Principled Estimation and Prediction with Competing Risks: a Bayesian Nonparametric Approach

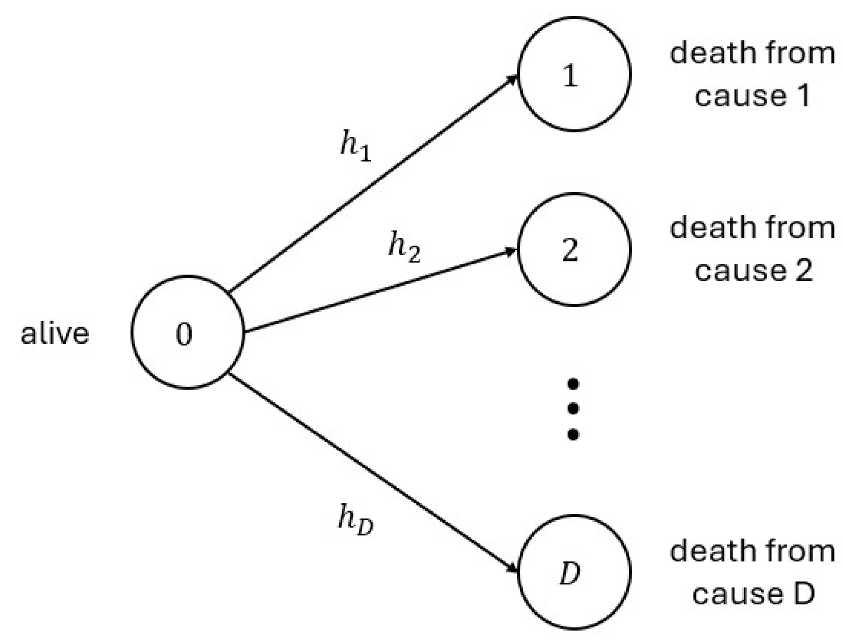
Claudio Del Sole [1] Antonio Lijoi [2] Igor Prünster [2]
[1]University of Milano–Bicocca, Milan, Italy [2]Bocconi University, Milan, Italy,
claudio.delsole@unimib.it

Competing risks in survival analysis

In **survival analysis**, researchers may be interested in different types of events (sources of risk), which are **competing events** if the occurrence of an event prevents the occurrence of other events [4].

Multi–state approach to competing risks

Competing risks data are modelled through multi–state models with a transient state (**alive**) and multiple absorbing states (**causes of death**):



- the **time–to–event** $T \in \mathbb{R}^+$ is the **time of transition** away from state 0;
- the **cause of death** $\Delta \in \{1, \dots, D\}$ is the target **absorbing state**;
- the cause–specific **hazard rates** $h_1(t), \dots, h_D(t)$ are the **transition rates**.

Functionals of interest in competing risks

The main quantities of interest in a competing risks framework are:

- the **survival function**, i.e. the probability of surviving every competing event up to a certain time,
$$S(t) = \mathbb{P}(T \geq t) = \exp\left(-\sum_{\delta=1}^{D}\int_0^t h_\delta(u)\,du\right);$$
- the cause–specific **cumulative incidence functions**, i.e. probabilities of experiencing a certain type of event within a certain time;
- **(prediction viewpoint)** the probabilities of experiencing a certain type of event, given the survival time, termed prediction curves,
$$\pi_\delta(t) = \mathbb{P}(\Delta = \delta \mid T = t), \qquad \delta = 1, \dots, D.$$

Modeling mixture hazard rates

In a Bayesian setting, a prior is defined over **hazard rate functions** [3], i.e. the instantaneous risks of occurrence of each competing event, given survival up to that time:
$$\tilde{h}_\delta(t) = \int_{\mathbb{R}^+} k(t;x)\,\mathrm{d}\tilde{\mu}_\delta(x), \qquad \delta = 1, \dots, D,$$
where $k(t,x)$ is a deterministic kernel and $\tilde{\mu}_1, \dots, \tilde{\mu}_D$ are increasing random processes.

The model for an **exchangeable sequence** of time-to-event and event type pairs is
$$(T_1, \Delta_1), \dots, (T_n, \Delta_n) \mid \tilde{\boldsymbol{\mu}} \overset{\text{i.i.d.}}{\sim} \tilde{p}, \qquad \tilde{\boldsymbol{\mu}} = (\tilde{\mu}_1, \dots, \tilde{\mu}_D) \sim \mathcal{Q},$$
where the directing random probability measure $\tilde{p}$ depends on random processes through hazard rates:
$$\tilde{p}(dt, \delta) = \underbrace{\int_{\mathbb{X}} k(t;x)\,\mathrm{d}\tilde{\mu}_\delta(x)}_{\text{hazard rate for cause } \delta}\ \exp\left(-\sum_{\ell=1}^{D}\int_0^t \underbrace{\int_{\mathbb{X}} k(s;x)\,\mathrm{d}\tilde{\mu}_\ell(x)}_{\text{hazard rate for cause } \ell}\,ds\right)dt.$$

Hierarchical prior specification

The prior specification $\mathcal{Q}$ introduces dependence among hazard rates through a **hierarchical structure** of completely random measures (increasing Lévy processes) [2]:
$$\tilde{\mu}_1, \dots, \tilde{\mu}_D \mid \tilde{\mu}_0 \overset{\text{i.i.d.}}{\sim} \text{CRM}(\tilde{\nu}), \qquad \tilde{\mu}_0 \sim \text{CRM}(\nu_0),$$
having homogeneous **Lévy intensities**
$$\mathrm{d}\tilde{\nu}(s,x) = \rho(s)\,\mathrm{d}\tilde{\mu}_0(x), \qquad \mathrm{d}\nu_0(s,x) = \rho_0(s)\,\mathrm{d}P_0(x).$$
A natural choice for hierarchical CRMs is the **hierarchical gamma** process.

Latent variables and partition structure

The marginal, predictive and posterior distributions are conveniently described via the introduction of two sequences of **latent variables**:
$$\boldsymbol{X} = (X_1, \dots, X_n), \qquad \boldsymbol{Z} = (Z_1, \dots, Z_n).$$
Because of the discreteness of CRMs, variables in each sequence admit **ties** with positive probability → **nested partition structure** (Chinese restaurant franchise metaphor [5]).

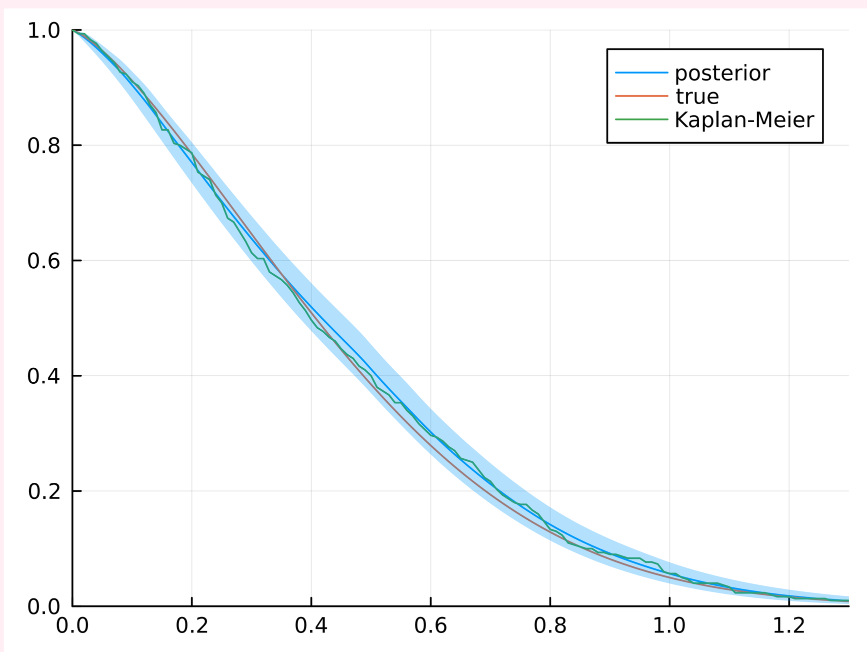
Posterior characterization

The posterior distribution of random processes, given observations and latent variables, is **structurally conjugate**, as the hierarchical form is preserved a posteriori:
$$\tilde{\mu}_\delta(x) \mid (\boldsymbol{T}, \boldsymbol{\Delta}, \boldsymbol{X}, \boldsymbol{Z}),\ \tilde{\mu}_0 \sim \tilde{\mu}^*(x) + \sum_{j=1}^{k} J_{\delta j}\,(x \geq X_j^*),$$
$$\tilde{\mu}_0(x) \mid (\boldsymbol{T}, \boldsymbol{\Delta}, \boldsymbol{X}, \boldsymbol{Z}) \sim \tilde{\mu}_0^*(x) + \sum_{j=1}^{k} I_j\,(x \geq X_j^*),$$
where $\tilde{\mu}^*$ and $\tilde{\mu}_0^*$ are CRMs with **non–homogeneous** Lévy intensities, while $J_{\delta j}$'s and I_j's are **independent** random variables.

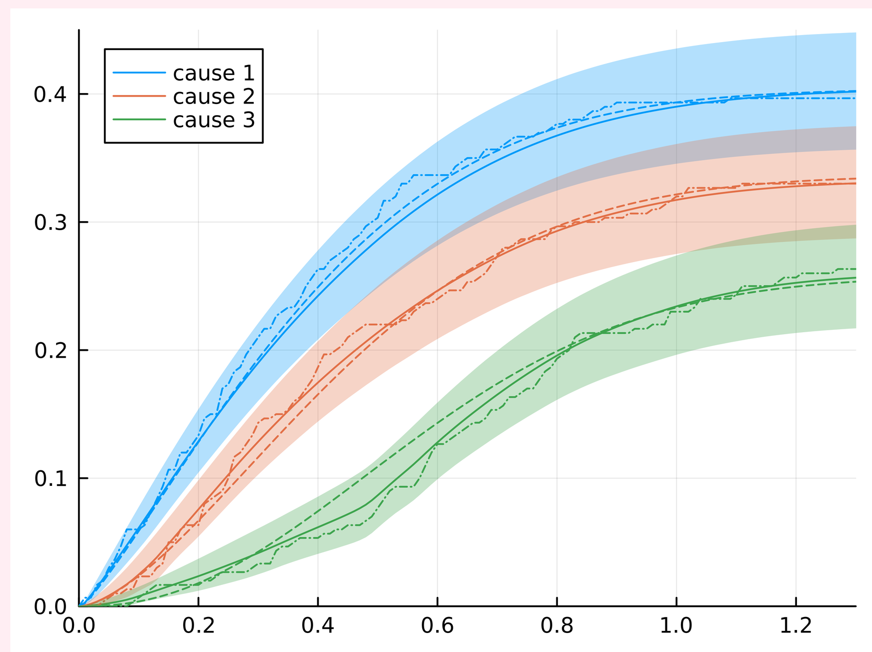
Numerical illustration on simulated data

Consider three independent competing risks and record the minimum time–to–event and the corresponding event type, for $n = 300$ observations.

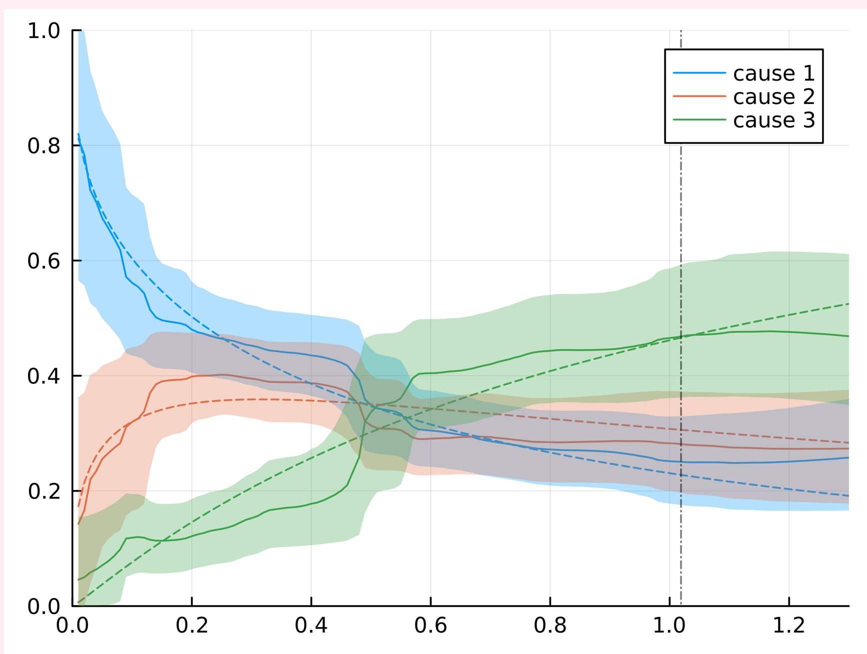
Survival function



Cumulative incidence functions



Prediction curves



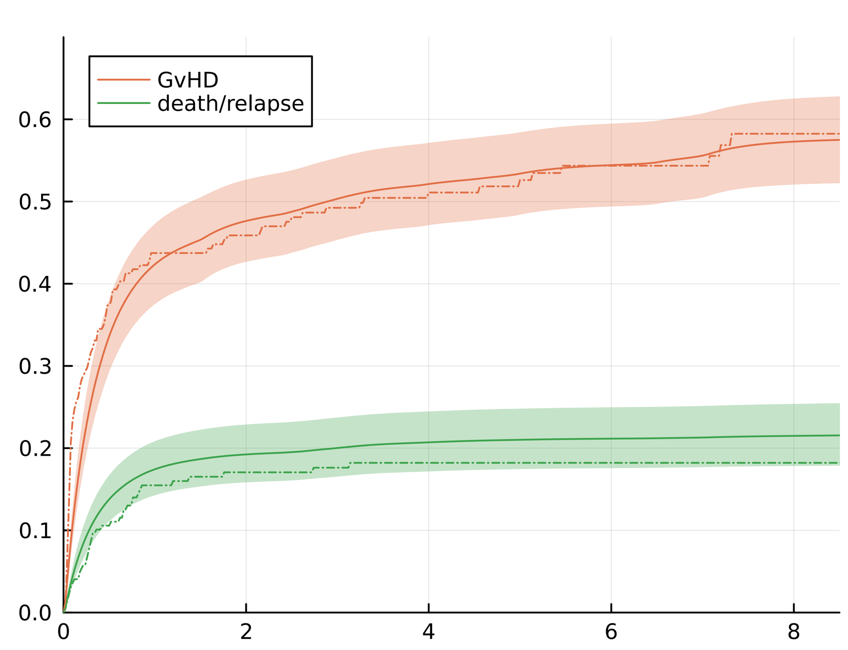
- **Full conditional distributions** of latent variables $(\boldsymbol{X}, \boldsymbol{Z})$ are derived from the marginal distribution, and exploited to devise a **Gibbs sampling scheme**.
- **Posterior estimates** of quantities of interest are obtained at each step, conditionally on latent variables.

Application to bone marrow transplant data

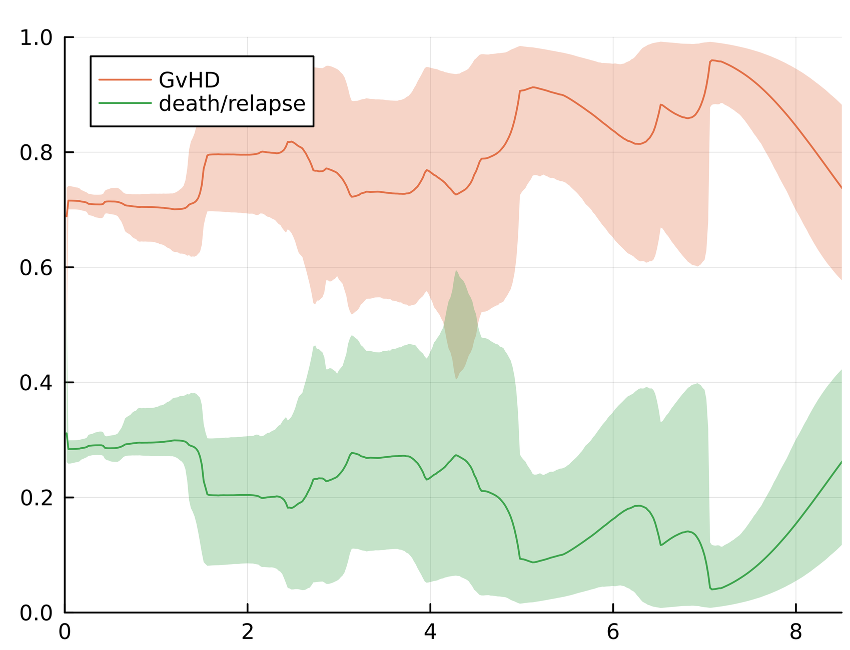
The dataset includes data for 400 patients diagnosed with acute **myeloid leukemia**, who underwent a bone marrow transplantation:

- the primary event of interest is occurrence of **Graft–versus–Host–Disease** (GvHD);
- **death or relapse** without GvHD are competing events.

Cumulative incidence functions



Prediction curves



Results are displayed for patients receiving graft from bone marrow cells.

[1] Del Sole, C., Lijoi A. and Prünster I. (2025+). *Principled estimation and prediction with competing risks: a Bayesian nonparametric approach.* Submitted.

[2] Camerlenghi, F., Lijoi A. and Prünster I. (2021). *Survival analysis via hierarchically dependent mixture hazards.* The Annals of Statistics 49(2), 863–884.

[3] Dykstra, R. L. and P. Laud (1981). *A Bayesian nonparametric approach to reliability.* The Annals of Statistics 9 (2), 356–367.

[4] Geskus, R. B. (2024). *Competing risks: Concepts, methods, and software.* Annual Review of Statistics and Its Application 11(1), 227–254.

[5] Teh, Y.M., Jordan, M.I., Beal, M.J. and Blei, D.M. (2006). *Hierarchical Dirichlet processes.* Journal of the American Statistical Association 101(476), 1566–1581.