in order to administer treatments?
- ▶ A: Disease progression model ...

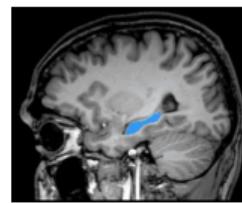
Brain MRI



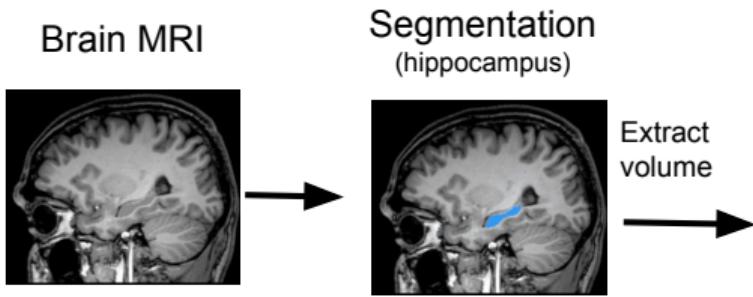
Brain MRI



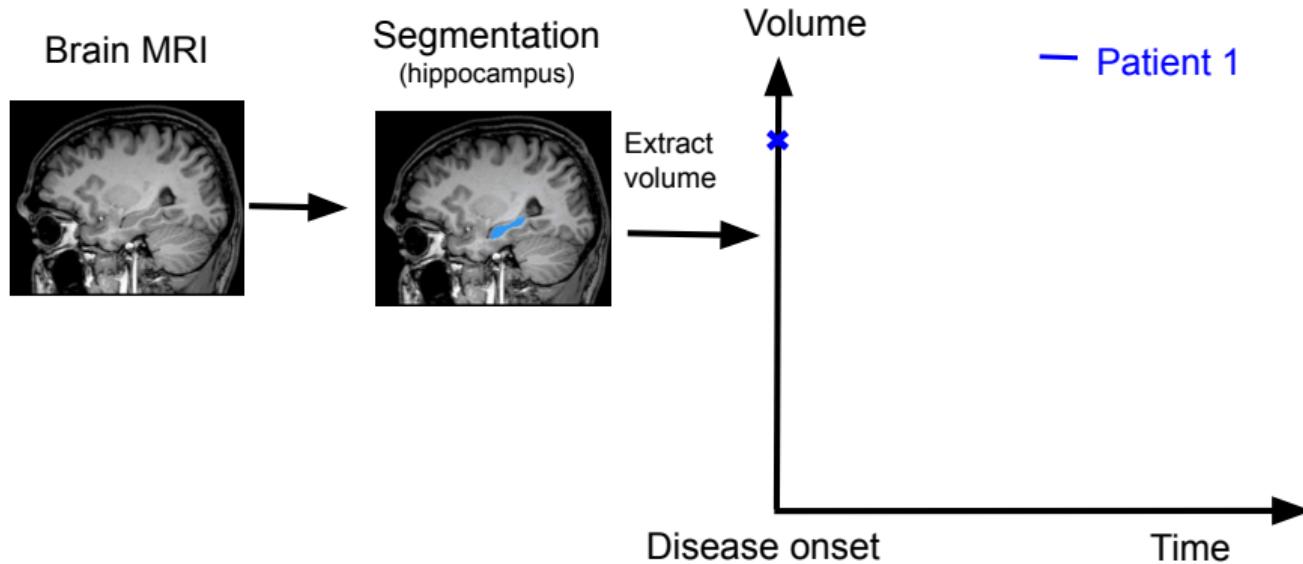
Segmentation
(hippocampus)



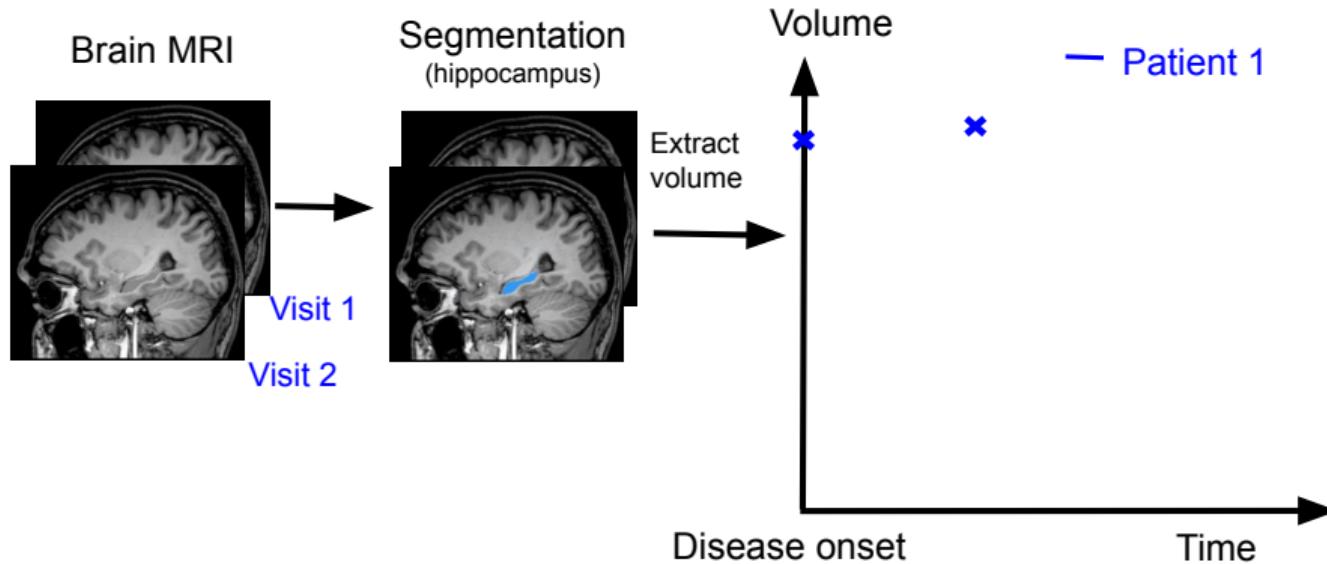
Building a Disease Progression Model



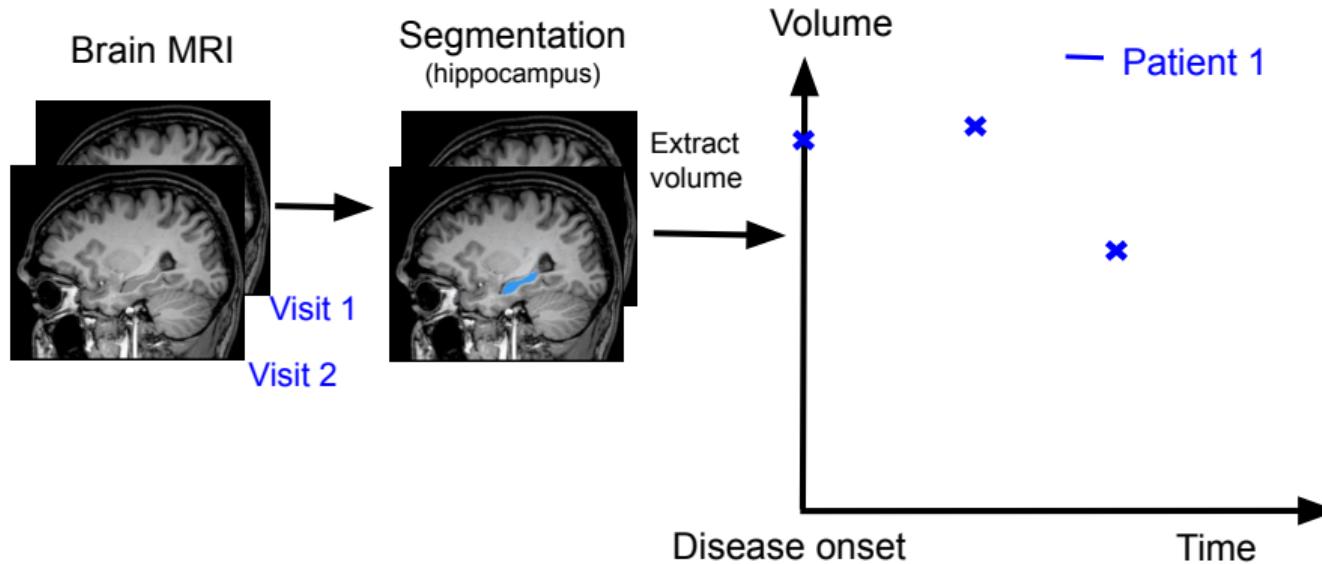
Building a Disease Progression Model



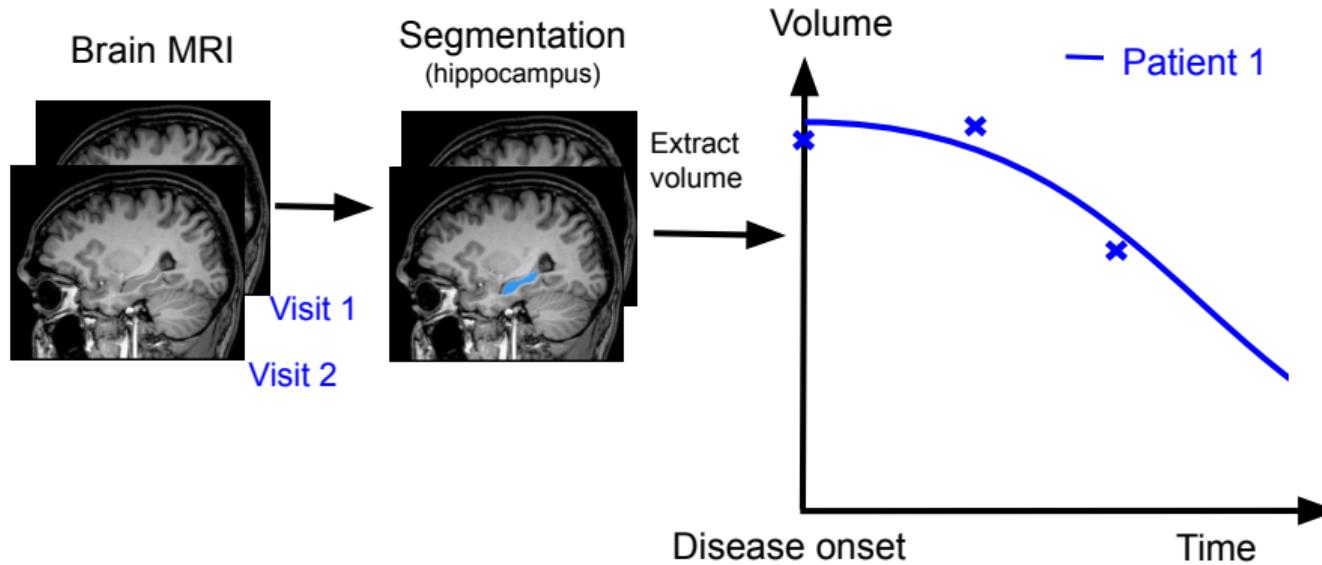
Building a Disease Progression Model



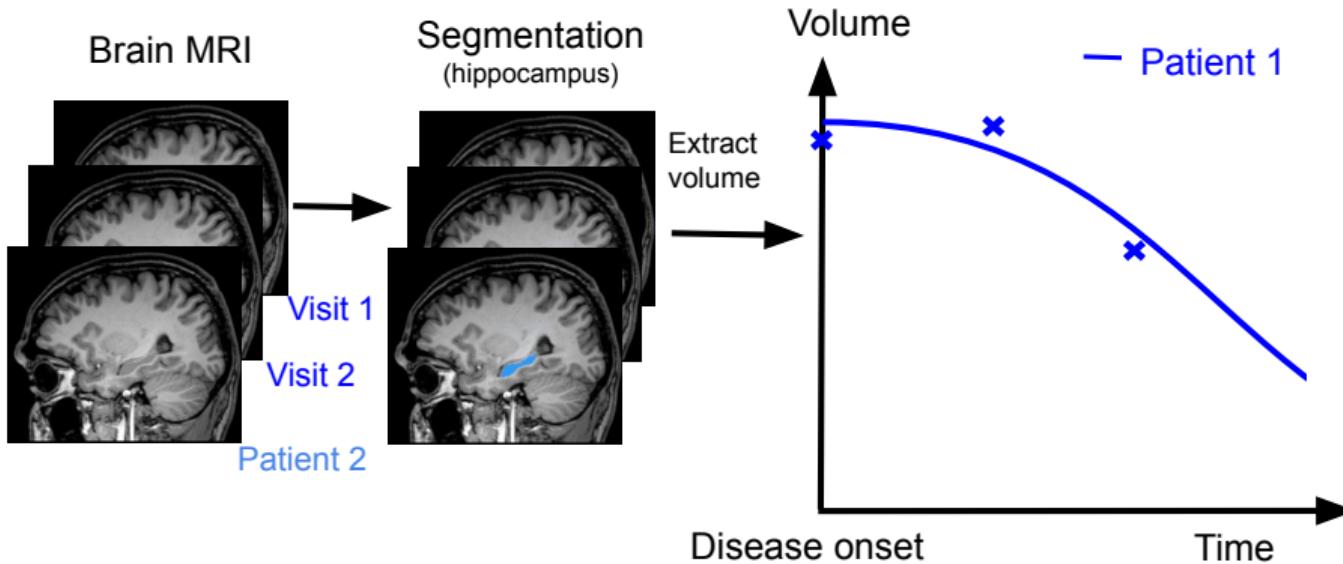
Building a Disease Progression Model



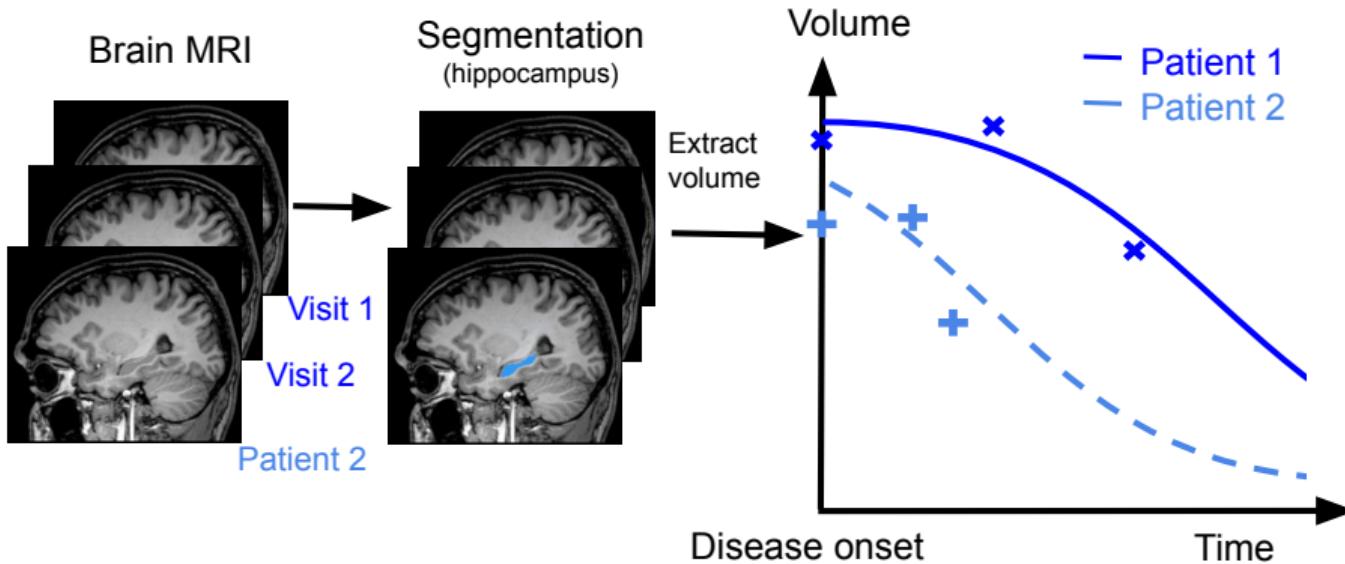
Building a Disease Progression Model



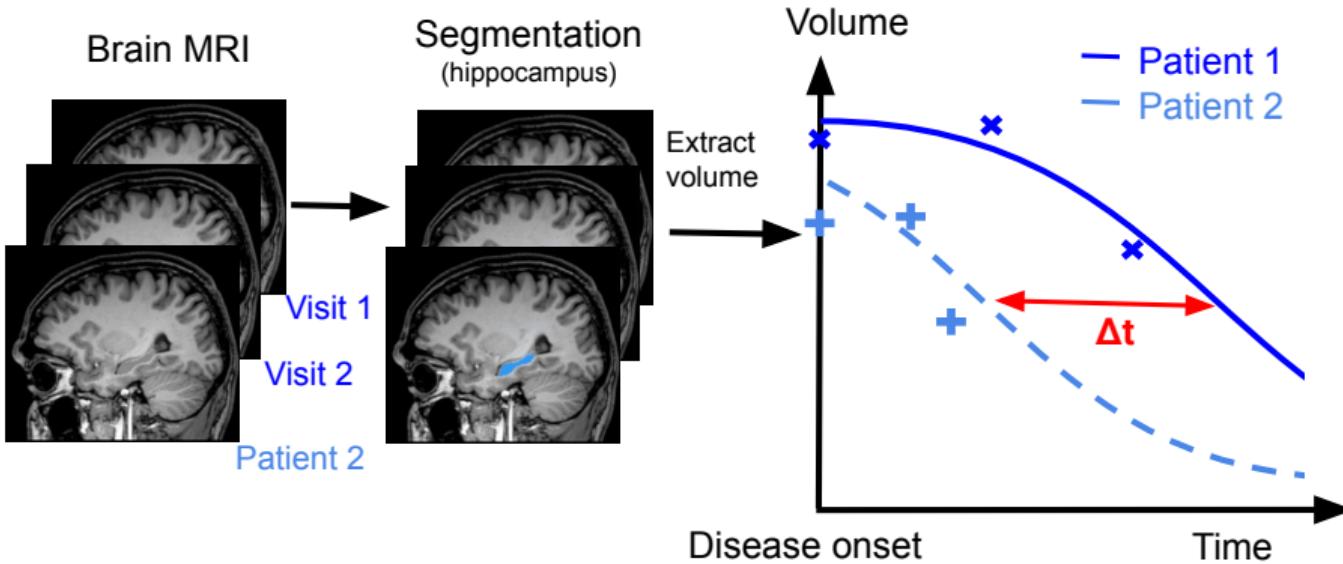
Building a Disease Progression Model



Building a Disease Progression Model

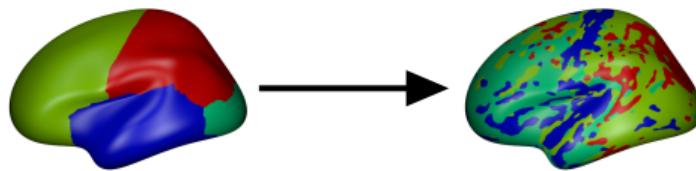


Building a Disease Progression Model

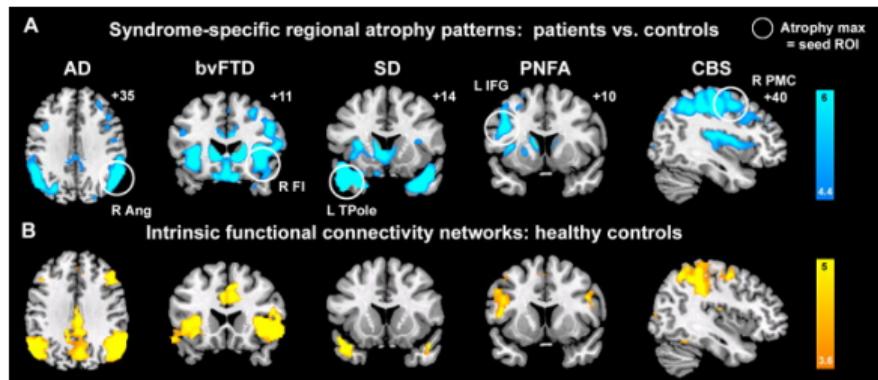


Aim: Build a disease progression model for voxelwise data

- Aim: Move from ROI-based analysis to voxelwise

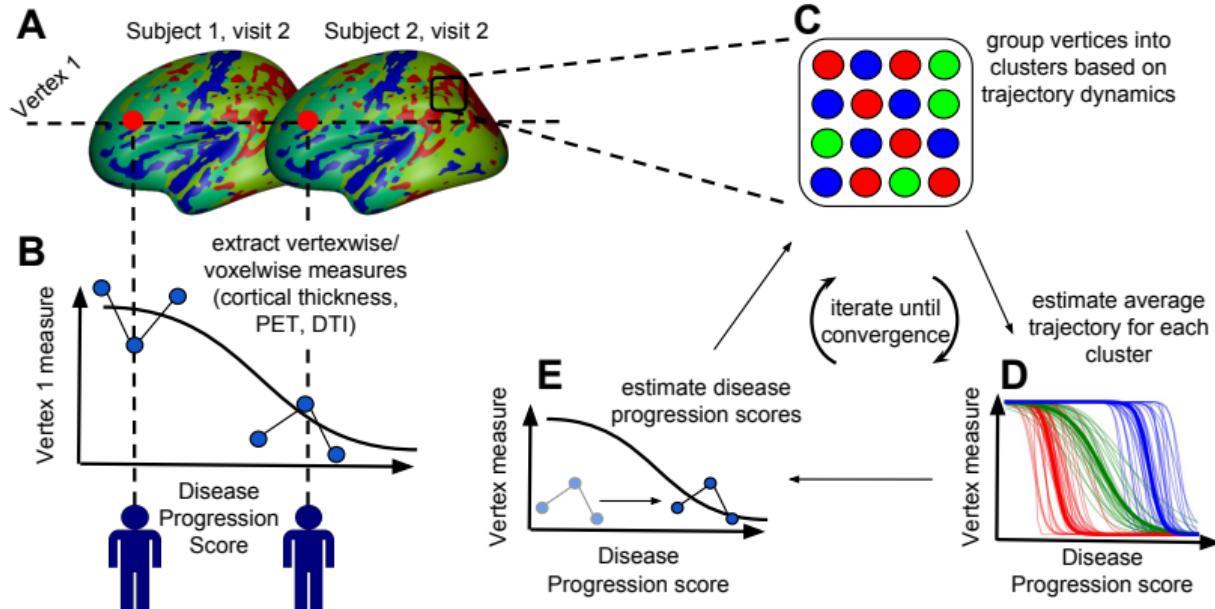


1. Atrophy correlates with functional networks, which are not spatially connected (Seeley et al., Neuron, 2009)
2. Better biomarker prediction and disease staging



(a) Seeley et al., Neuron, 2009

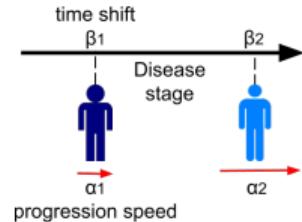
DIVE clusters vertices/voxels with similar trajectories of pathology



Building the model using a generative Bayesian framework

1. Model disease progression score for one subject i at visit j :

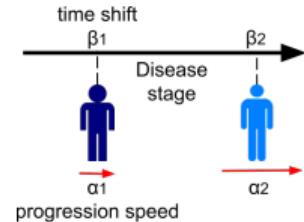
$$s_{ij} = \alpha_i t_{ij} + \beta_i$$



Building the model using a generative Bayesian framework

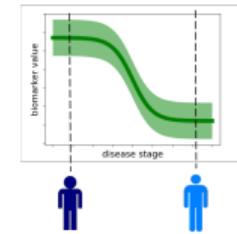
1. Model disease progression score for one subject i at visit j :

$$s_{ij} = \alpha_i t_{ij} + \beta_i$$



2. Model biomarker trajectory of one vertex (point) l on the brain:

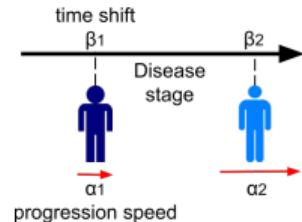
$$p(V_l^{ij} | \alpha_i, \beta_i, \theta_k, \sigma_k) \sim N(f(\alpha_i t_{ij} + \beta_i; \theta_k), \sigma_k)$$



Building the model using a generative Bayesian framework

1. Model disease progression score for one subject i at visit j :

$$s_{ij} = \alpha_i t_{ij} + \beta_i$$

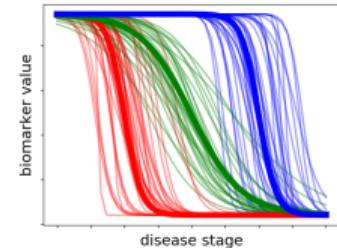
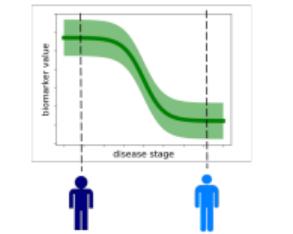


2. Model biomarker trajectory of one vertex (point) I on the brain:

$$p(V_I^{ij} | \alpha_i, \beta_i, \theta_k, \sigma_k) \sim N(f(\alpha_i t_{ij} + \beta_i; \theta_k), \sigma_k)$$

3. Extend to all vertices and subjects:

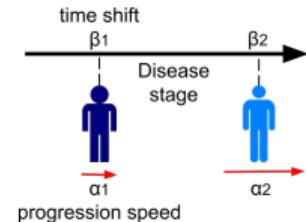
$$p(V, Z | \alpha, \beta, \theta, \sigma) = \prod_I^L \prod_{(i,j) \in I} N(V_I^{ij} | f(\alpha_i t_{ij} + \beta_i; \theta_Z), \sigma_{Z_I})$$



Building the model using a generative Bayesian framework

1. Model disease progression score for one subject i at visit j :

$$s_{ij} = \alpha_i t_{ij} + \beta_i$$



2. Model biomarker trajectory of one vertex (point) I on the brain:

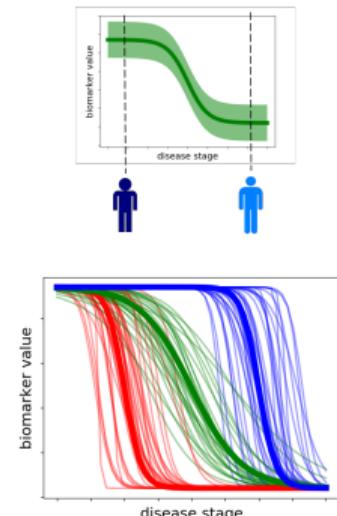
$$p(V_I^{ij} | \alpha_i, \beta_i, \theta_k, \sigma_k) \sim N(f(\alpha_i t_{ij} + \beta_i; \theta_k), \sigma_k)$$

3. Extend to all vertices and subjects:

$$p(V, Z | \alpha, \beta, \theta, \sigma) = \prod_I^L \prod_{(i,j) \in I} N(V_I^{ij} | f(\alpha_i t_{ij} + \beta_i; \theta_{Z_I}), \sigma_{Z_I})$$

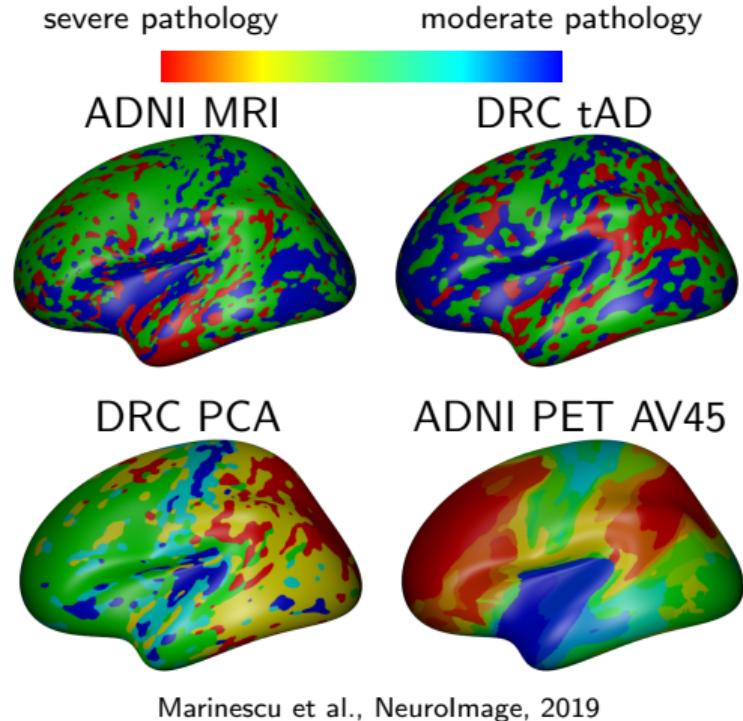
4. Marginalise over the hidden variables Z_I (cluster assignments):

$$p(V | \alpha, \beta, \theta, \sigma) = \prod_{I=1}^L \sum_{k=1}^K p(Z_I = k) \prod_{(i,j) \in I} N(V_I^{ij} | f(\alpha_i t_{ij} + \beta_i ; \theta_k), \sigma_k)$$



DIVE Finds Plausible Atrophy Patterns on Four Datasets

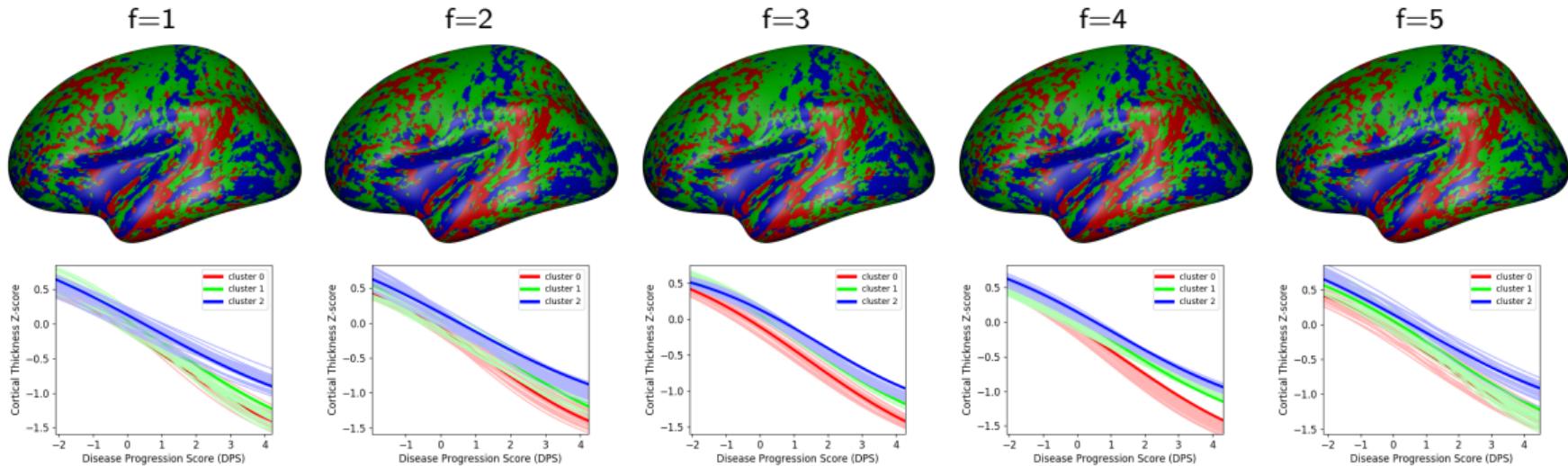
- ▶ Similar patterns of tAD atrophy in independent datasets: ADNI and UCL DRC
- ▶ Distinct patterns of atrophy in different diseases (tAD and PCA) and modalities (MRI vs PET)



Validation - Model Robustly Estimates Atrophy Patterns

Method: Tested the consistency of the spatial clustering in ADNI using 10-fold CV

Results: Good agreement in terms of spatial distribution (dice score 0.89)

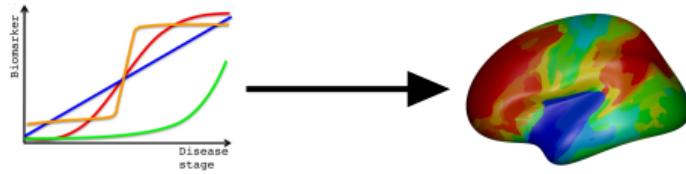


Marinescu et al., Neuroimage, 2019

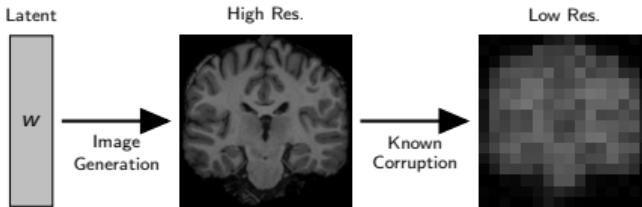
- ▶ We modelled the continuous progression of Alzheimer's disease and related dementias
- ▶ Used clustering, in EM framework, to model spatial correlation
- ▶ However, such models require good quality data, to extract biomarkers
- ▶ How can we do such modelling for scans with limited resolution and contrast?

1. Disease progression modelling of Alzheimer's disease

1.1 Towards unsupervised clustering of biomarker trajectories



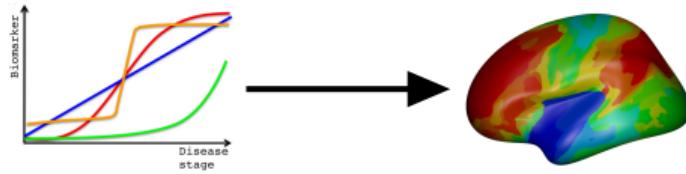
2. Image Reconstruction using Deep Generative Models



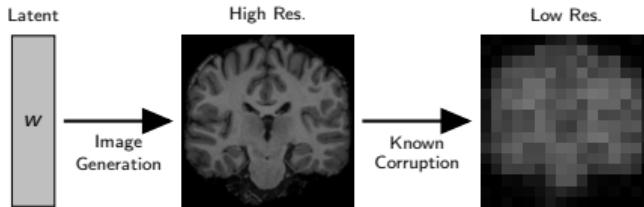
3. Future work

1. Disease progression modelling of Alzheimer's disease

1.1 Towards unsupervised clustering of biomarker trajectories



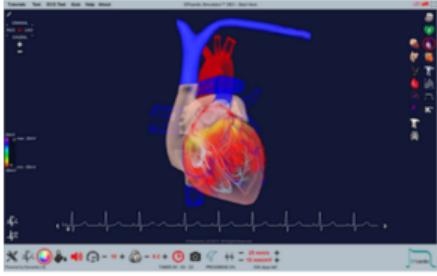
2. Image Reconstruction using Deep Generative Models



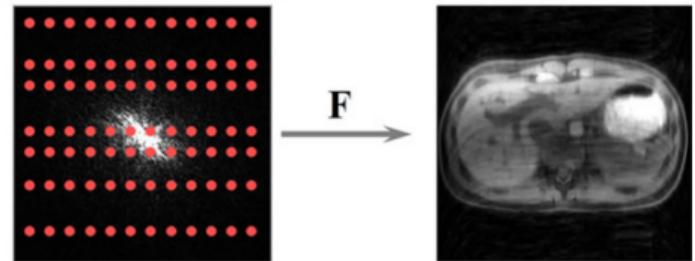
3. Future work

Future work

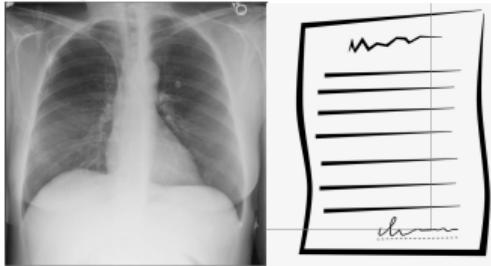
Biological simulators



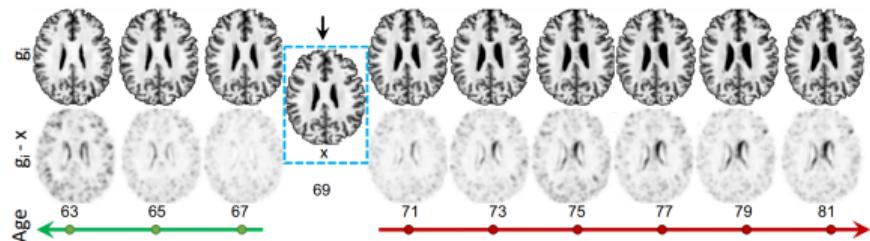
Better and faster reconstruction of medical images
Undersampled k-space Acquired Image



Multimodal modelling
images + text + structural data



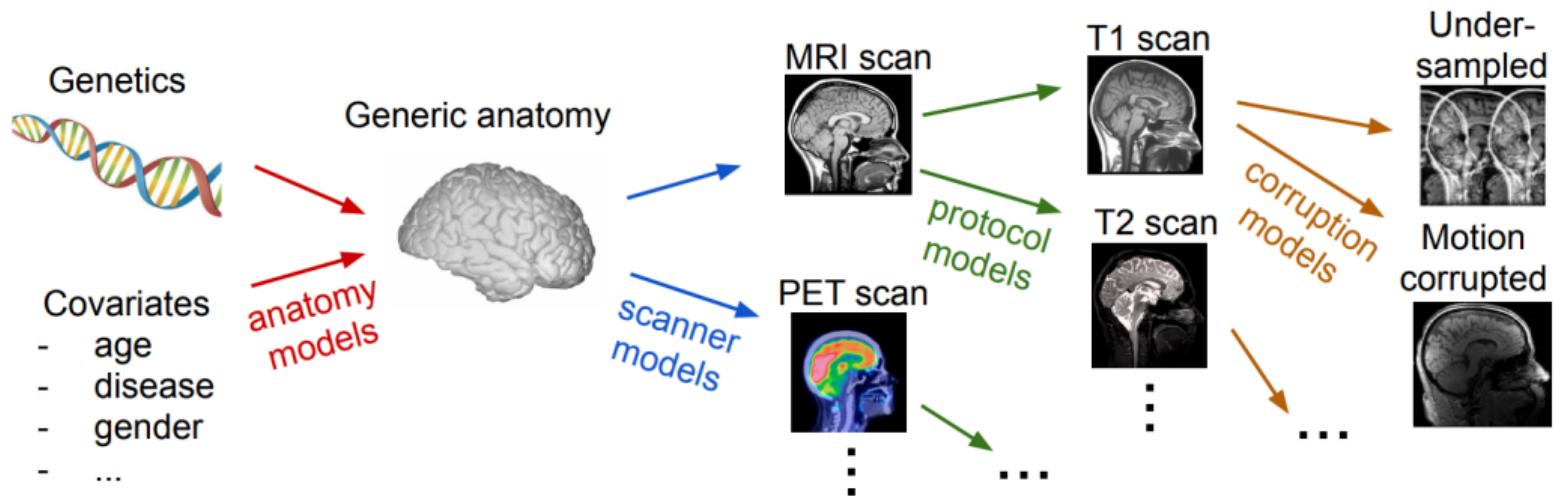
Disease Progression Modelling



Future work: Brain tissue and anatomy simulator

Simulator for brain anatomy from genetics:

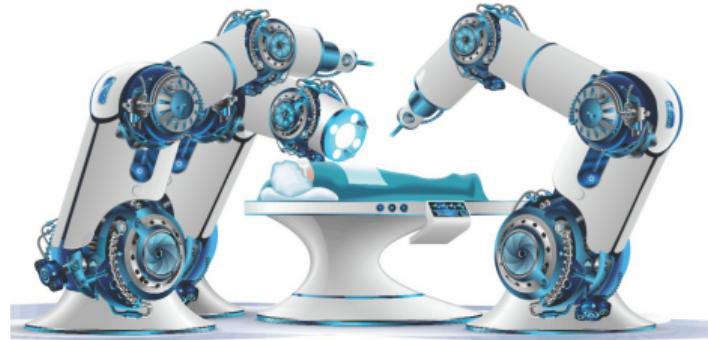
- ▶ Using deep generative models
- ▶ Accounting for distributions shifts
- ▶ Following causal principles



Early diagnosis and prognosis



Robotic Surgery



AI augmenting humans



Drug development



Step 5: Modelling Spatial Correlation using Markov Random Fields

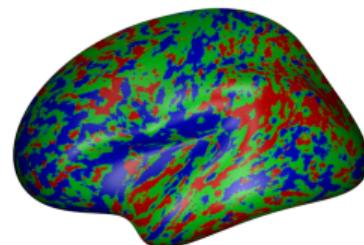
Motivation

- ▶ measurements from neighbouring vertices are inherently correlated
- ▶ can "fill-in holes", eliminate noisy cluster assignments due to noise

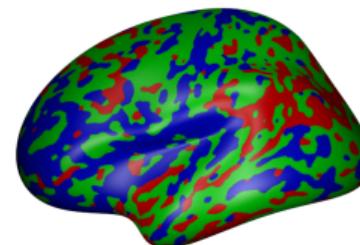
$$p(V, Z | \alpha, \beta, \theta, \sigma) = \prod_{l=1}^L \prod_{(i,j) \in I_l} N(V_l^{ij} | f(\alpha_i t_{ij} + \beta_i | \theta_{Z_l}), \sigma_{Z_l}) \prod_{l_1 \sim l_2} \Psi(Z_{l_1}, Z_{l_2})$$

where

- ▶ $\Psi(Z_{l_1} = k_1, Z_{l_2} = k_2) = \begin{cases} \exp(\lambda) & \text{if } k_1 = k_2 \\ \exp(-\lambda) & \text{otherwise} \end{cases}$
- ▶ λ - MRF parameter



(a) Without MRF



(b) With MRF, $\alpha = 5$.

Model Fitting with Expectation-Maximisation (EM)

► E-step:

- ▶ Estimate vertex assignment to clusters $z_{lk}^{(u)} = \zeta_{lk}(\lambda^{(u)})$:

$$\lambda^{(u)} = \arg \max_{\lambda} \sum_{l=1}^L \sum_{k=1}^K \zeta_{lk}(\lambda) \left[D_{lk} + \lambda \sum_{l_2 \in N_l} \zeta_{l_2 k}(\lambda) - \lambda^2 \sum_{l_2 \in N_l} (1 - \zeta_{l_2 k}(\lambda)) \right]$$
$$\zeta_{lk}(\lambda) \approx \exp \left(D_{lk} + \sum_{l_2 \in N_l} \log \left[\exp(-\lambda^2) + z_{l_2 k}^{(u-1)} (\exp(\lambda) - \exp(-\lambda^2)) \right] \right)$$

where:

$$D_{lk} = -\frac{1}{2} \log (2\pi (\sigma_k^{(u)})^2 |I|) - \frac{1}{2(\sigma_k^{(u)})^2} \sum_{i,j \in I} (V_i^{ij} - f(\alpha_i^{(u)} t_{ij} + \beta_i^{(u)} | \theta_k^{(u)}))^2$$

► M-step:

- ▶ Update trajectories:

$$\theta_k = \arg \min_{\theta_k} \left[\sum_{l=1}^L z_{lk} \sum_{(i,j) \in I} (V_i^{ij} - f(\alpha_i t_{ij} + \beta_i | \theta_k))^2 \right] - \log p(\theta_k) \quad (1)$$

- ▶ Update subject progression scores:

$$\alpha_i, \beta_i = \arg \min_{\alpha_i, \beta_i} \left[\sum_{l=1}^L \sum_{k=1}^K z_{lk} \frac{1}{2\sigma_k^2} \sum_{j \in I_l} (V_i^{ij} - f(\alpha_i t_{ij} + \beta_i | \theta_k))^2 \right] - \log p(\alpha_i, \beta_i) \quad (2)$$