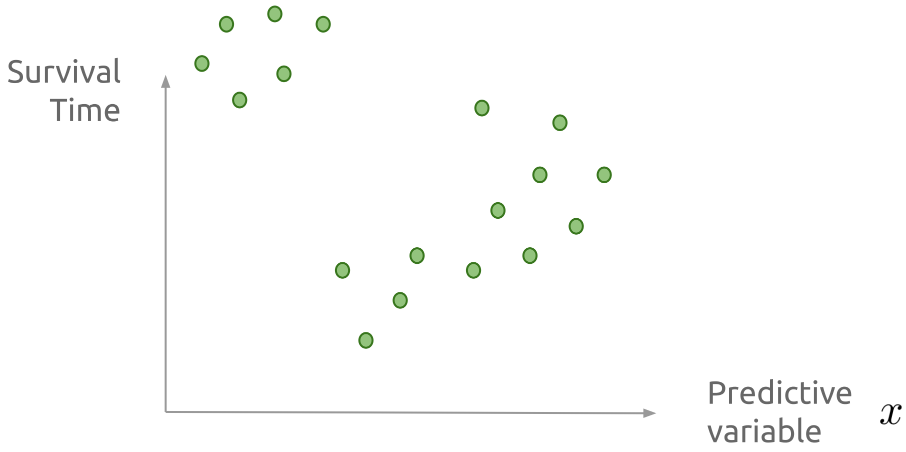
# Neural Fine-Gray

**Vincent Jeanselme**, Chang Ho Yoon,
Brian Tom and Jessica Barrett

CHIL - June 2023

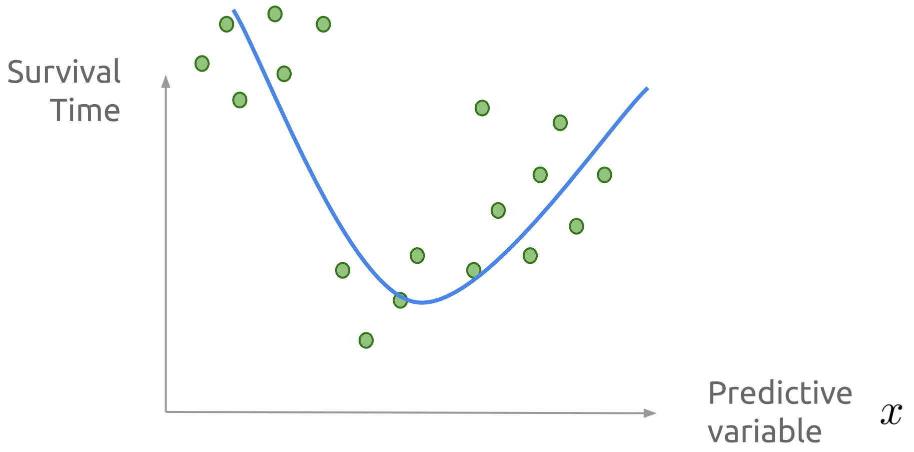
# What is survival analysis ?

**Time-to-event** modelling



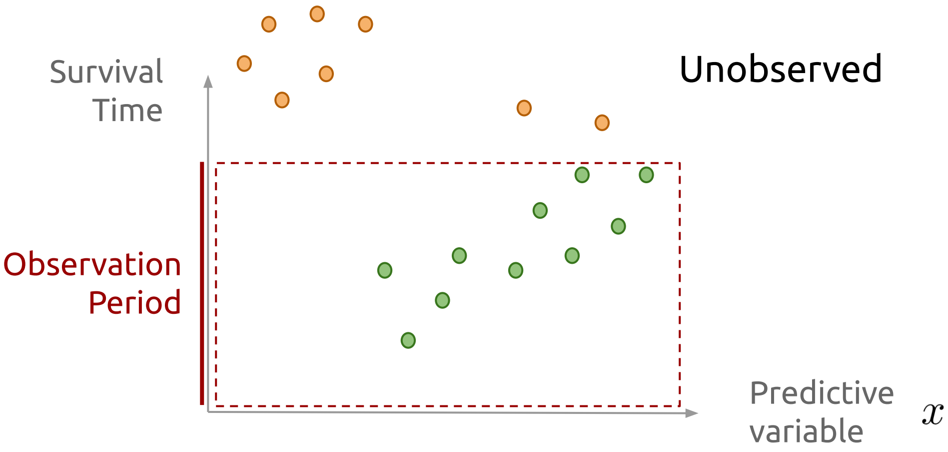
# What is survival analysis ?

**Time-to-event** modelling



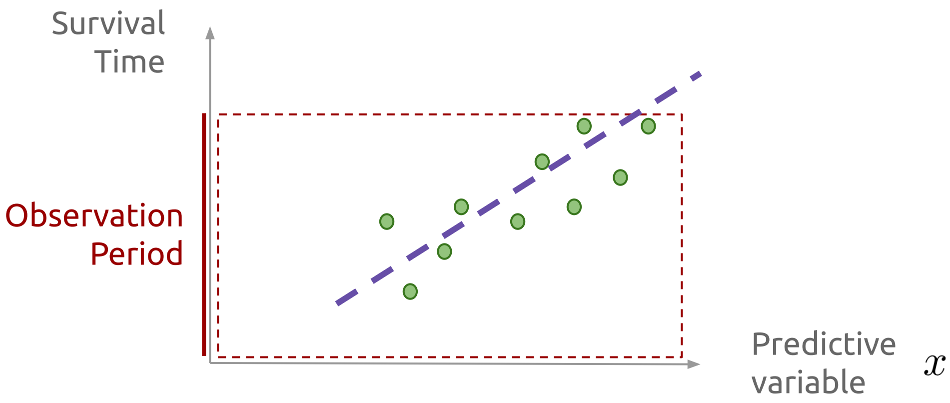
# What is survival analysis ?

**Time-to-event** modelling with **censored** patients



# What is survival analysis ?

**Time-to-event** modelling with **censored** patients



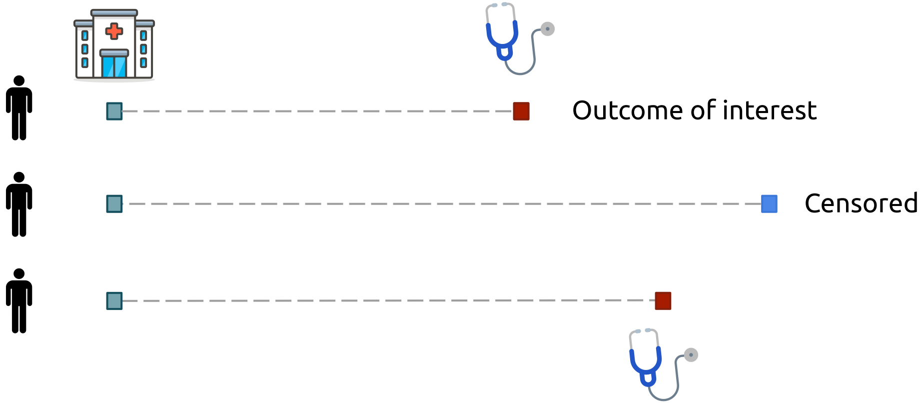
# What is survival analysis ?

Maximise the **likelihood** of **both** observed and censored outcomes

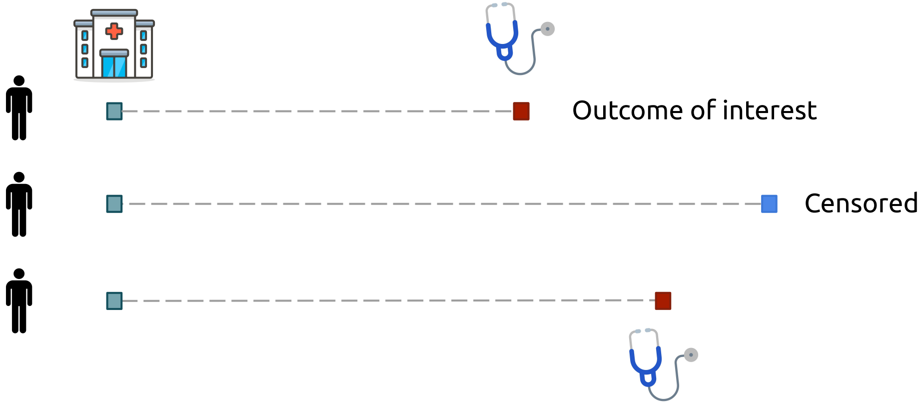
Model the **survival** function

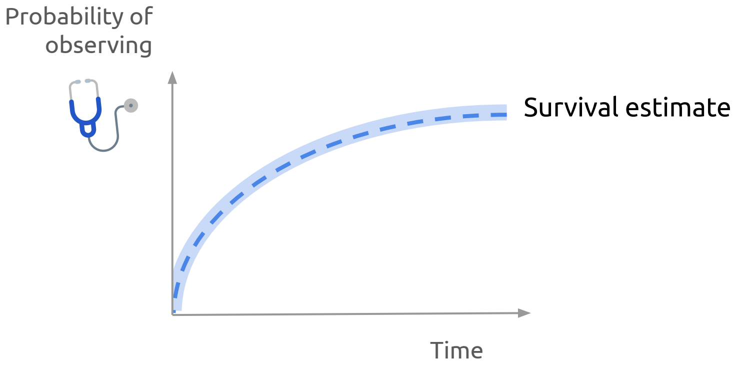
$$S(t \mid x) \;=\; \mathbb{P}(T \geq t \mid x)$$

# Motivation

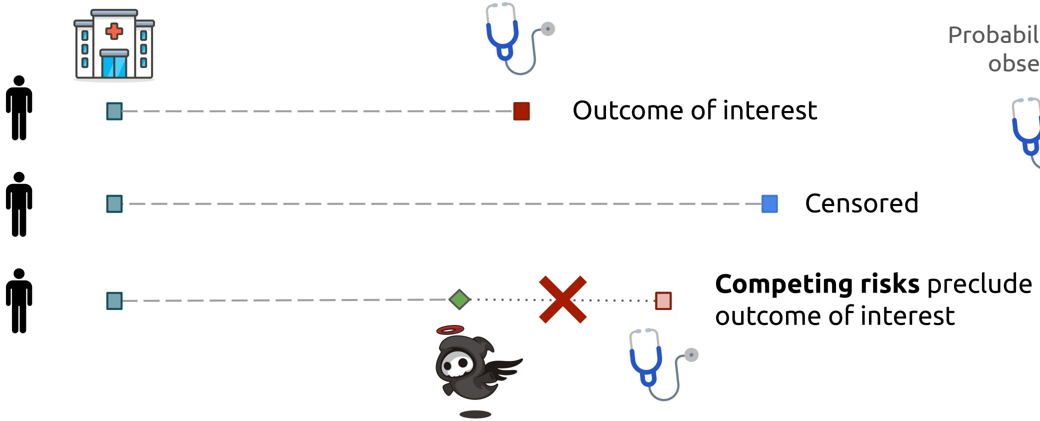


# Motivation

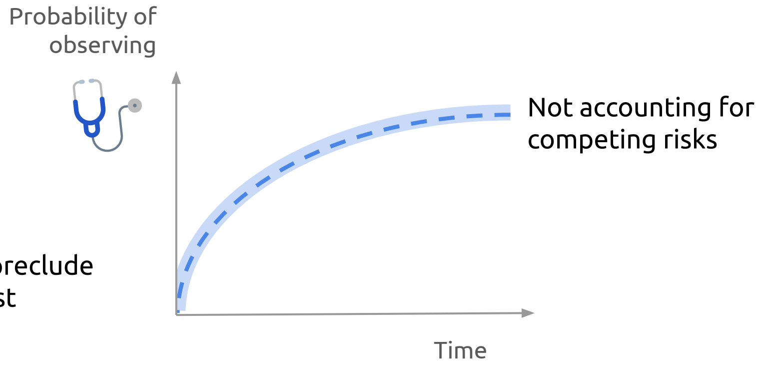

Outcome of interest
Censored


Probability of observing
Survival estimate
Time

# Motivation

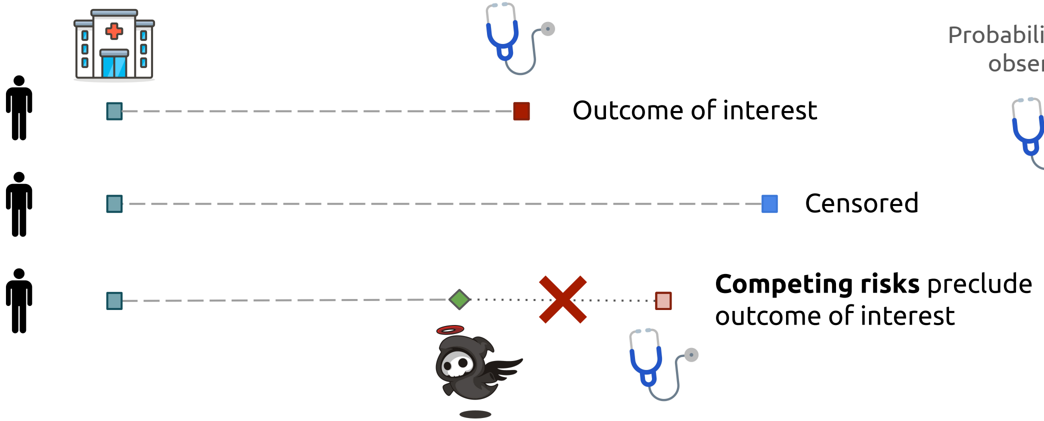


Outcome of interest
Censored
**Competing risks** preclude outcome of interest

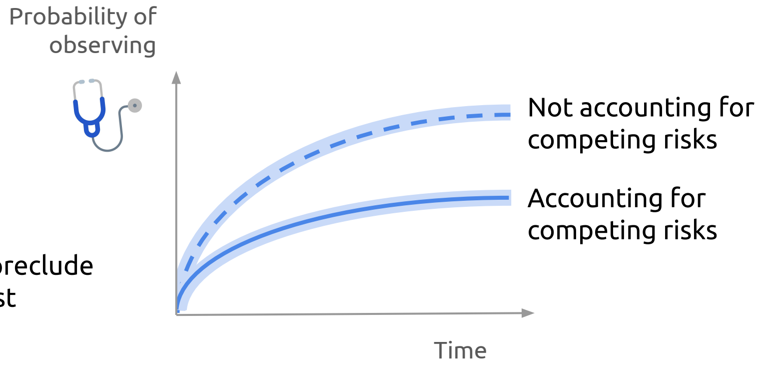



Probability of observing
Not accounting for competing risks
Time

# Motivation



Outcome of interest
Censored
**Competing risks** preclude outcome of interest

Probability of observing
Not accounting for competing risks
Accounting for competing risks
Time

# Existing strategies

1. Cox model
2. Fine-Gray

3. DeepHit
4. Deep Survival Machine
5. DeSurv

# Survival Analysis

$$
\begin{aligned}
S(t \mid x) &= \mathbb{P}(T \geq t \mid x) \\
&= e^{-\underline{\Lambda(t \mid x)}} \\
&= e^{-\int_0^t \underline{\lambda(u \mid x)} du}
\end{aligned}
$$

# Cox Model

Baseline Hazard

$$\lambda(s \mid x) \;=\; \lambda_0(s)\; \exp\left(\alpha^T x\right)$$

Hazard

Covariate Drift

# Competing risks

With multiple risks, one is interested in estimating the **cumulative incidence function**:

$$\begin{aligned} F_r(t|x) &= \mathbb{P}(T < t, \text{risk} = r|x) \\ &= \int_0^t \lambda_r(u|x) e^{-\int_0^u \sum_r \lambda_r(s) ds} du \end{aligned}$$

# Fine-Gray

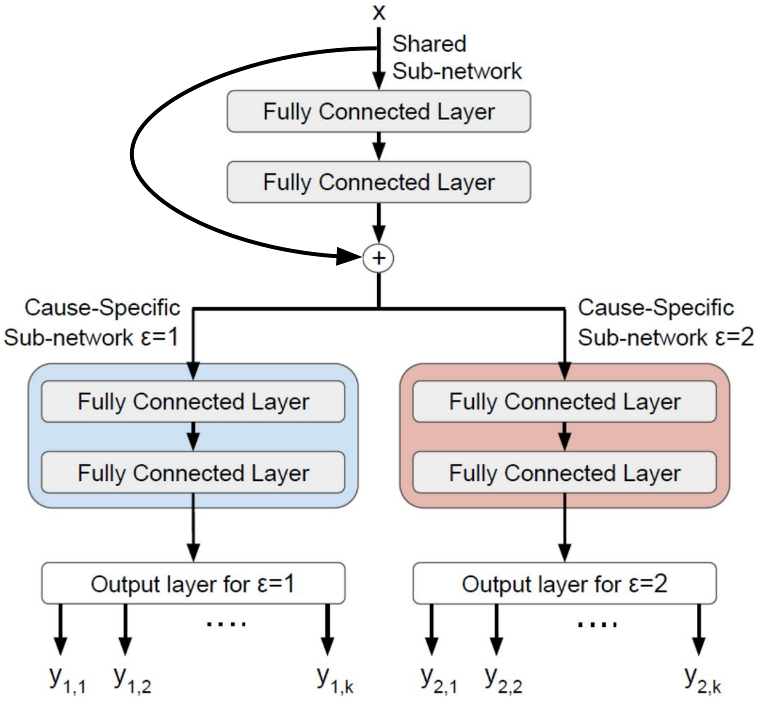
Instead of considering each risk **separately** to then estimate the cumulative incidence function, Fine-Gray proposes to **account for** the different risks by modelling the subdistribution hazard:

$$h_r(t|x) = \lim_{\delta t \to 0} \frac{\mathbb{P}(t < T < t + \delta t,\ \text{risk} = r | (T \geq t)\ \cup\ (T < t\ \cap\ \text{risk} \neq r),\ x)}{\delta t}$$

**A proportional hazards model for the subdistribution of a competing risk** by J. Fine and R. Gray, 1999.

# DeepHit

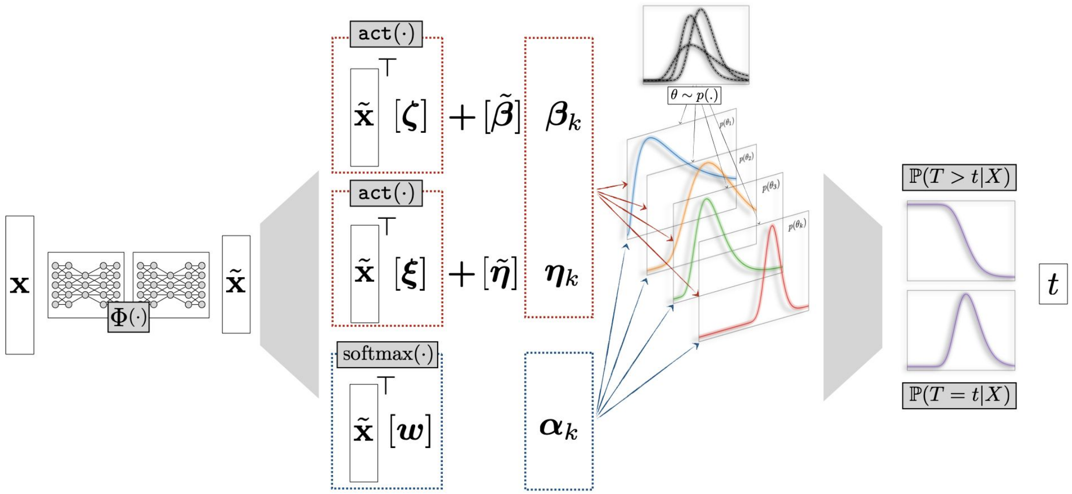
Time **discretisation** with softmax over time and risks




**Deephit: A deep learning approach to survival analysis with competing risks** by C. Lee, 2018

# Deep Survival Machine

Mixture of **parametric distributions** parameterized by neural networks

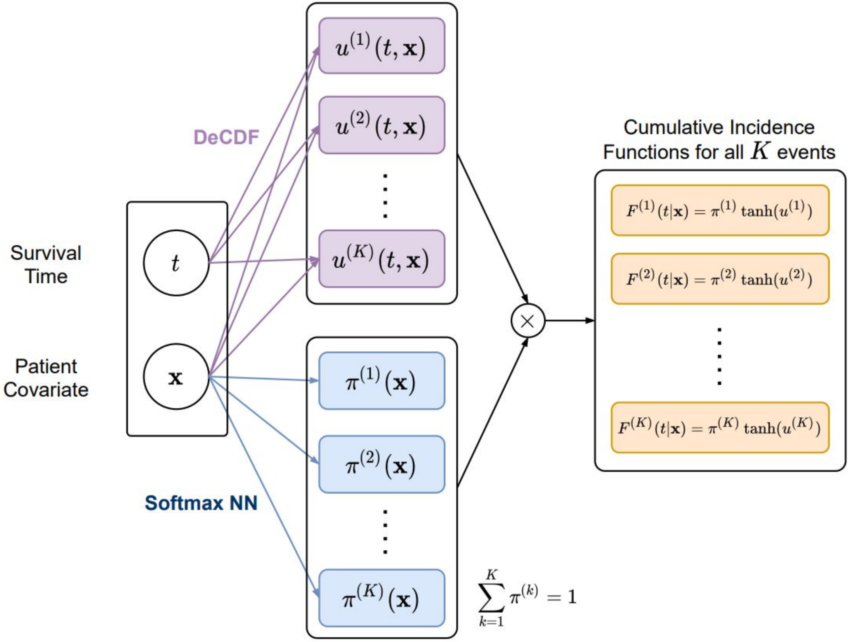



**Deep survival machines: Fully parametric survival regression and representation learning for censored data with competing risks** by C. Nagpal, et al., 2020

# DeSurv

A neural network models the **derivative** of F, and another weigh the outcomes. The cumulative incidence is obtained by solving an **ODE**.

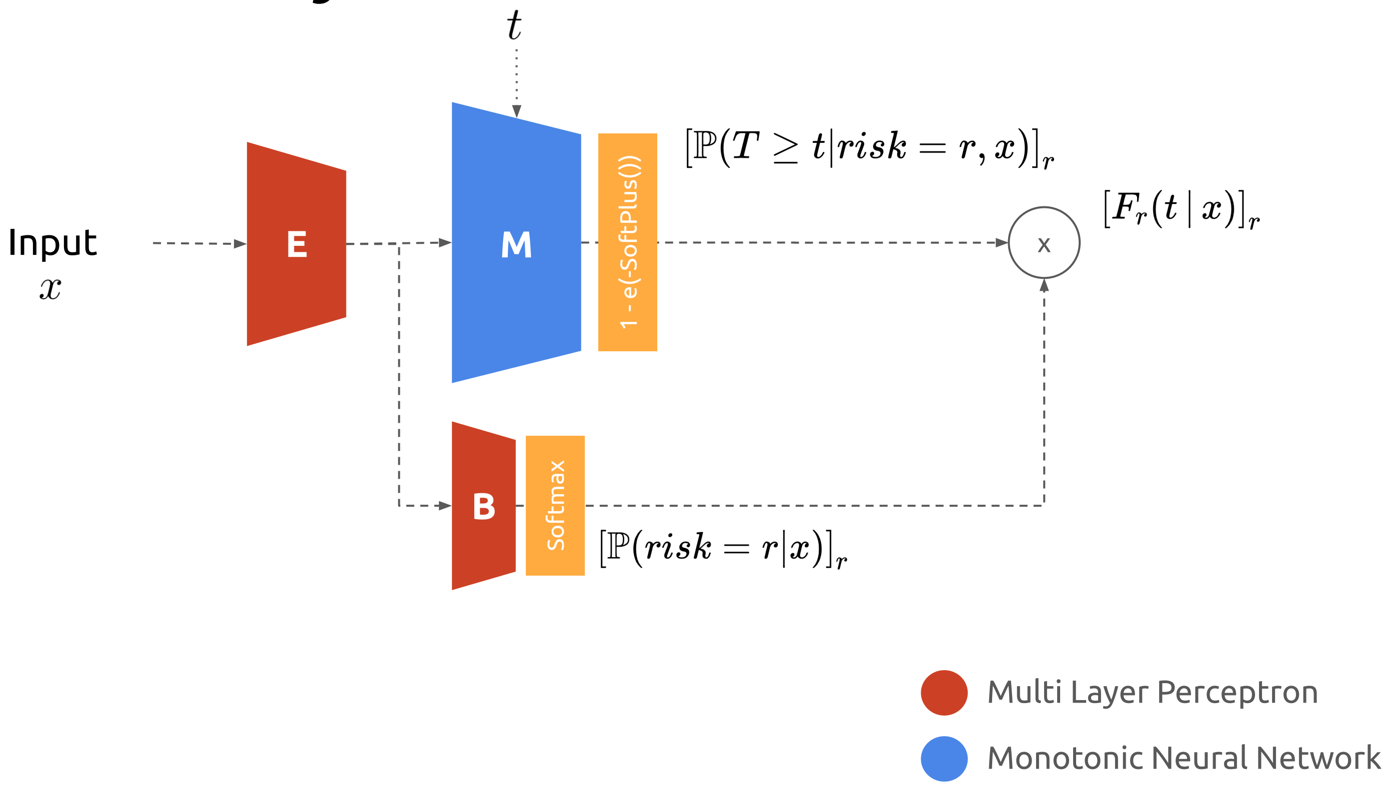
Importantly, one needs **both** the **derivative** of F and **F** to compute the **likelihood** of the observed data.



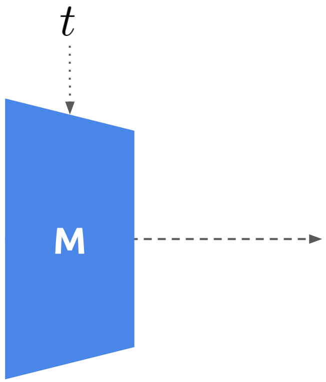

**Derivative-Based Neural Modelling of Cumulative Distribution Functions for Survival Analysis** by D. Danks and C. Yau, 2022

# Proposed strategy

# Neural Fine-Gray

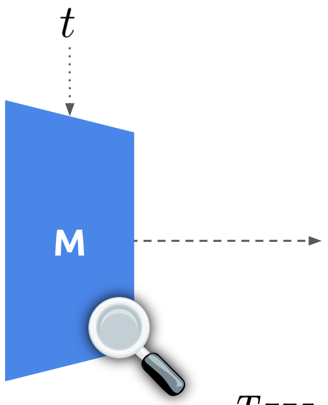


# Monotonic Network




$$\forall d \in \mathbb{R}^+, \mathrm{M}(t + d) \geq \mathrm{M}(t)$$

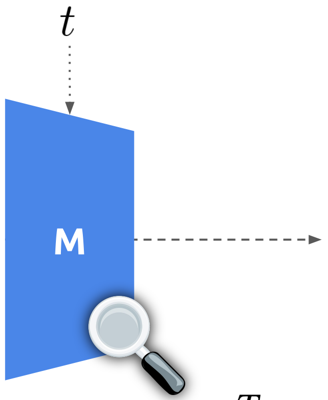
# Monotonic Network




$$\forall d \in \mathbb{R}^+, \mathrm{M}(t + d) \geq \mathrm{M}(t)$$

$$t^T W + b$$

# Monotonic Network




$$\forall d \in \mathbb{R}^+, \mathrm{M}(t + d) \geq \mathrm{M}(t)$$

$$t^T W + b$$

$$W = w^2 > 0$$

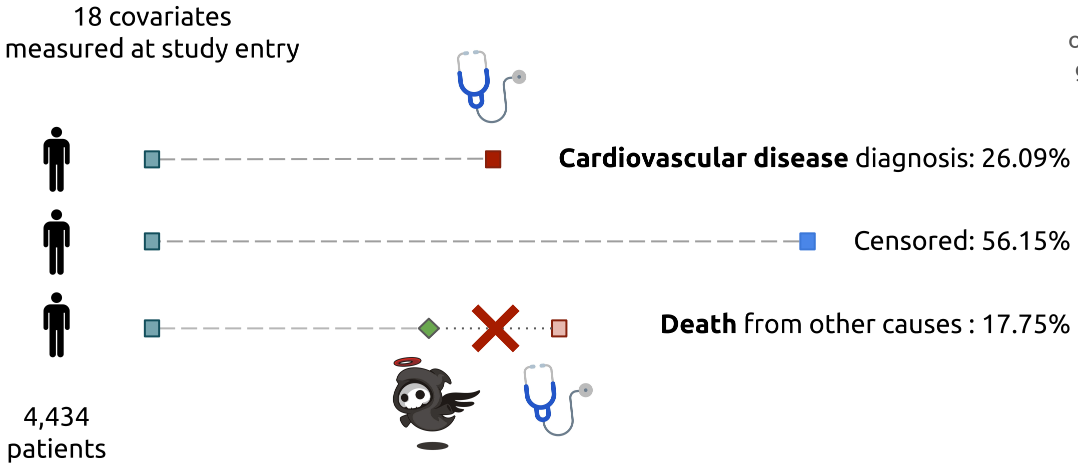
Positively weighted neural networks are **universal monotonic approximators**.

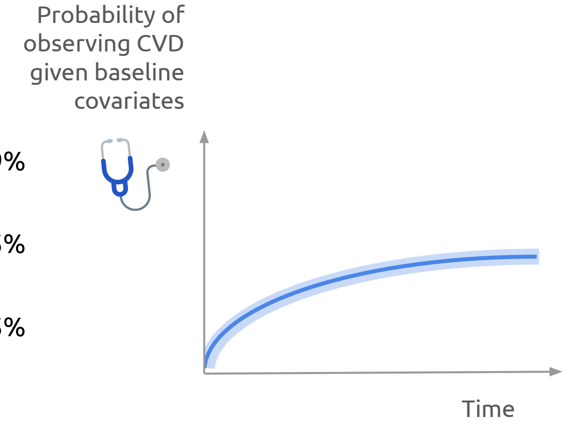
**Monotonic networks** by J. Sill, 1997 & **Monotonic Multi-layer Perceptron Networks as Universal Approximators** by Lang, 2005

# Cardiovascular risk

1. Experimental settings
2. Results
3. Importance of modelling competing risks

# Experimental settings




Evaluation: 5-fold cross-validation

- Time-dependent C Index
- Time-dependent Brier Score

# Results

| Model | C-index *(Discrimination)* | | | Brier Score *(Calibration)* | | |
|---|---|---|---|---|---|---|
| | $q_{0.25}$ | $q_{0.5}$ | $q_{0.75}$ | $q_{0.25}$ | $q_{0.5}$ | $q_{0.75}$ |
| **Neural Fine Gray** | **0.872 (0.024)** | **0.812 (0.029)** | **0.782 (0.018)** | 0.050 (0.003) | **0.095 (0.010)** | **0.128 (0.004)** |
| DeepHit | 0.855 (0.026) | 0.781 (0.026) | 0.743 (0.014) | 0.053 (0.003) | 0.102 (0.007) | 0.141 (0.002) |
| DSM | 0.866 (0.023) | 0.806 (0.023) | 0.778 (0.014) | 0.057 (0.005) | 0.104 (0.006) | 0.141 (0.002) |
| DeSurv | **0.872 (0.027)** | 0.807 (0.031) | 0.775 (0.022) | **0.049 (0.005)** | **0.095 (0.009)** | 0.129 (0.003) |
| Fine Gray | 0.842 (0.025) | 0.794 (0.024) | 0.772 (0.015) | 0.057 (0.006) | 0.099 (0.007) | 0.131 (0.003) |

Time quantiles

In addition to **state-of-the-art** performance,
the proposed method offers a ***n / 2*** **computational gain** in comparison to DeSurv
*n* being the number of points used for the numerical integration (n = 15).

# Importance

## Why model competing risks?

| Model | C-index | | | Brier Score | | |
|---|---|---|---|---|---|---|
| | $q_{0.25}$ | $q_{0.5}$ | $q_{0.75}$ | $q_{0.25}$ | $q_{0.5}$ | $q_{0.75}$ |
| **Neural Fine Gray** | **0.872 (0.024)** | **0.812 (0.029)** | **0.782 (0.018)** | **0.050 (0.003)** | **0.095 (0.010)** | **0.128 (0.004)** |
| Non - Competing | 0.862 (0.029) | 0.807 (0.032) | 0.780 (0.020) | 0.053 (0.004) | 0.099 (0.011) | 0.129 (0.005) |

Accounting for competing risks **improves** risk predictions.

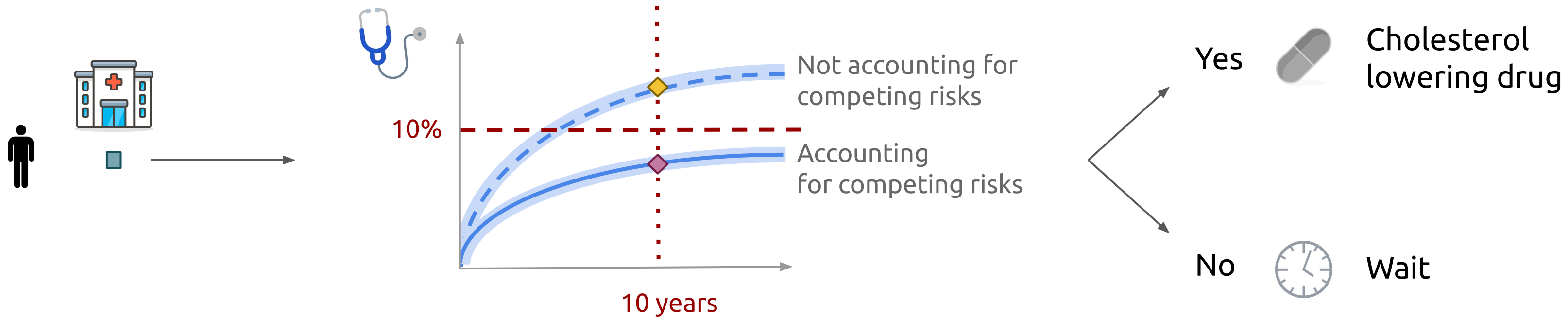
# Importance

## Who benefits?

| Age groups | Difference in Brier Score $q_{0.25}$ | $q_{0.5}$ | $q_{0.75}$ |
|---|---|---|---|
| < 40 | -0.000 (0.000) | -0.001 (0.002) | 0.000 (0.005) |
| 40 - 50 | -0.001 (0.001) | -0.002 (0.003) | -0.002 (0.001) |
| 50 - 60 | -0.003 (0.005) | -0.004 (0.003) | -0.006 (0.007) |
| **60 +** | **-0.013 (0.011)** | **-0.022 (0.018)** | **-0.007 (0.024)** |

Patients the **most at risk** for the competing risk may benefit the most.

# Impact on medical practice

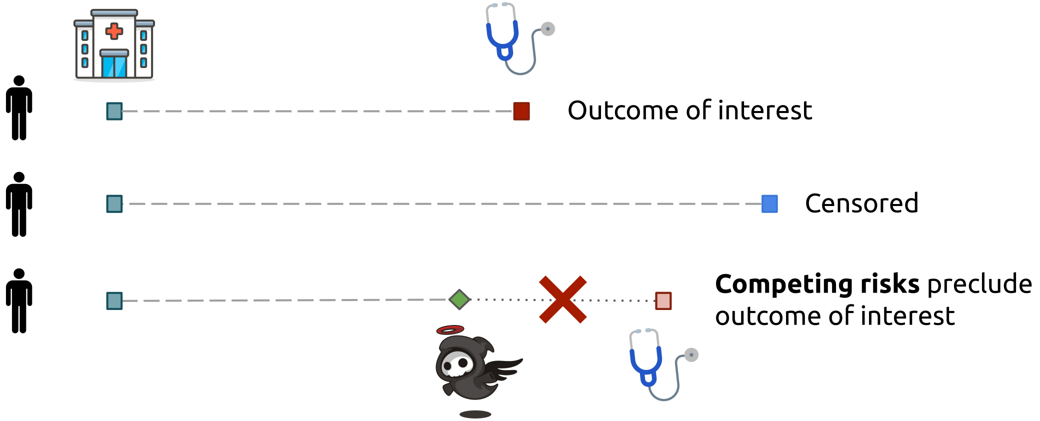



Has the patient a risk > 10% in the
10 next years ?

# Conclusion

Code

Paper




**Competing risks *must* be accounted for in medical analyses**


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