MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

MANCHESTER 1824 The University of Manchester

Using observation processes to predict survival: A deep learning approach to joint modelling

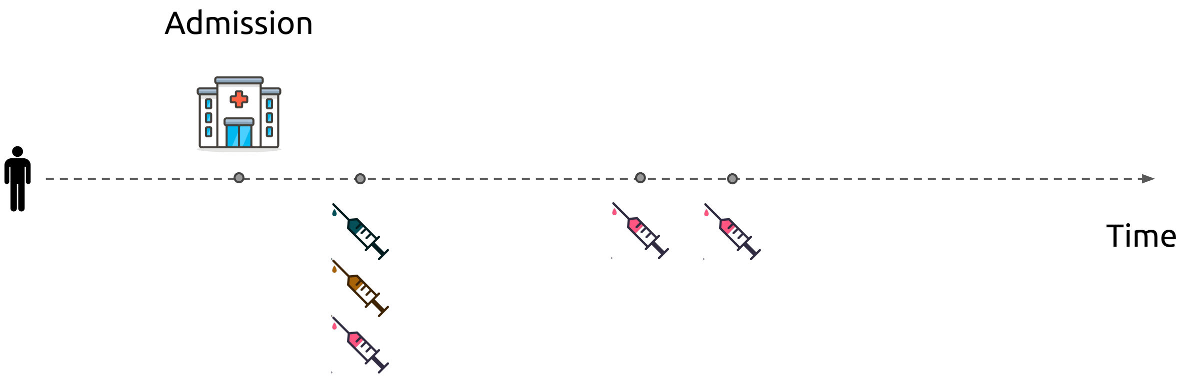
Vincent Jeanselme, Glen Martin, Matthew Sperrin,
Niels Peek, Brian Tom and Jessica Barrett

International Biometric Conference
10^{th} - 15^{th} July 2022

Clinical Presence

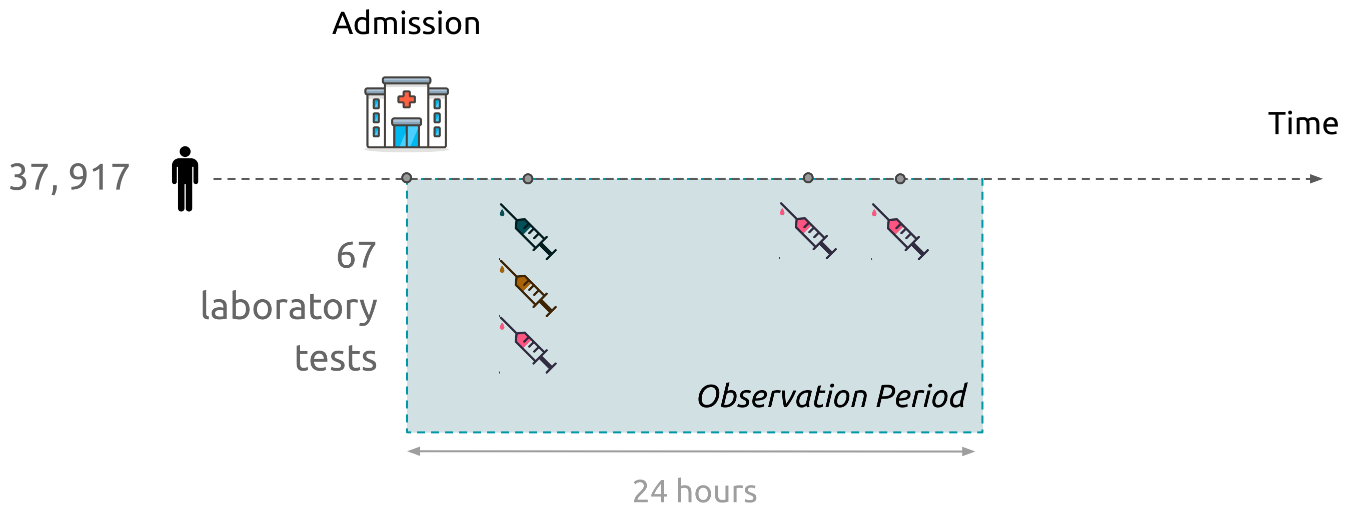
Clinical Presence

Clinical Presence



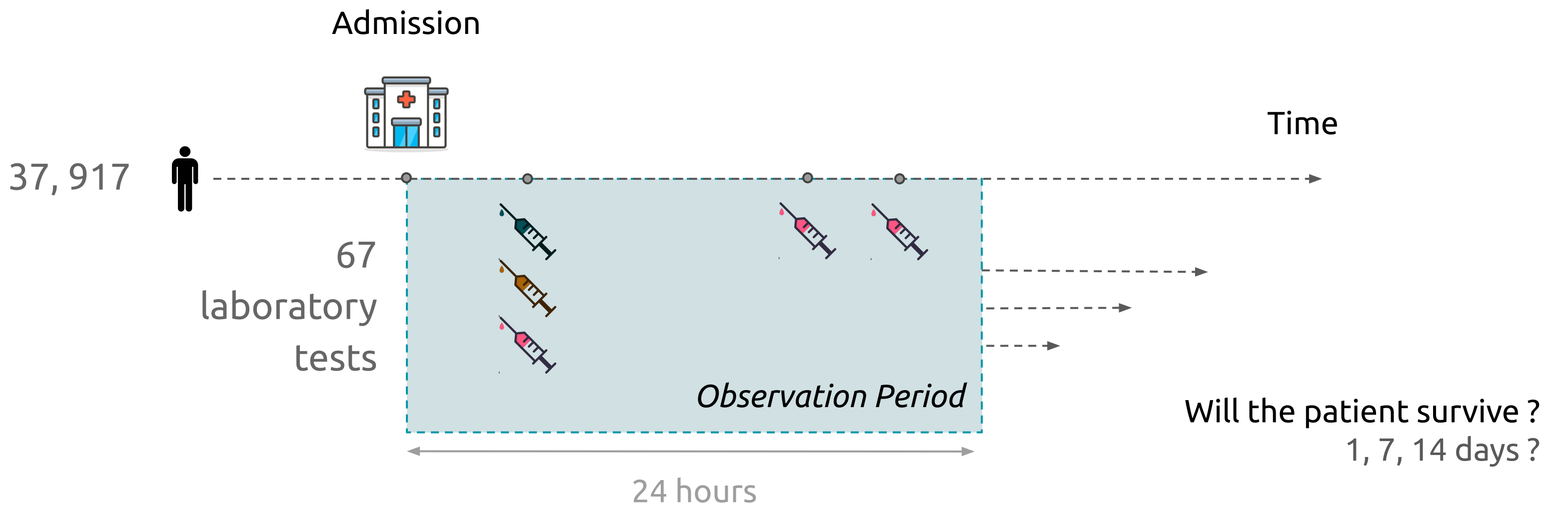
*The **sampling process** is imprinted by the interaction between **patients** and the **healthcare system**.*

Motivating Example

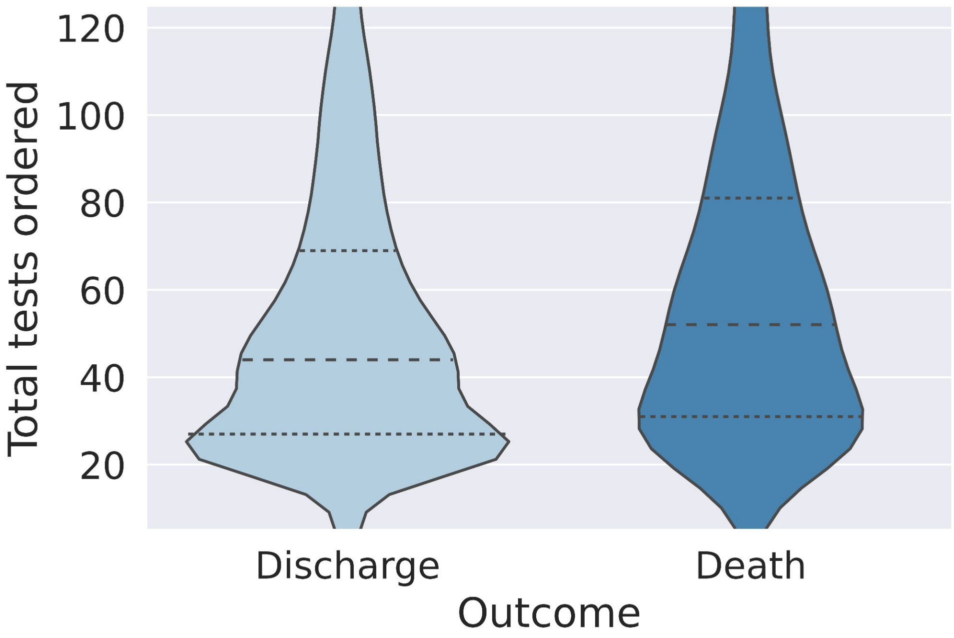


MIMIC-III, a freely accessible critical care database by A Johnson & al., 2016

Motivating Example



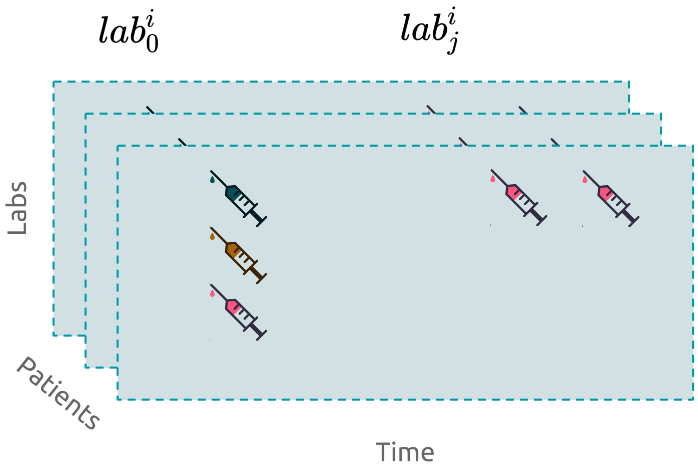
Evidence of Clinical Presence



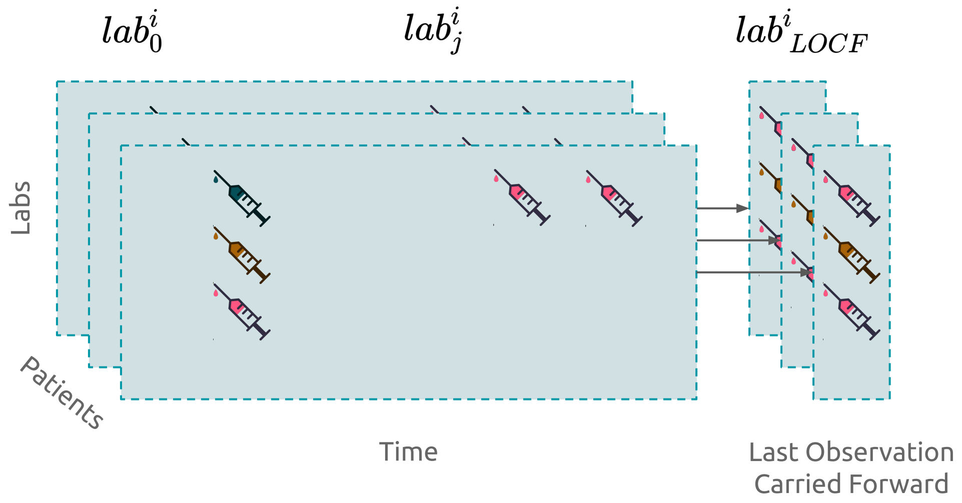
Proposed Approach

1. CoxPH
2. DeepSurv
3. Recurrent Neural Network
4. **DeepJoint**

Notation

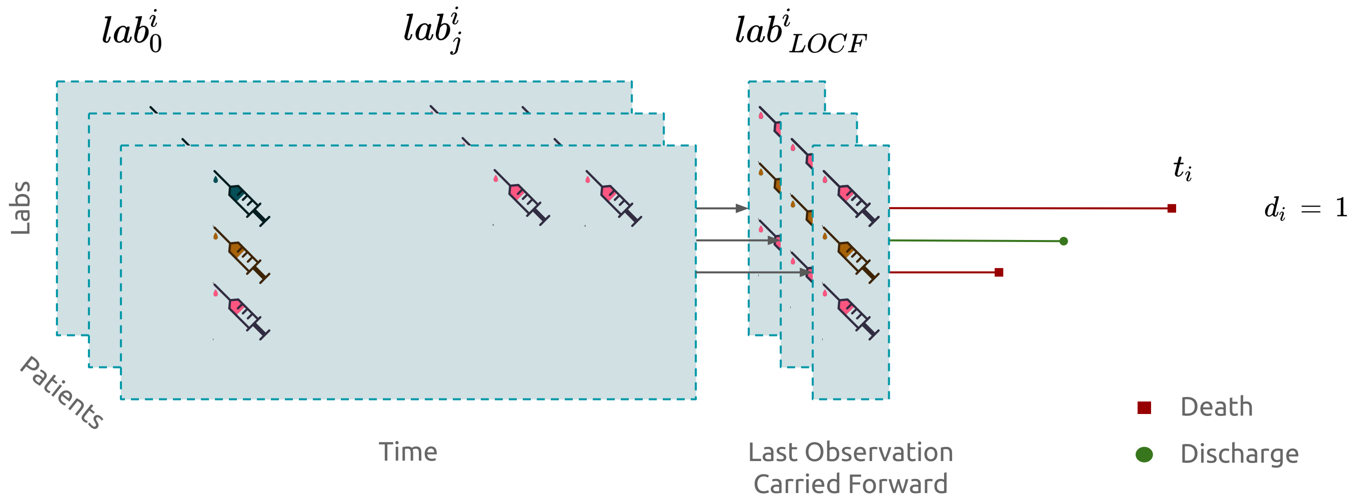


Notation



lab_0^i
lab_j^i
$lab^i{}_{LOCF}$
Labs
Patients
Time
Last Observation
Carried Forward

Notation



lab_0^i
lab_j^i
$lab^i{}_{LOCF}$
t_i
$d_i = 1$
Labs
Patients
Time
Last Observation
Carried Forward
Death
Discharge

Cox Model

Baseline Hazard

$$\lambda(s \mid lab_{LOCF}) = \lambda_0(s) \exp\left(\alpha^T lab_{LOCF}\right)$$

Hazard

Covariates Drift

Regression Models and Life-Tables by D Cox, 1972

DeepSurv

$$\lambda(s \,|\, lab_{LOCF}) \;=\; \lambda_0(s)\, \exp\left(\alpha^T lab_{LOCF}\right)$$

$$\lambda(s \,|\, lab_{LOCF}) \;=\; \lambda_0(s)\, \exp\left(h(lab_{LOCF})\right)$$

Neural Network Interaction

DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network by J Katzman & al, 2018

DeepSurv

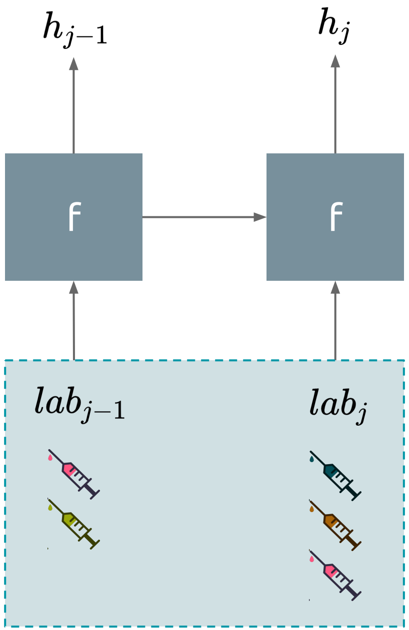
$$\lambda(s \mid lab_{LOCF}) = \lambda_0(s) \exp\left(\alpha^T lab_{LOCF}\right)$$

$$\lambda(s \mid lab_{LOCF}) = \lambda_0(s) \exp\left(h(lab_{LOCF})\right)$$

$$l_{DS} = \sum_{i,\, d_i=1} h\left(lab^{i}_{LOCF}\right) - \log \sum_{j,\, t_j > t_i} \exp h\left(lab^{j}_{LOCF}\right)$$

DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network by J Katzman & al, 2018

Recurrent Neural Network

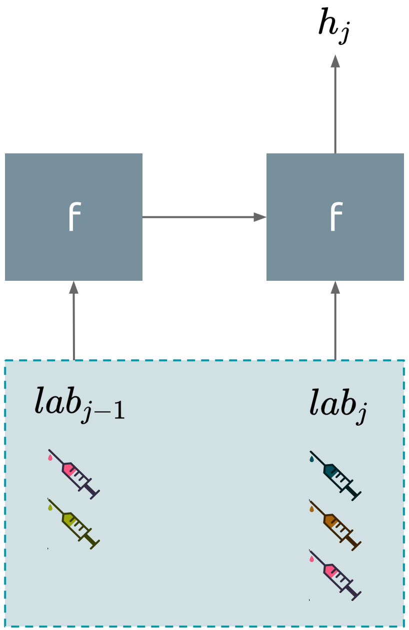


A Recurrent Neural Network (RNN) is a neural network that extracts a hidden representation of the data by taking advantage of its **sequential nature**

$$h_j = f(h_{j-1},\, lab_j)$$

Embedding

Recurrent Neural Network

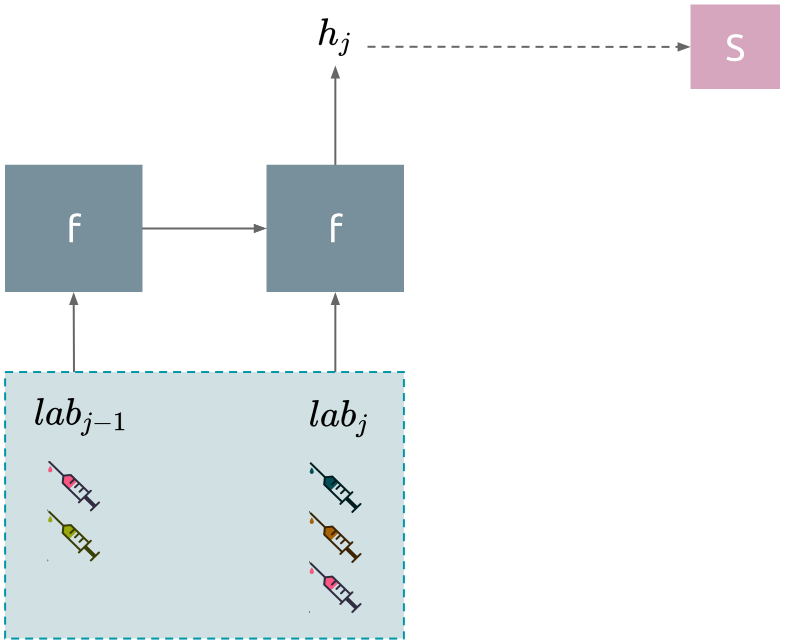


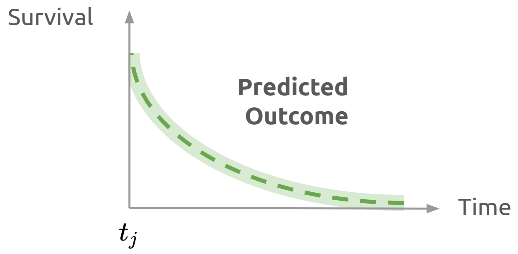
We leverage this embedding for modelling the survival outcome with DeepSurv

$$h_j = f(h_{j-1},\, lab_j)$$

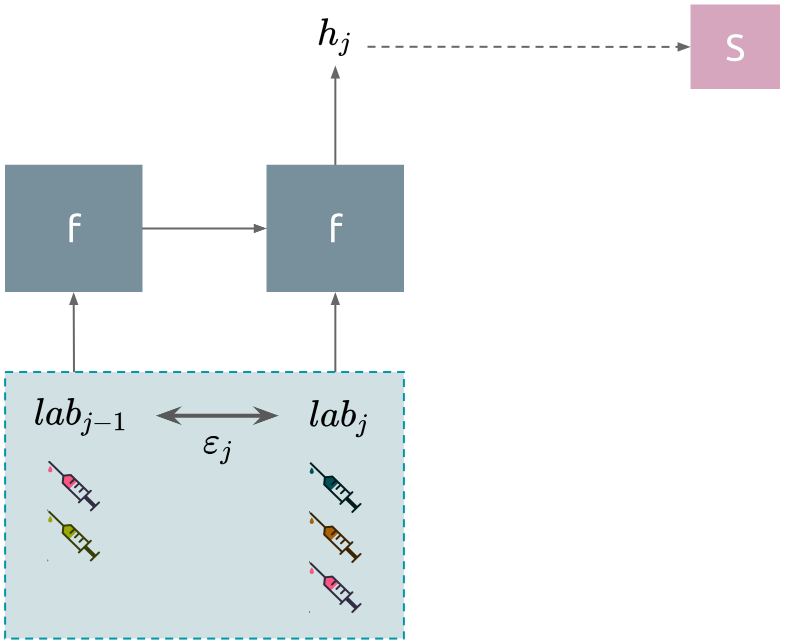
$$\lambda(s \,|\, \rightarrow lab_j) \;=\; \lambda_0(s)\, \exp\left(h_j\right)$$

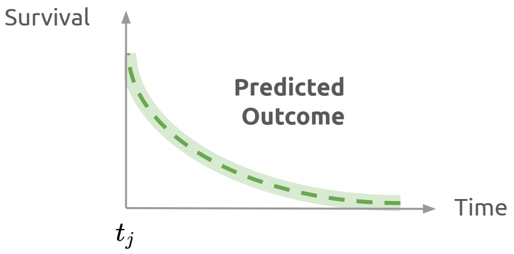
Recurrent Neural Network





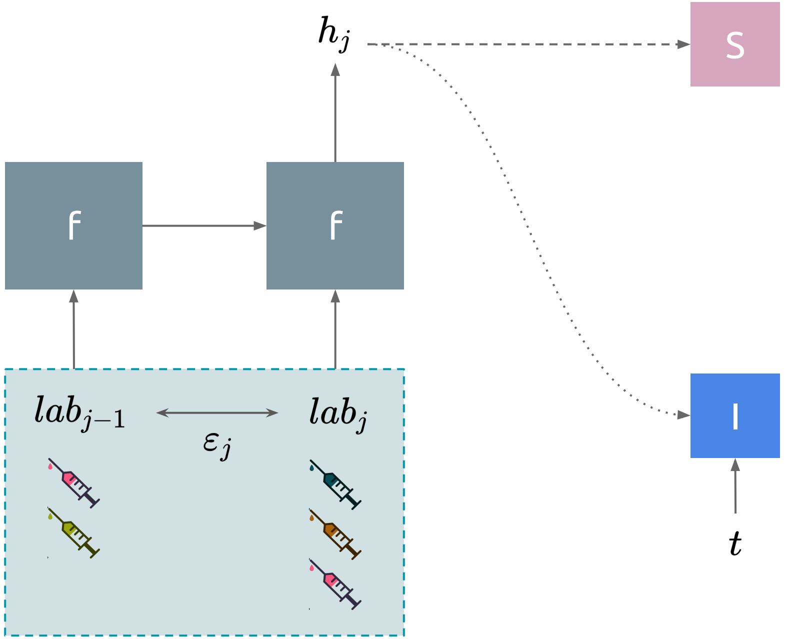
Proposed Approach





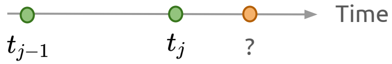
We propose to model the **interevent time** as a **temporal point process**

Proposed Approach



$$\Lambda_I(\varepsilon_j \,|\, h_{j-1}) = I(\varepsilon_j, \, h_{j-1})$$

Cumulative Hazard

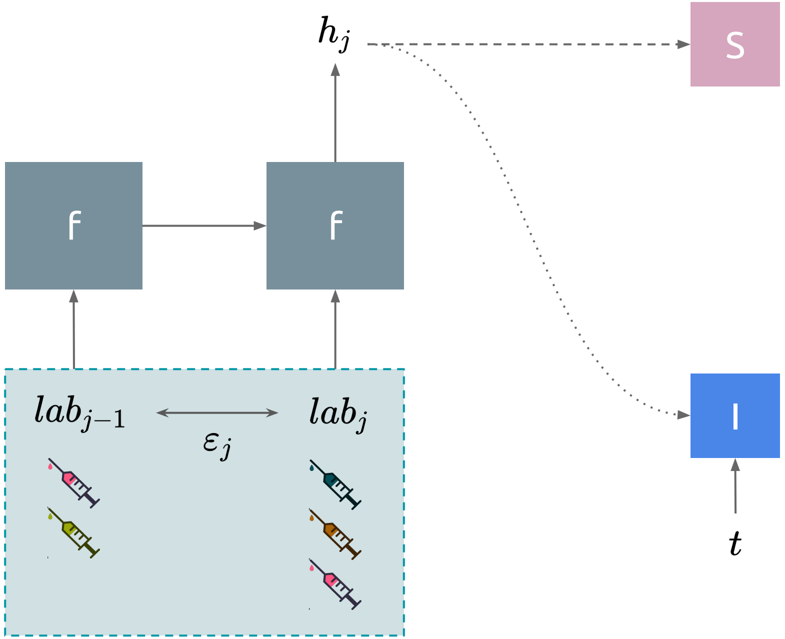


I models the **cumulative hazard** of observing an event

It is learnt by maximizing the likelihood by back propagation

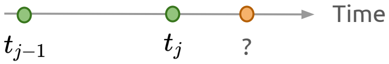
Fully Neural Network based Model for General Temporal Point Processes by T Omi & al., 2019

Proposed Approach



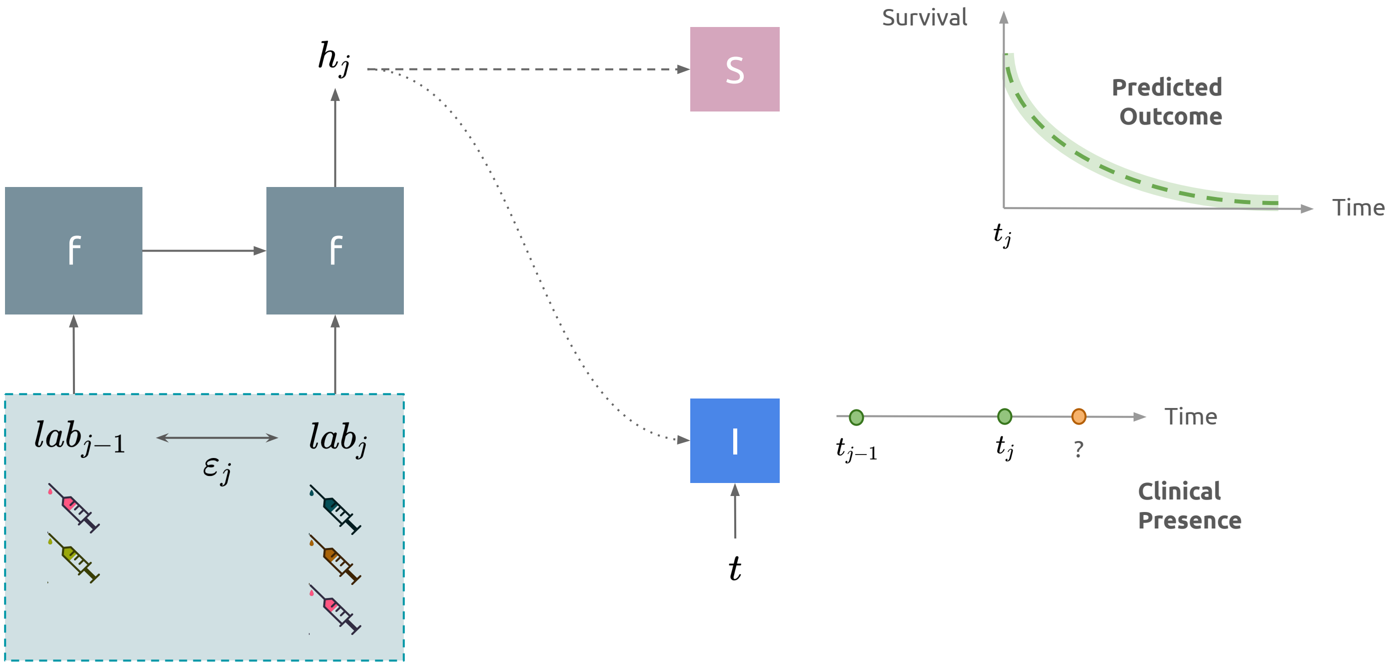
$$\Lambda_I(\varepsilon_j \mid h_{j-1}) = I(\varepsilon_j, h_{j-1})$$

Cumulative Hazard



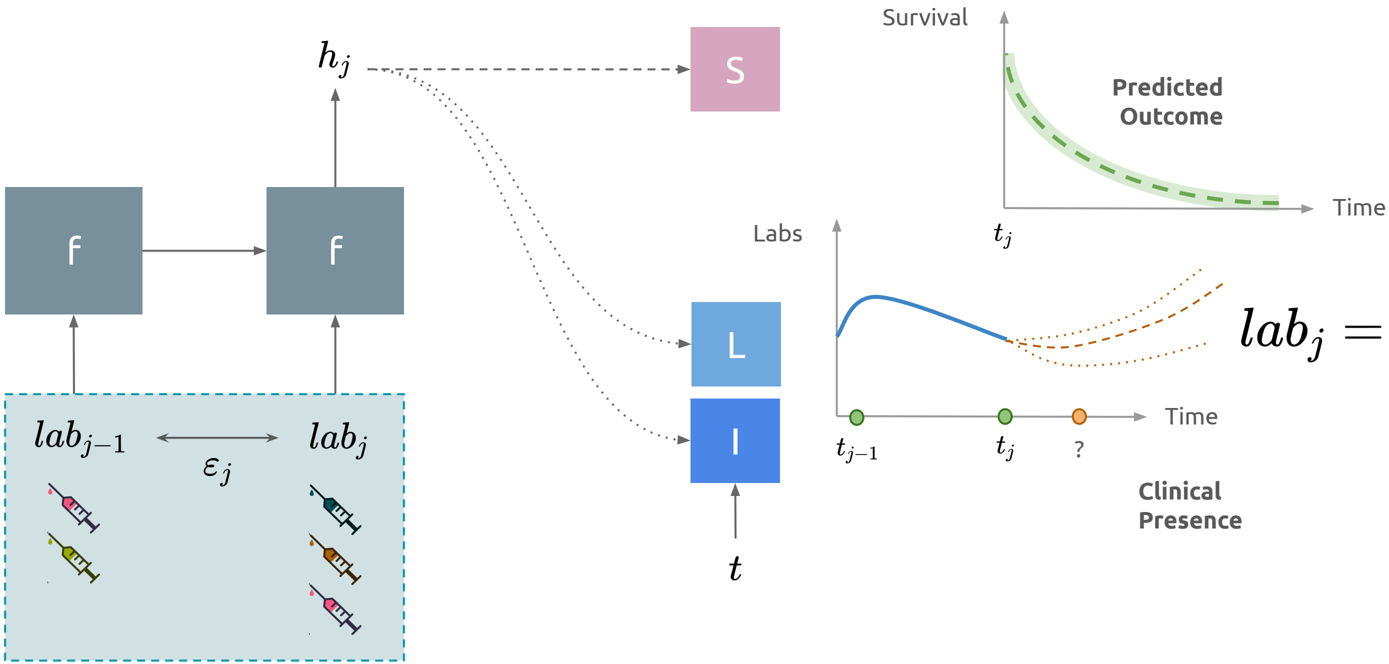
$$l_{TPP} = \sum_j \log \left. \frac{d\Lambda(\varepsilon \mid h_{j-1})}{d\varepsilon} \right|_{\varepsilon = \varepsilon_j} - \Lambda(\varepsilon_j \mid h_{j-1})$$

DeepJoint



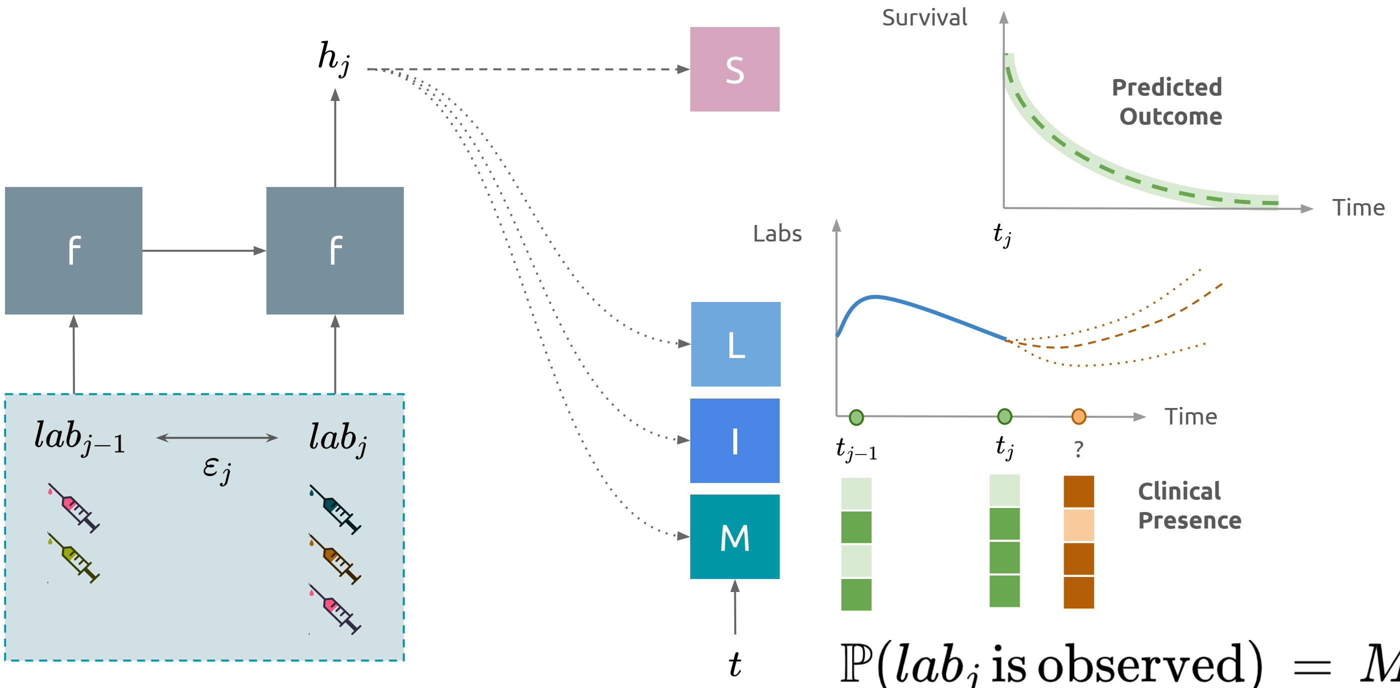
h_j
S
Survival
Predicted Outcome
Time
t_j
f
f
lab_{j-1}
ε_j
lab_j
I
t
Time
t_{j-1}
t_j
?
Clinical Presence

DeepJoint



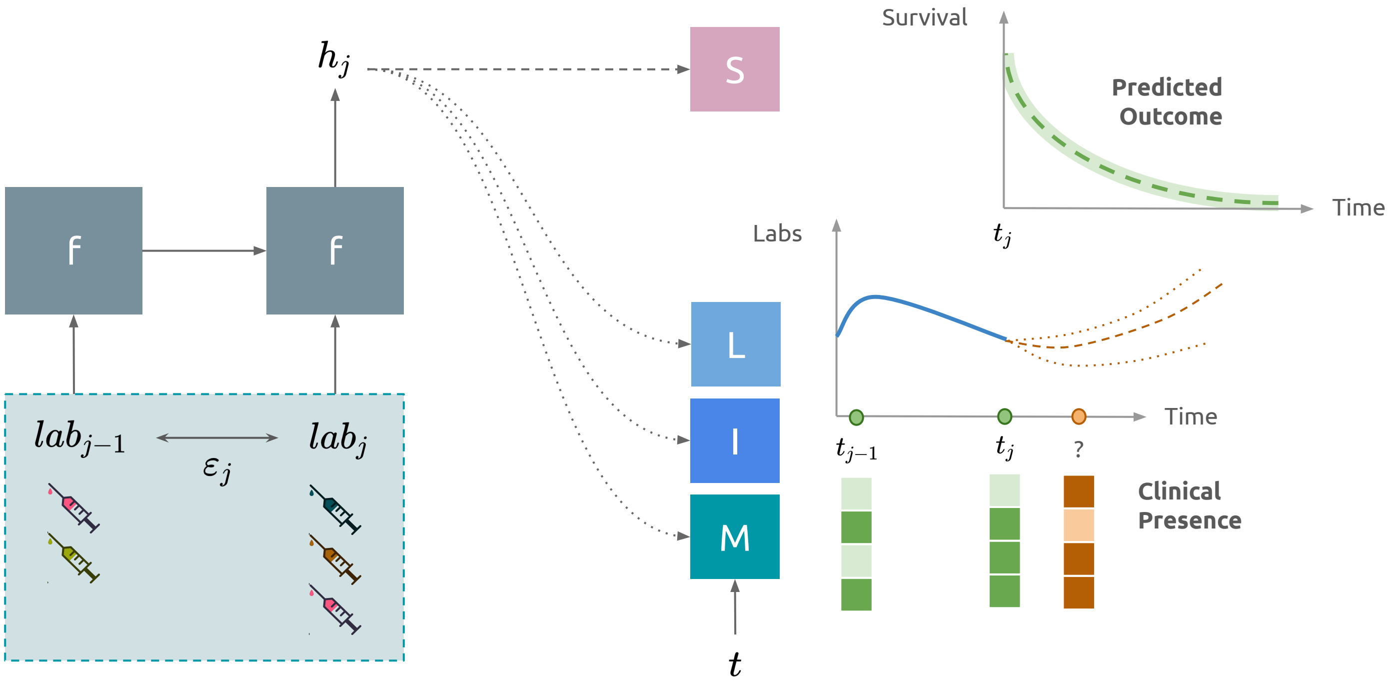
$$lab_j = L(\varepsilon_j,\, h_{j-1})$$

DeepJoint



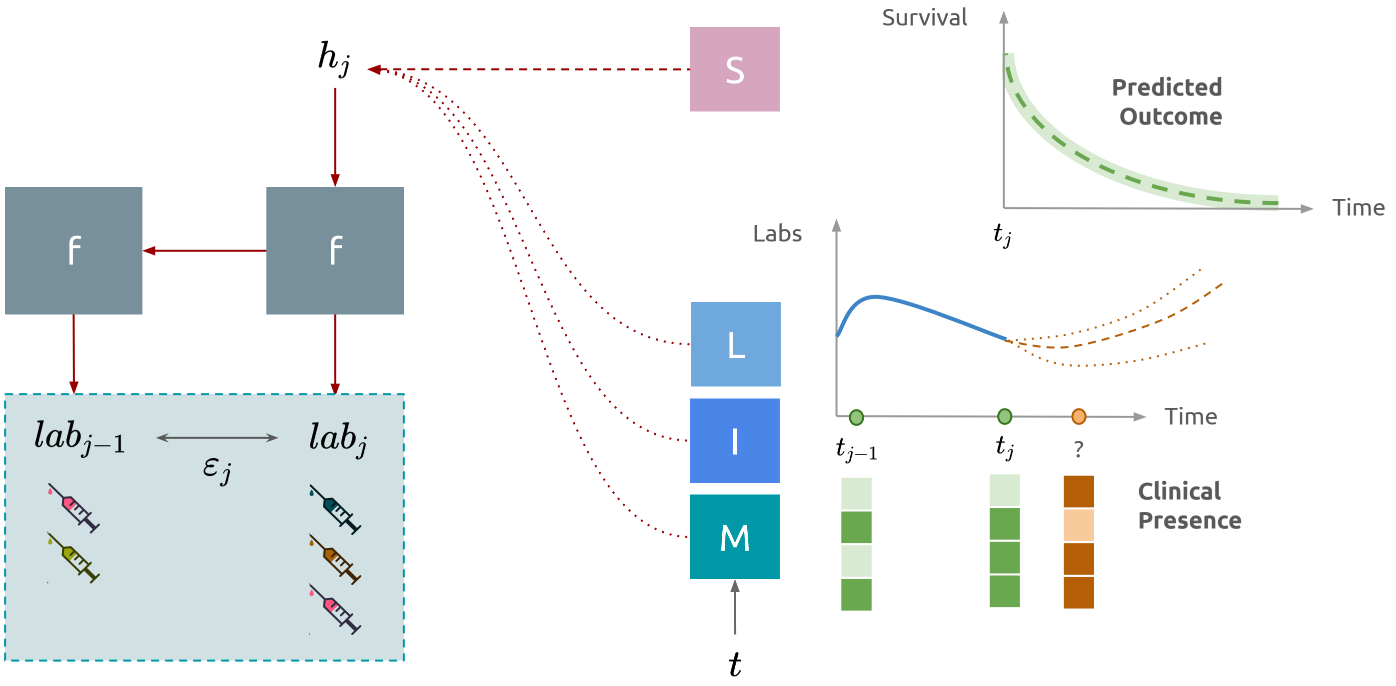
$$\mathbb{P}(lab_j \text{ is observed}) = M(\varepsilon_j, h_{j-1})$$

DeepJoint



h_j
S
f
f
L
I
M
lab_{j-1}
ε_j
lab_j
t
Survival
Predicted Outcome
Time
t_j
Labs
Time
t_{j-1}
t_j
?
Clinical Presence

DeepJoint



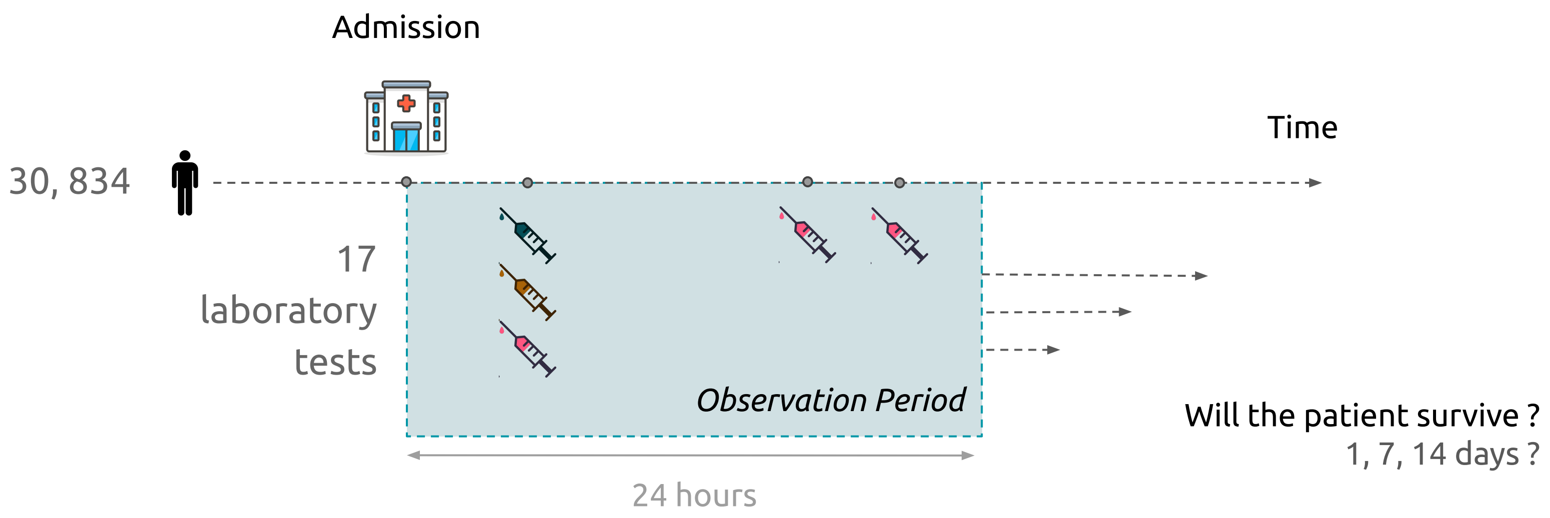
$$L_{tot} = L_{DS} + \alpha(L_{TPP} + L_L + L_M)$$

Application

MIMIC III Dataset

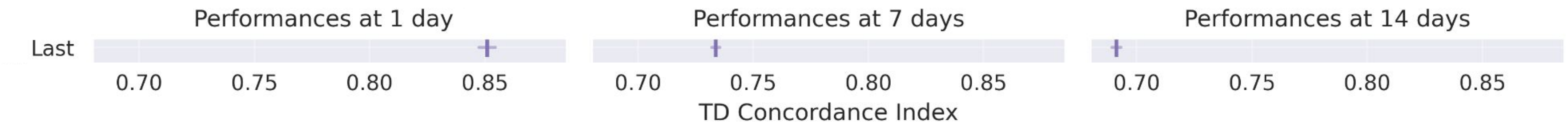
1. Predictive performance
2. Robustness performance

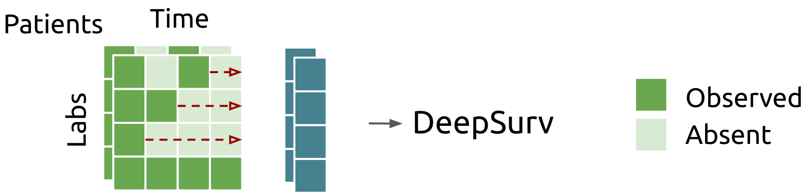
MIMIC Dataset - Preprocessed



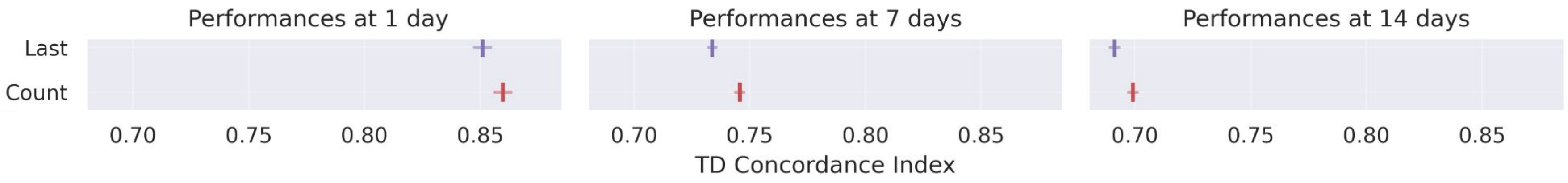
MIMIC-III, a freely accessible critical care database by A Johnson & al., 2016

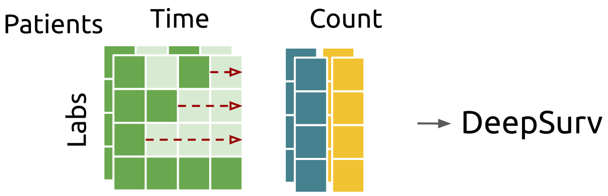
C-Index Performance



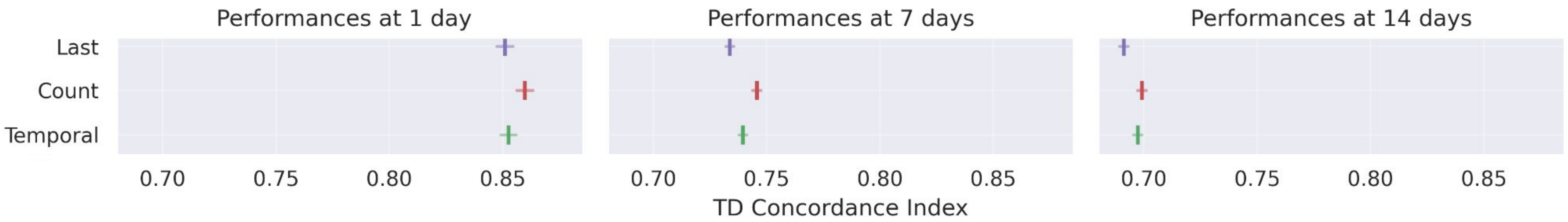


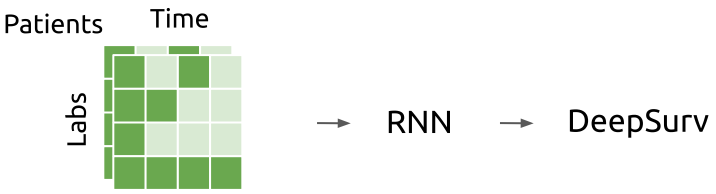
C-Index Performance



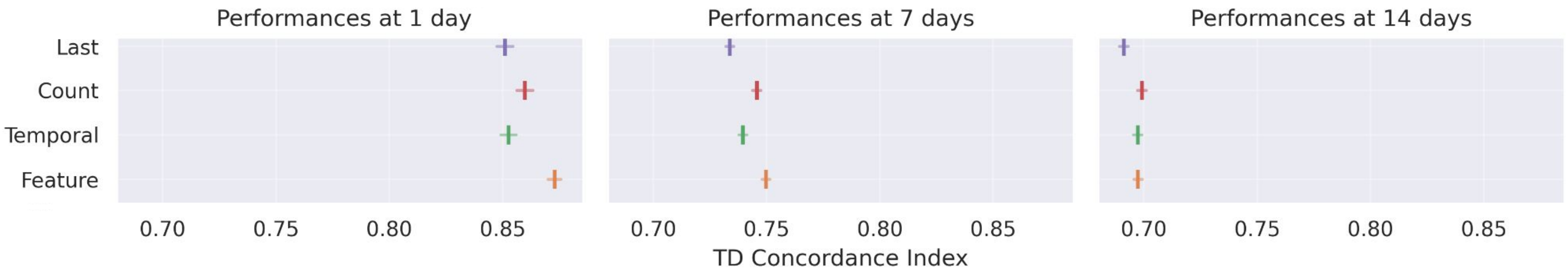


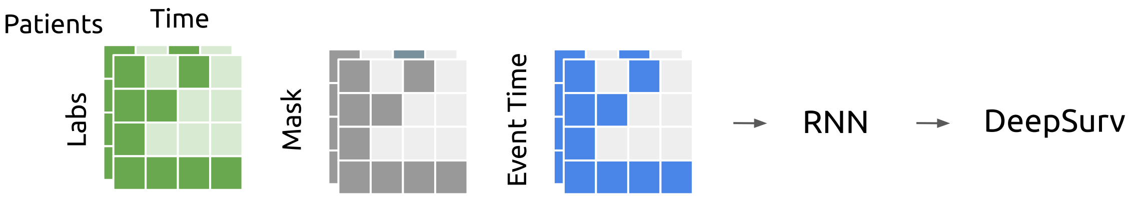
C-Index Performance



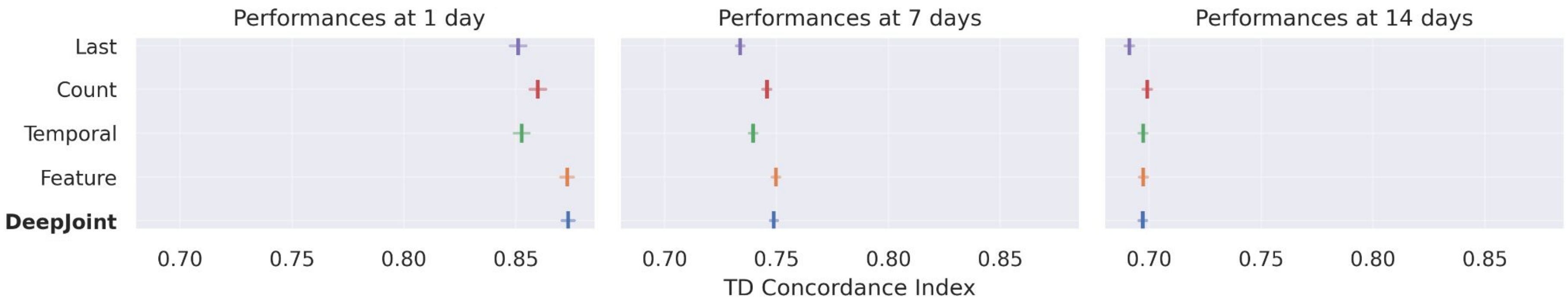


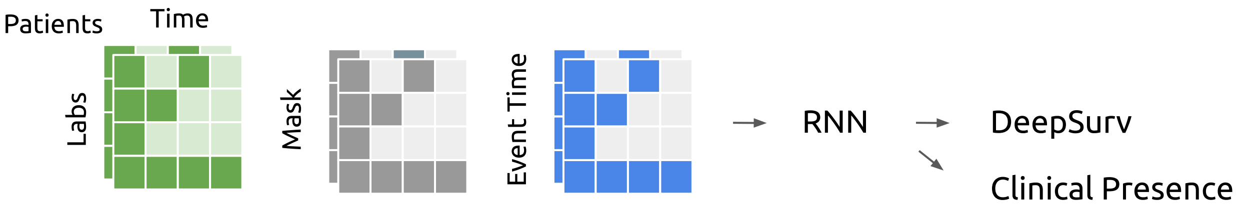
C-Index Performance





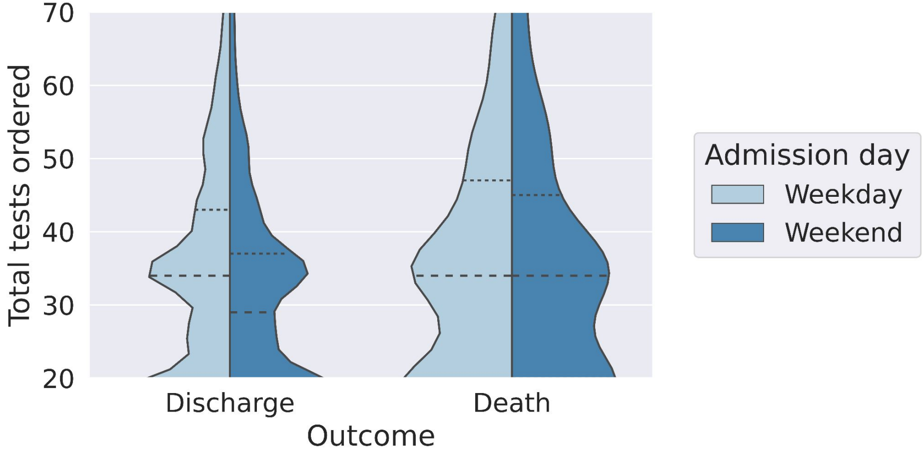
C-Index Performance





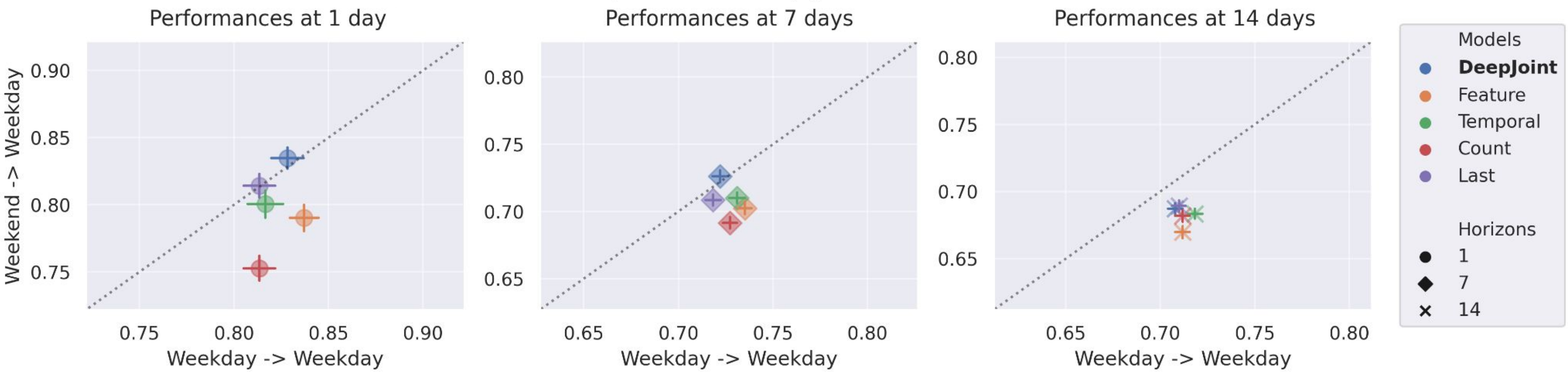
Change in Observation Process

Inspired by the **weekend effect**, we split the patients between weekdays and weekend admission to evaluate if a **shift in the observation process** would impact the survival predictive performances.



Robustness Evaluation

Robustness to Weekend Effect



Conclusions

- **Clinical presence** leads to **informative** observation process

- **Leveraging** this information **improves predictive** performance

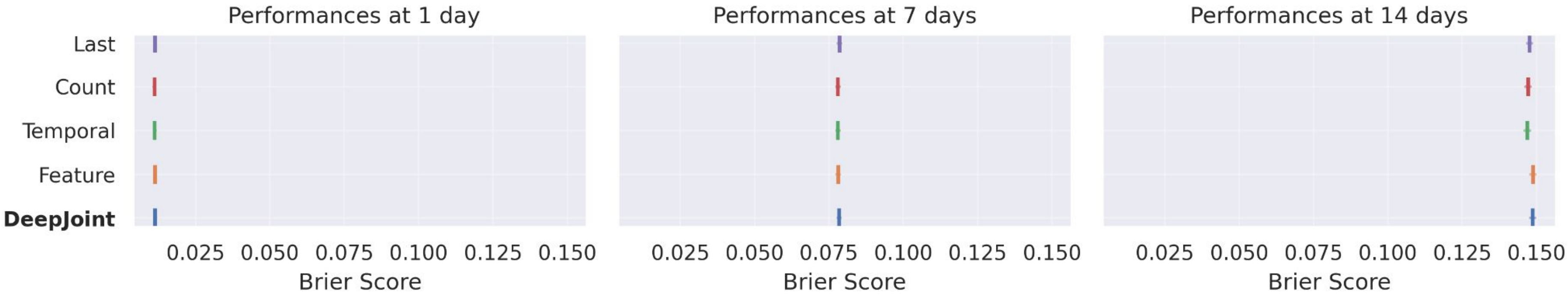
- **DeepJoint** results in an embedding **more robust** to change in observation process

Preprint

Code

Contact
vincent.jeanselme@mrc-bsu.cam.ac.uk
@JeanselmeV

Brier Performance



Weekend Performance