SurvLatent ODE : A **Neural ODE based time-to-event model** with competing risks for **longitudinal data** improves cancer-associated Venous Thromboembolism (VTE) prediction

Intae Moon[1,2], Stefan Groha[3,4,5], and Alexander Gusev[3,4,5]

[1]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA
[2]Division of Population Sciences, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA, USA
[3]Division of Genetics, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
[4]The Broad Institute of MIT and Harvard, Cambridge, MA, USA

Personal website QR code

MIT EECS

BROAD INSTITUTE

Dana-Farber Cancer Institute

Motivation

- Many risk prediction tools rely on **simple, linear scoring system** using a small number of features.
- Learning from electronic health records (EHR) data for predicting clinical outcomes is challenging due to data irregularities such as i) **irregularly sampled measurements**, ii) **loss to follow-up** (i.e. right-censored), and iii) **competing events**

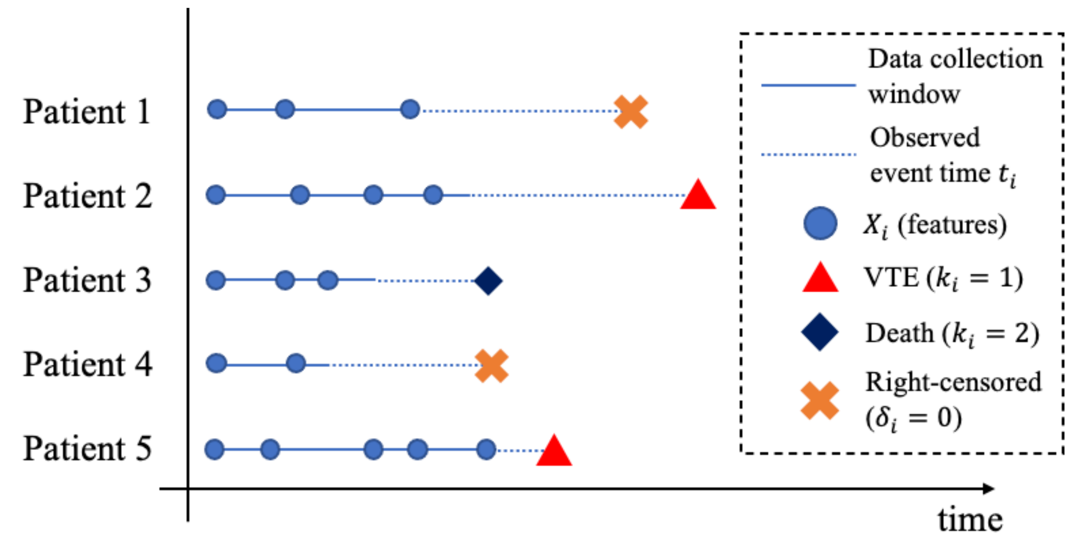


Fig. 1 : illustration of time-to-event data with time-varying features

Method overview

Fig. 2 : *Simplified* illustration of SurvLatent ODE

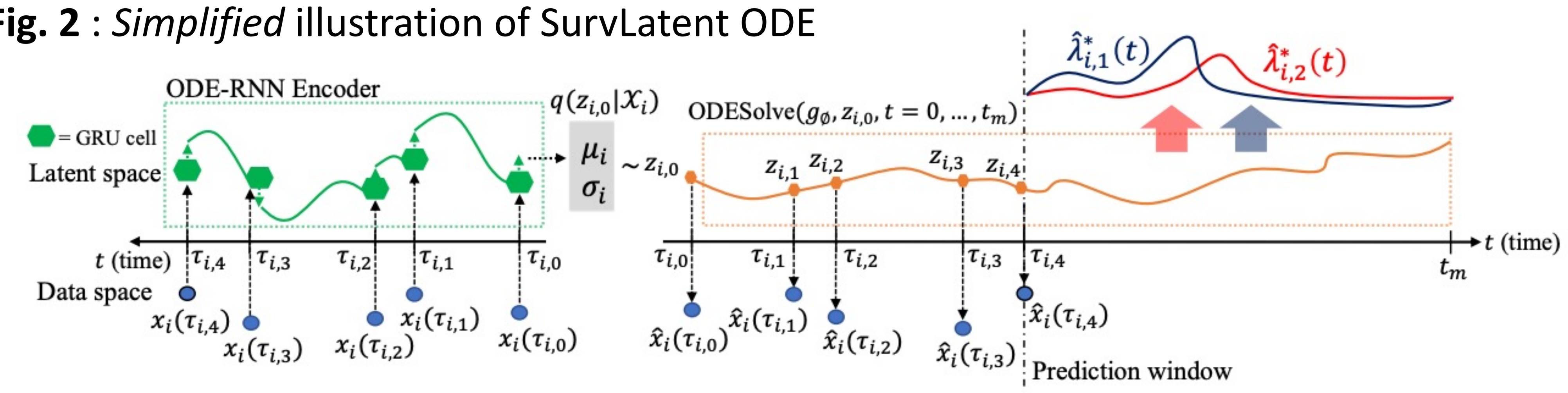
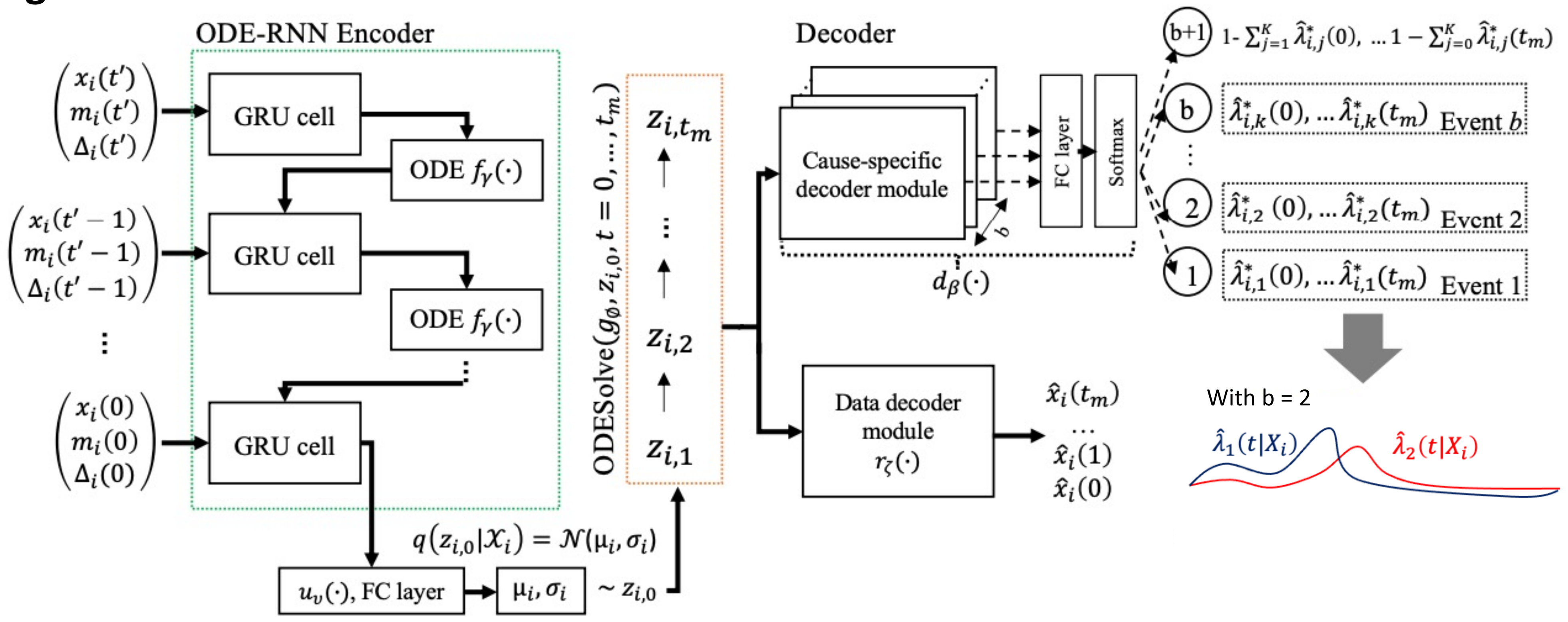


Fig. 3 : *Detailed* model architecture of SurvLatent ODE



Inference : the loss function handles **right-censoring** and **multiple competing events**

$$L_{\text{total}}(\mathcal{D}; \Phi, \zeta, \beta) = -\boxed{\text{ELBO}(\mathcal{X}; \Phi, \zeta)} - \boxed{\log(L_{\text{surv}}(\mathcal{D}; \Phi, \beta))}$$

$$\text{ELBO}(\mathcal{X}; \Phi, \zeta) = \mathbb{E}_{q(z_0|\mathcal{X};\Phi)}[\log(p(\mathcal{X}|z_0; \phi, \zeta)] - \text{KL}[q(z_0|\mathcal{X}; \Phi) \,||\, p(z_0)]$$

, where $p(z_0) \sim \mathcal{N}(0,1)$ and Φ, ζ, β are model parameters

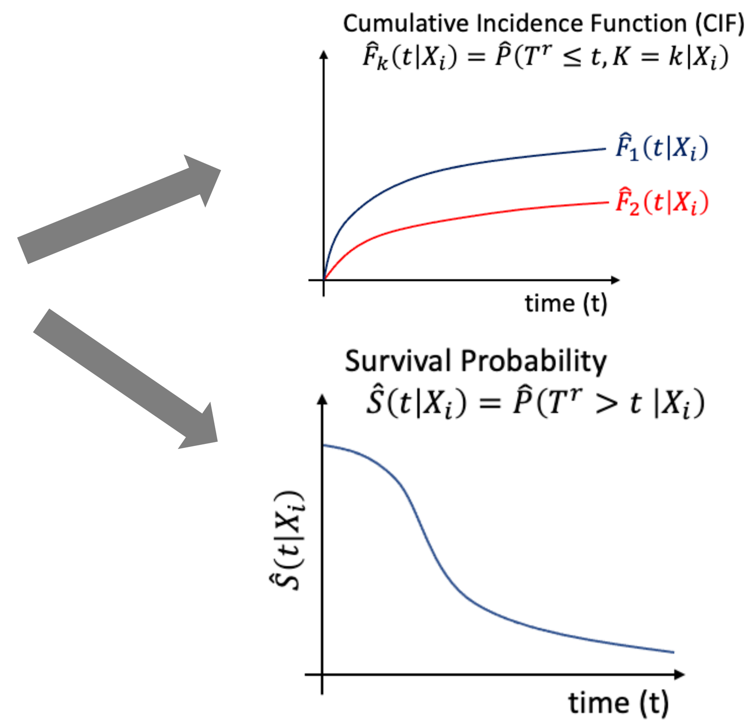
Survival dataset, $\mathcal{D}$, is a set of tuples $\{(t_i, k_i, \delta_i, X_i)\}_{i=1}^N$ (see Fig. 1)

$$L_{\text{surv}}(\mathcal{D}; \Phi, \beta) = \prod_{i \in \mathcal{D}} \hat{P}(T^r = t_i^r, K = k_i; \Phi, \beta)^{\delta_i} \times \hat{P}(T^r > t_i^r; \Phi, \beta)^{1-\delta_i}$$

$$= \prod_{i \in \mathcal{D}} \underbrace{\left[\hat{\lambda}^*_{i,k}(t_i^r; \Phi, \beta)\hat{S}(t_i^r - 1|X_i; \Phi, \beta)\right]^{\delta_i}}_{\text{Uncensored: } \delta_i = 1} \times \underbrace{\hat{S}(t_i^r|X_i; \Phi, \beta)^{1-\delta_i}}_{\text{Censored: } \delta_i = 0}$$

Method key takeaways

- Takes **irregularly sampled patient's data** and flexibly learns **dynamics of a patient-specific latent health trajectory** using **neural networks**
- Learns **hazard function trajectory $\hat{\lambda}^*_k(t)$** which enables handling **loss to follow-up** (i.e. right censoring) and **predicting time-to-event**.
- Utilizes a **multi-task learning framework** to handle **competing risks** where the latent trajectory $Z^t(t)$ is shared across multiple events while each cause-specific decoder module captures signals specific to each outcome.



Experimental results

- Model comparison in terms of key strategies in handling longitudinal data

	Handles time-varying features	Handles Competing risks	*Handling missing measurements* Include missing pattern in input	Mean imputation/ forward-filling	*Learning latent state dynamics* Exponential decay	Neural Networks	Generative model (VAE)
SurvLatent ODE (Proposed model)	V	V	V			V	V
Surv RNN-VAE (Modified from Che et al. (2018))	V		V		V		V
RDSM (Nagpal et al., 2021)	V			V			
Dynamic-Deephit (Lee et al., 2020)	V	V	V				
Cox PH[1] (Cox, 1972)				V			

Experient 1 (MIMIC-III data)

(I) **SurvLatent ODE** outperforms conventional as well as SOTA (state-of-the-art) time-to-event models for predicting time to mortality.

Time to hospital mortality prediction (MIMIC-III):

- Time-varying measurements of the first 36 hours of the admission
- N_{train} = 11,950 (55%), N_{valid} = 3,259 (15%), N_{test} = 6,519 (30%)

	Time-dependent AUC(t) 25th percentile (Hour 35)	50th percentile (Hour 81)	75th percentile (Hour 150)	Brier Score, BS(t) 25th percentile (Hour 35)	50th percentile (Hour 81)	75th percentile (Hour 150)
SurvLatent ODE (Proposed model)	**0.920 (0.009)**	**0.883 (0.009)**	**0.831 (0.010)**	**0.0220 (0.0013)**	**0.0442 (0.0019)**	0.0789 (0.0029)
Surv RNN-VAE	0.535 (0.022)**	0.535 (0.016)**	0.521 (0.014)**	0.0281 (0.0017)**	0.0571 (0.0023)**	0.0950 (0.003)**
RDSM	0.836 (0.017)**	0.817 (0.013)**	0.784 (0.011)**	0.0241 (0.0018)*	0.0449 (0.0023)	**0.0618 (0.0025)**
Dynamic-Deephit	0.891 (0.009)**	0.860 (0.009)*	0.808 (0.010)*	0.0247 (0.0018)*	0.0492 (0.0024)**	0.0816 (0.0032)
Cox PH	0.826 (0.017)**	0.806 (0.013)**	0.762 (0.012)**	0.0234 (0.0017)	0.0465 (0.0023)	0.0789 (0.0032)

Predicting time to **hospital mortality**

Experient 2 (Dana-Farber Cancer Institute VTE dataset)

(II) **SurvLatent ODE** outperforms SOTA as well as current clinical standards, *Khorana scores*, for predicting time to VTE.

Time to VTE prediction with **death** as a **competing event**:

- **Venous Thromboembolism (VTE)** is a frequent, yet **fatal complication in patients with active cancer.**
  - Preventive measures (e.g. thromboprophylaxis) are effective but come with side effects.
- Utilized Dana-Farber Cancer Institute dataset with total 54 features collected up to a year after admission.
  - N_{train} = 4,797 (55%), N_{valid} = 1,307 (15%), N_{test} = 2,630 (30%)

	Time-dependent AUC$_k$(t) 25th percentile (Day 47)	50th percentile (Day 113)	75th percentile (Day 266)	Brier Score, BS$_k$(t) 25th percentile (Day 47)	50th percentile (Day 113)	75th percentile (Day 266)
SurvLatent ODE (Proposed model)	**0.782 (0.031)**	**0.781 (0.021)**	**0.758 (0.020)**	**0.0222 (0.0026)**	**0.0426 (0.0034)**	**0.0631 (0.0040)**
Dynamic-Deephit	0.729 (0.032)*	0.770 (0.023)	0.722 (0.022)*	0.0223 (0.0027)	0.0428 (0.0036)	0.0641 (0.0042)*
CS Cox	0.702 (0.034)*	0.728 (0.023)*	0.706 (0.020)*	0.0228 (0.0028)	0.0444 (0.0038)*	0.0666 (0.0045)**
FG Cox	0.672 (0.033)*	0.695 (0.024)*	0.686 (0.019)**	0.0227 (0.0028)	0.0445 (0.0038)*	0.0667 (0.0045)**
Khorana scores[2]	0.625 (0.051)	0.592 (0.036)	0.581 (0.029)	N/A	N/A	N/A
Khorana scores (imputed)	0.627 (0.036)**	0.628 (0.026)**	0.628 (0.022)**	N/A	N/A	N/A

Predicting time to **Venous Thromboembolism (VTE)**

	Time-dependent AUC$_k$(t) 25th percentile (Day 141)	50th percentile (Day 269)	75th percentile (Day 446)	Brier Score, BS$_k$(t) 25th percentile (Day 141)	50th percentile (Day 269)	75th percentile (Day 446)
SurvLatent ODE (Proposed model)	**0.772 (0.018)**	**0.762 (0.014)**	**0.761 (0.012)**	0.0785 (0.0028)	0.1270 (0.0031)	0.1626 (0.0032)
Dynamic-Deephit	0.762 (0.018)	0.742 (0.014)*	0.741 (0.012)**	**0.0662 (0.0041)**	0.1198 (0.0047)	0.1600 (0.0044)
CS Cox	0.711 (0.020)**	0.710 (0.014)**	0.710 (0.013)**	0.0675 (0.0046)	0.1204 (0.0055)	0.1602 (0.0054)
FG Cox	0.706 (0.020)**	0.706 (0.014)**	0.706 (0.013)**	0.0667 (0.0045)	**0.1179 (0.0050)**	**0.1580 (0.0048)**

Predicting time to **all-cause mortality** as the competing event

(III) **SurvLatent ODE provides clinically meaningful and interpretable latent representations.**

Fig. 4 : Cumulative incidence trajectories for the VTE event across the latent clusters.
Fig. 5 : Heatmap visualization of feature-wise coefficient for each latent cluster

Found significant associations of the elevated risk group with well-known risk factors for VTE (shown in red and yellow boxes) such as SGOT, White Blood Cell counts, Platelet counts, etc.

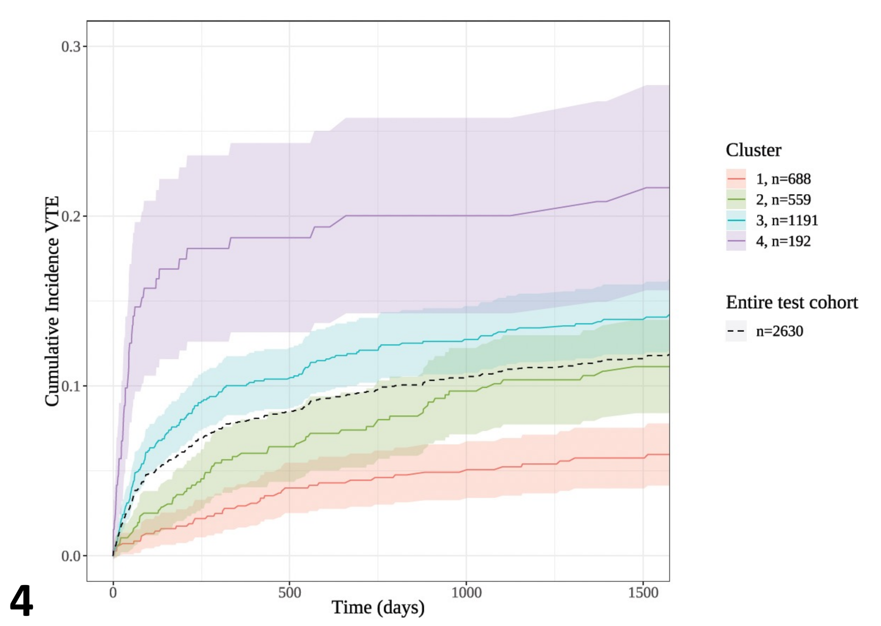


Fig. 4

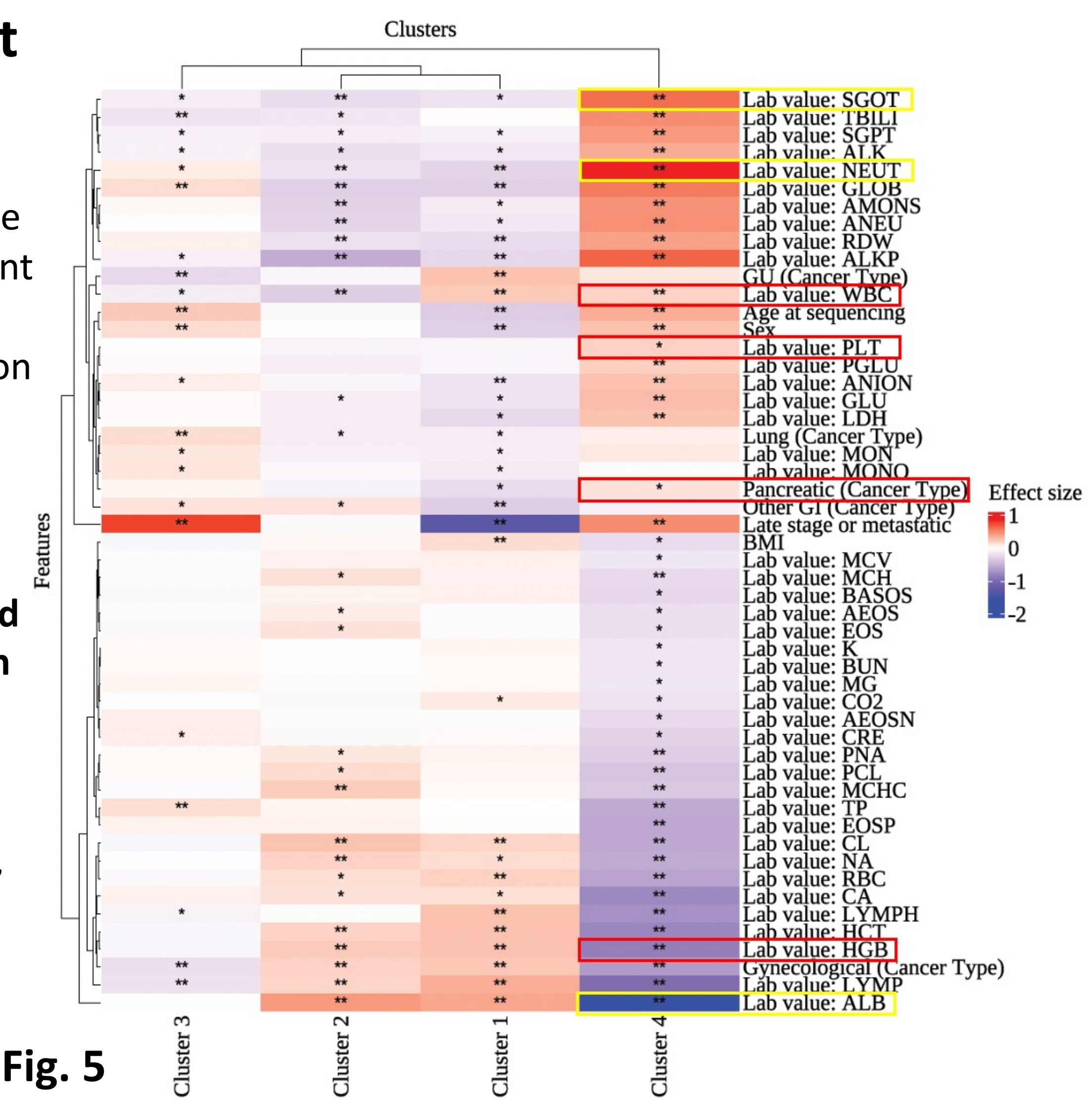


Fig. 5

Conclusion

First demonstration of the ODE-based variational autoencoder time-to-event model for longitudinal data, which

- Handles **irregularly sampled data**
- Handles **right-censored** patients
- Flexibly estimates hazard functions for the event of interest as well as **competing events** via a multi-task learning framework

Future work

- Extends to a multimodal framework which includes patient's tumor genetics data
- Well-calibrated uncertainty of time-to-event predictions as well as latent dynamics