Attentive State-Space Modeling of Disease Progression

Ahmed M. Alaa
Mihaela van der Schaar

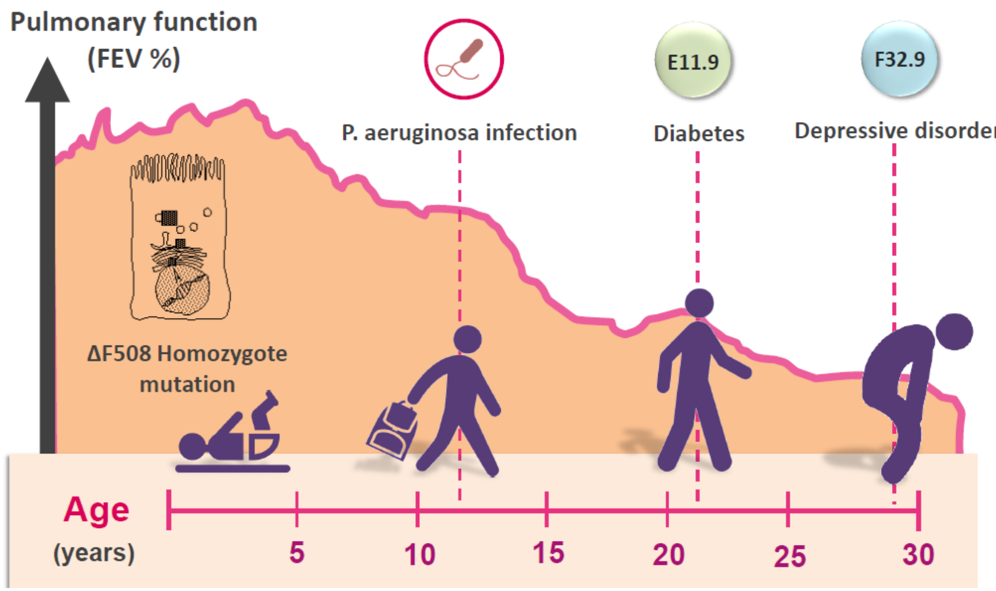
NeurIPS 2019

Disease Trajectory Modeling

- Cross-sectional modeling: Simple prognostic questions (how long will a patient *survive*?) are not enough!

- Longitudinal modeling: addresses a more adequate question: ***what will happen to the patient in future?***

Inform patient carers about the patient's future needs and interactions with health services



Two Central Goals of Longitudinal Models

- Goal A: **Predicting individual-level disease trajectories**
  - Early diagnosis
  - What are the risks of mortality, comorbidities, complications, infections in the future?
- Goal B: **Understanding disease progression mechanisms.**
  - Underlying latent structure of disease evolution
  - Causal pathways and comorbidity networks
  - Patients' subgroup analysis
  - Refined phenotypes and ontologies

Accuracy-interpretability Trade-off

- **Goals A and B** have very **different** (and **conflicting**) modeling requirements!

Goal A

Complex black-box models optimized for a given prediction task

Example: Recurrent neural networks (RNNs)

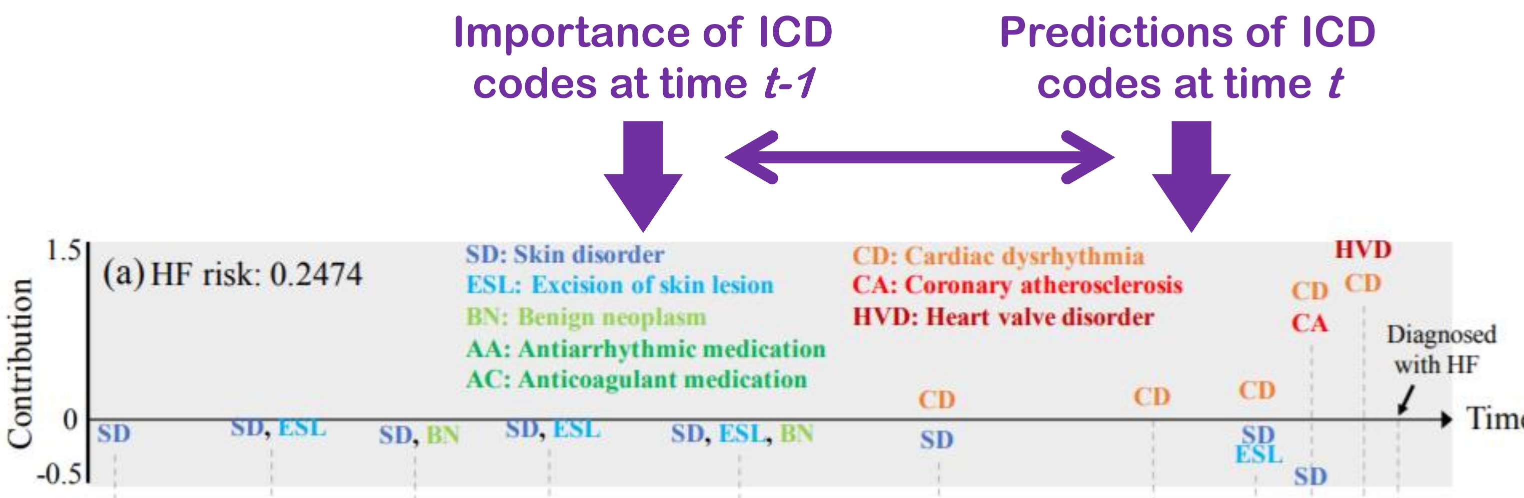
Goal B

Models with structures that are understandable to humans and mimic an underlying physical process

Example: Hidden Markov model (HMM)

Different Forms of Interpretability

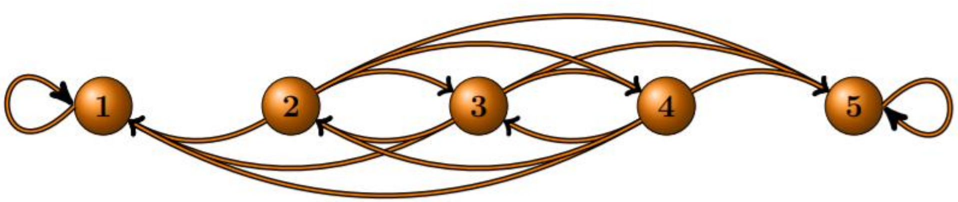
- **Form A: Instance-wise variable importance**
  - **<u>For every patient</u>:** identify important variables for future predictions at every time step based on the patient's history.
  - **Useful for transparency and actionability but not sufficient (no knowledge extraction).**
  - **Example RETAIN [E. Choi, 2017].**



Different Forms of Interpretability

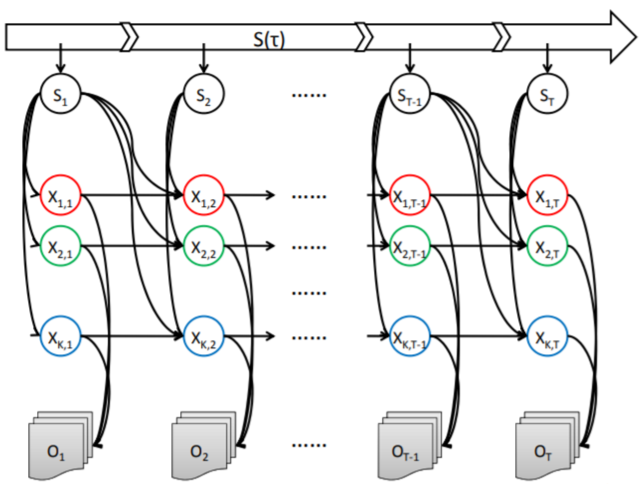
- **Form B: Discovering latent structures in disease progression**
  - **Learn a hidden representation of the disease:** identify disease states and manifestations of states.

Population-level representation of disease states



Generative models

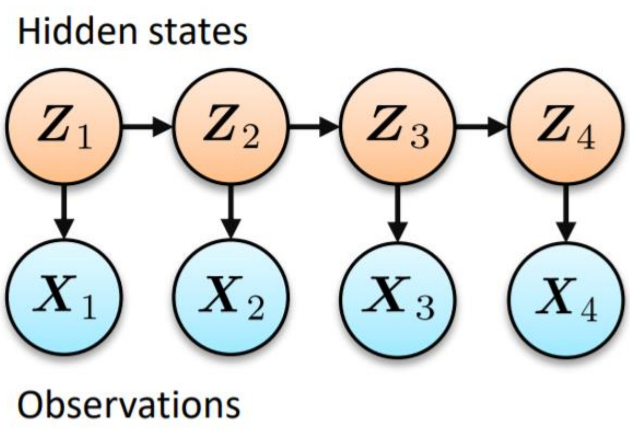
Segmentation and annotation of individual-level trajectories



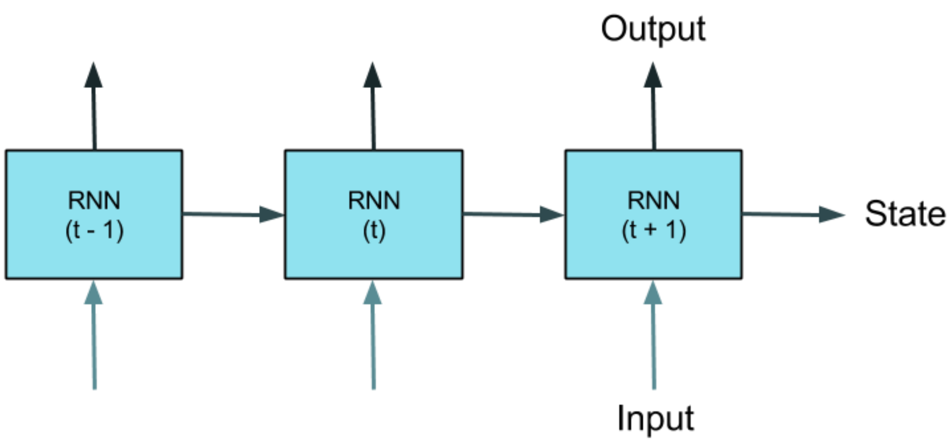
Attentive State-space Modeling

- We want Form A and Form B + accurate predictions
- Our contribution: attentive state-space models
- Main idea:

Maintain probabilistic structure of an HMM

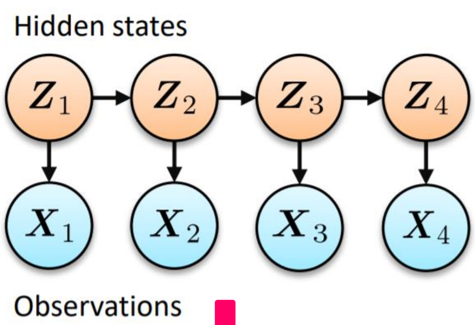


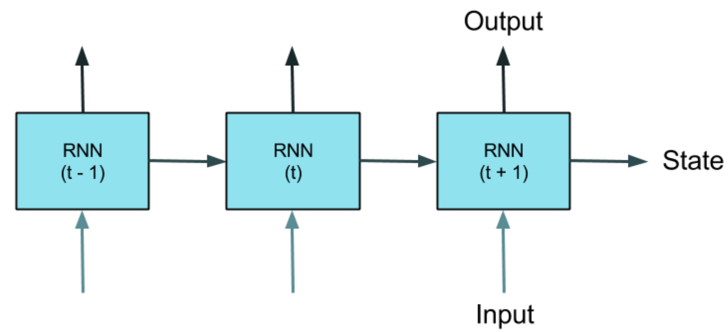
But use an RNN to model the state dynamics



Attentive State-space Modeling

- Attentive state-space models





$$p_\theta(\vec{x}_T, \vec{z}_T) = \prod_{t=1}^{T} \underbrace{p_\theta(x_t \mid z_t)}_{\textbf{Emission}} \underbrace{p_\theta(z_t \mid \vec{x}_{t-1}, \vec{z}_{t-1})}_{\textbf{Transition}}$$

Contextual Attention

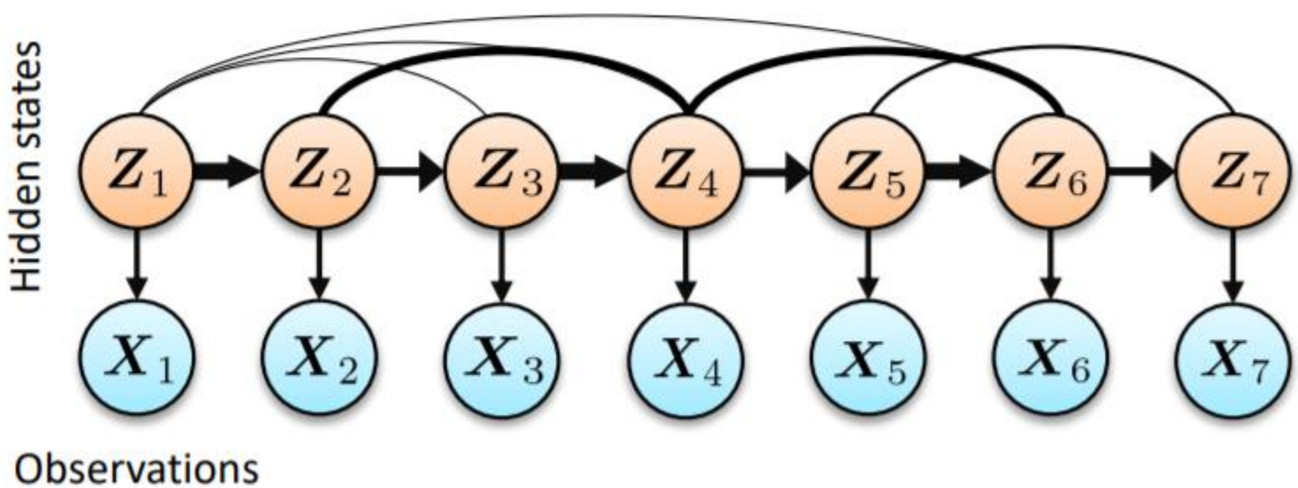
- **Use attention mechanism to keep state dynamics interpretable**
- **Observable variables:** $(\boldsymbol{X}_1, \dots, \boldsymbol{X}_t)$
- **Disease states:** $(\boldsymbol{Z}_1, \dots, \boldsymbol{Z}_t)$

$$\overbrace{(\boldsymbol{\alpha}_1^t, \dots, \boldsymbol{\alpha}_t^t)}^{\text{Attention weights}} = \mathbf{RNN}\overbrace{(\boldsymbol{X}_1, \dots, \boldsymbol{X}_t)}^{\text{Patient context}}$$

$$P(\boldsymbol{Z}_m = z \,|\, \mathcal{F}_{t_{m-1}}) = P(\boldsymbol{Z}_m = z \,|\, \underbrace{\boxed{\{(\boldsymbol{Z}_k = z_k, \boldsymbol{\alpha}_k^m, \Delta_k)\}_{k=1}^{m-1}}}_{\textbf{Sufficient statistics}}) = \sum_{k=1}^{m-1} \underbrace{\boxed{\boldsymbol{\alpha}_k^m}}_{\textbf{Attention weights}} \overbrace{(e^{\Delta_k \boldsymbol{\Lambda}})}^{\text{Baseline Markov chain}}{}_{z, z_k}$$

Contextual Attention

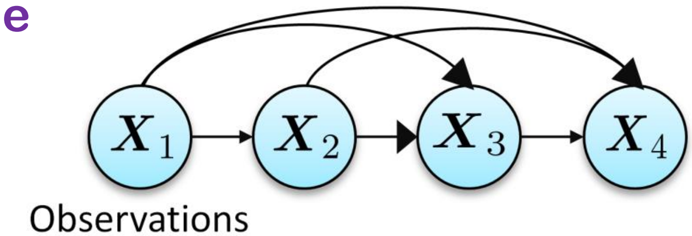
- The attentive state-space model can be interpreted as a soft version of a variable-order Markov model
- Memory is shaped by the current patient's context (clinical events, treatments, etc)



Comparison with Existing Models

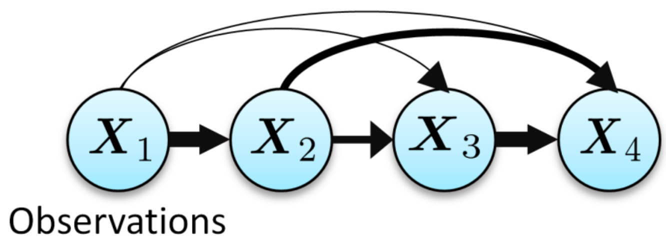
- Difference with RNNs, HMMs, HSMMs, Joint models.

Observable models



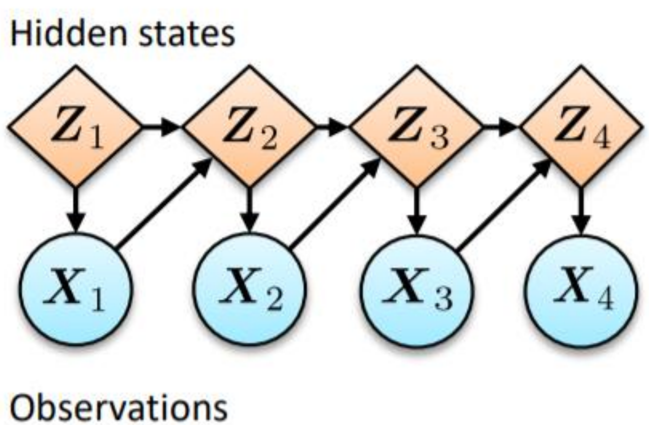
No latent structure

RETAIN



<u>Interpretable</u> predictions, <u>Uninterpretable</u> latent structure

RNN

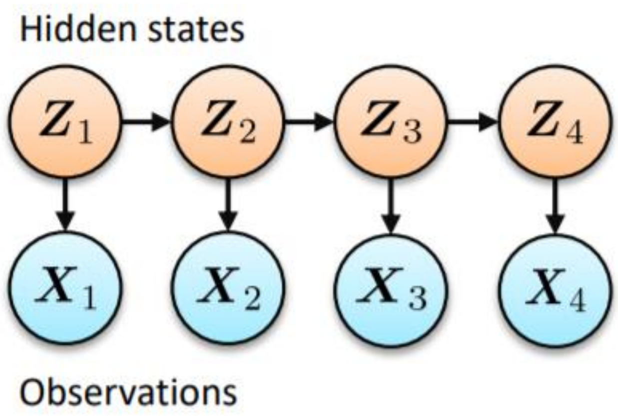


<u>Uninterpretable</u> predictions, <u>Uninterpretable</u> latent structure

Comparison with Existing Models

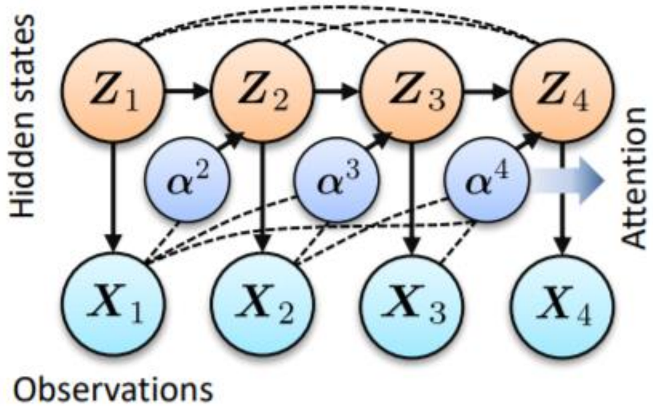
- Difference with RNNs, HMMs, HSMMs, Joint models.

HMM



<u>Memoryless</u> predictions,
<u>Interpretable</u> latent structure

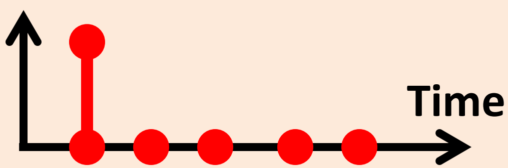
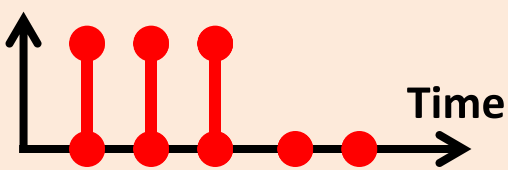
PASS



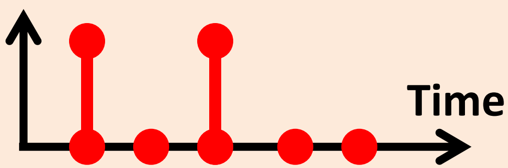
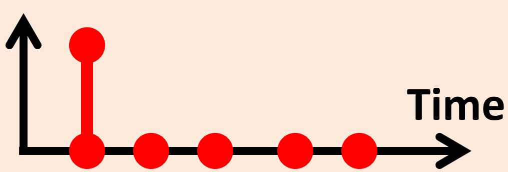
<u>Memoryful and interpretable</u> predictions,
<u>Interpretable</u> latent structure

Generality of Attentive State Space Model

- Attentive state space models reduce to various classical sequential models for certain settings of attention mechanism.

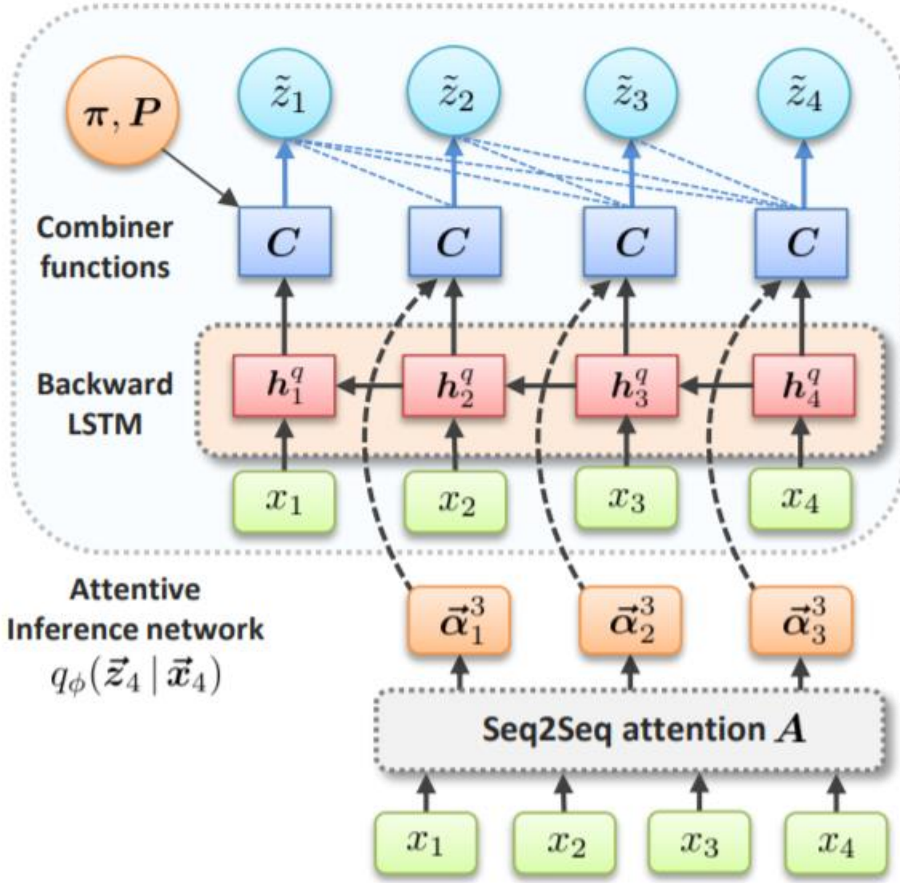
Attention mechanism	Model
1 step back attention 	Hidden Markov Model
Fixed multi-step attention 	Higher-order Hidden Markov Model

Generality of Attentive State Space Model

Attention mechanism	Model
• Phased attention with binarized attention weights  	**Variable-order Markov Model (context tree)**
• 1 step back attention with non-linear transitions  	**Deep Markov Model**

Learning and Inference

- **Sequence-to-sequence architecture for the attention mechanism**



Experiments

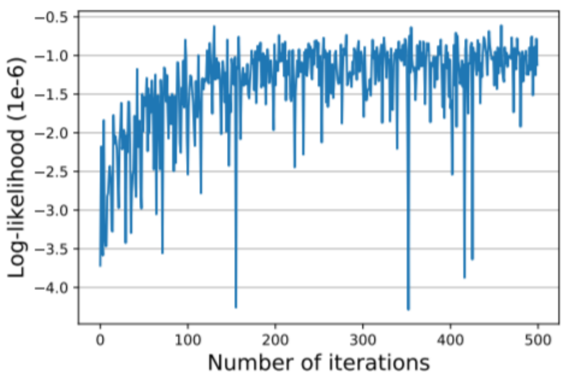
- We used data from a cohort of patients enrolled in the UK CF registry, a database held by the UK CF trust.

- The dataset records annual follow-ups for 10,263 patients over the period from 2008 and 2015, with a total of 60,218 hospital visits.

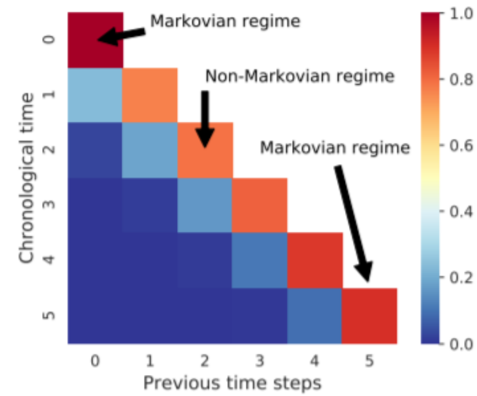
- Each patient is associated with 90 variables, including information on 36 possible treatments, diagnoses for 31 possible comorbidities and 16 possible infections, in addition to biomarkers and demographic information.

Experiments

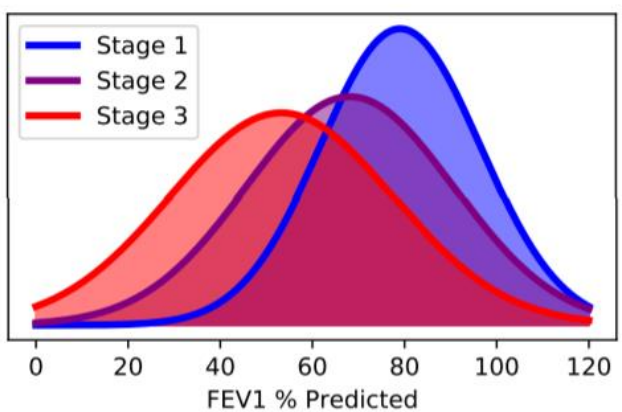
Learning curve

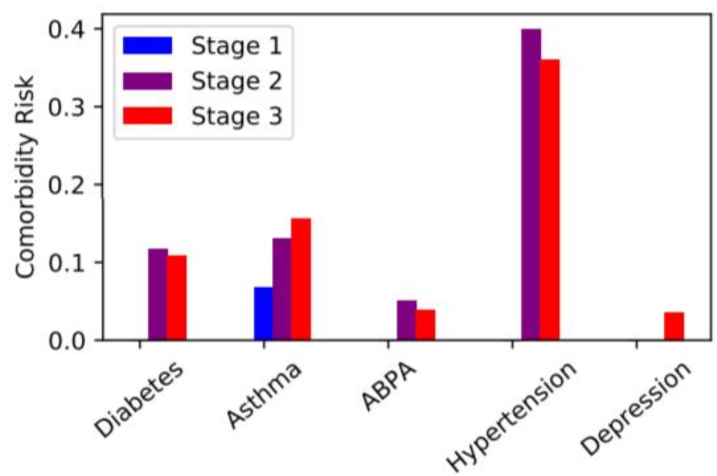


Individual trajectory



CF Phenotypes





Experiments

- Attentive state-space models do not only extract interpretable representations, but improve accuracy of prediction of comorbidities as well.

Model	Diabetes AUC-ROC	ABPA AUC-ROC	Depression AUC-ROC	Pancreatitus AUC-ROC	P. Aeruginosa AUC-ROC
Attentive SS	**0.709 $\pm$ 0.02**	**0.787 $\pm$ 0.01**	**0.751 $\pm$ 0.03**	**0.696 $\pm$ 0.04**	**0.680 $\pm$ 0.01**
HMM	0.625 $\pm$ 0.02	0.686 $\pm$ 0.03	0.667 $\pm$ 0.08	0.625 $\pm$ 0.04	0.610 $\pm$ 0.02
RNN	0.634 $\pm$ 0.03	0.727 $\pm$ 0.10	0.575 $\pm$ 0.01	0.590 $\pm$ 0.06	0.654 $\pm$ 0.01
LSTM	0.675 $\pm$ 0.03	0.740 $\pm$ 0.07	0.609 $\pm$ 0.12	0.578 $\pm$ 0.05	0.671 $\pm$ 0.01
RETAIN	0.610 $\pm$ 0.06	0.718 $\pm$ 0.05	0.580 $\pm$ 0.09	0.600 $\pm$ 0.08	0.676 $\pm$ 0.02