A Gaussian processes model for survival analysis with time dependent covariates and interval censoring

Aki Vehtari[1] and Heikki Joensuu[2]

[1]Department of Biomedical Engineering and Computational Science (BECS), Aalto University, Finland

[2]Department of Oncology, Helsinki University Central Hospital, Finland

Introduction

- Cancer patients are frequently followed up longitudinally with computed tomography (CT), which involves radiation hazards, and the optimal follow-up schedules are often unknown
- We developed a method that adjusts the timing of CT scans with the hazard of cancer recurrence in time

Material

- The clinical data generated in the phase III Scandinavian Sarcoma Group XVIII/ Arbeitsgemeinschaft Internistische Onkologie trial that compares 1-year to 3-year duration of adjuvant imatinib (medicine taken orally) in the treatment of patients with gastrointestinal stromal tumour (GIST)
- For the individual i, where $i = 1, \ldots, n$, we have survival time y_i with a censoring indicator δ_i, where $\delta_i = 0$ if the ith observation is uncensored and $\delta_i = 1$ if the observation is right or interval censored. For interval censored survival time, y_i is known to fall into an interval $[y_{i,\mathrm{lo}}, y_{i,\mathrm{up}}]$.
- Three time-associated covariates were the time from the date of randomisation, the time from the date of completion of adjuvant imatinib (considered to be 0 before completion of adjuvant therapy), and an indicator variable for adjuvant imatinib (considered to be 1 before completion of adjuvant imatinib, and 0 after its completion)
- Two covariates observed at the time of the study entry were tumour mitotic count and tumour location (gastric vs. non-gastric).

Methods

- We present a Gaussian processes model for inhomogoneus Poisson process survival analysis with interval censored data.
- The benefit of Gaussian process model is that time dependent and baseline covariates may have full interactions.
- To allow the form of the hazard function to depend on the covariates, we use a generic hazard model

$$h_i(t) = \exp(\eta(t, \boldsymbol{x}_i)), \tag{1}$$

  $\boldsymbol{x}_i$ is the $d \times 1$ vector of covariates for the ith patient, and η is nonlinear function depending on time t and the covariates $\boldsymbol{x}_i$.
- We assume a piecewise log-constant hazard in time. The time axis is partitioned into K non-overlapping intervals with equal lengths: $0 = s_0 < s_1 < s_2 < \ldots < s_K$, where $s_K \geq y_i$ for all $i = 1, \ldots, n$. In the interval k, where $k = 1, \ldots, K$, hazard is assumed to be constant in time and for the ith individual the hazard rate in the kth time interval is

$$h_i(t) = \exp(\eta(\tau_k, \boldsymbol{x}_{ik})),\ t \in (s_{k-1}, s_k], \tag{2}$$

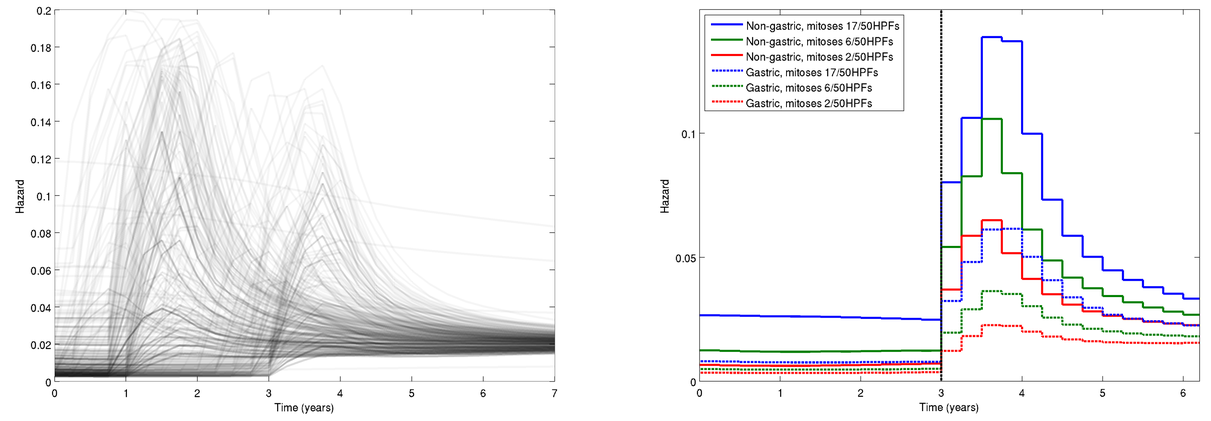
  where $\tau_k = (s_k - s_{k-1})/2$ is the mean of kth time interval and $\boldsymbol{x}_{ik}$ denotes possibly time varying covariates.
- Using the piecewise log-constant assumption for the hazard rate function, the contribution of the possibly right censored ith observation (y_i, δ_i) for the likelihood is

$$l_i = [\exp(\eta_{iK_i})]^{(1-\delta_i)} \exp\left(-\sum_{k=1}^{K_i}(s_k - s_{k-1})\exp(\eta_{ik})\right), \tag{3}$$

  where $\eta_{ik} = \eta(\tau_k, \boldsymbol{x}_{ik})$ and K_i is such that $s_{K_i-1} < y_i \leq s_{K_i}$.
- This likelihood is equivalent to the likelihood of K_i Poisson distributed data points, with means $(s_k - s_{k-1})\exp(\eta_{ik})$, of which $K_i - 1$ first ones are observed to be 0, and the last one observed to be 0 or 1 according to whether the survival time y_i is observed or censored.
- A zero-mean Gaussian process (GP) prior is set for η. The covariance function of GP defines the smoothness and scale properties of the latent function. We use a sum of constant and non-stationary neural network covariance functions. A neural network covariance function was chosen, since it is suitable for modeling non-stationary, saturating and interaction effects.
- Weakly informative priors were chosen for covariance function parameters.
- The conditional distribution $p(\theta|\boldsymbol{y}, \boldsymbol{x})$ was obtained by marginalising over the latent variables $\boldsymbol{\eta}$ using a Gaussian approximation.
- The covariance function parameters were sampled from the conditional distribution $p(\theta|\boldsymbol{y}, \boldsymbol{x})$ using hyperrectangle multivariate slice sampling.
- The unknown y_i for interval censored observations were sampled by first sampling latent values η_{ik} from the Gaussian approximation of the conditional distribution $p(\eta|\theta, \boldsymbol{y}, \boldsymbol{x})$, then computing the hazard h given the latent values, and finally sampling y_i from the conditional density $p(y_i|h_i)$ at interval $[y_{i,\mathrm{lo}}, y_{i,\mathrm{up}}]$.
- Inference for the model was made using GPstuff toolbox

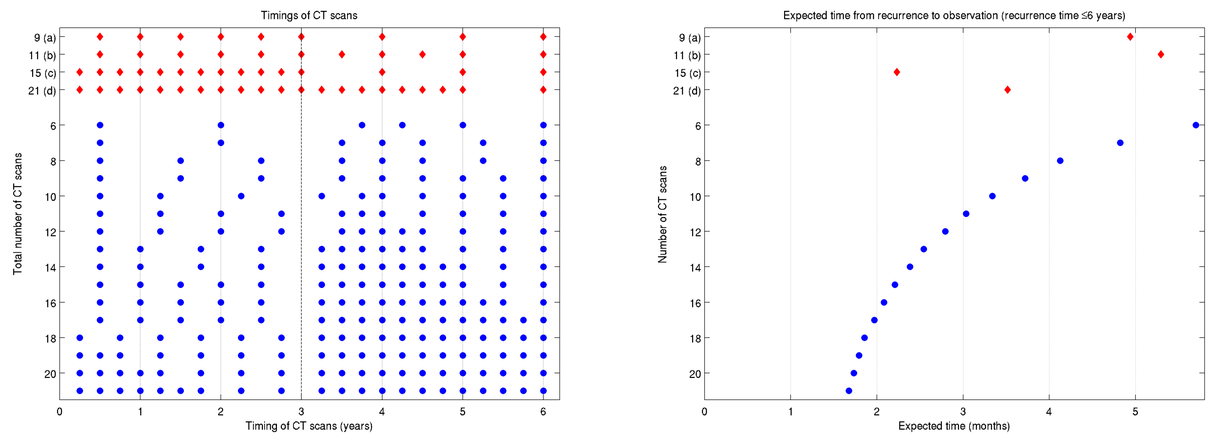
Results

- The model was used to find the optimal points in time for performing each CT scan by minimizing the expected time between the observed date of tumour recurrence and the model-predicted date of recurrence for six prototype patients treated for 36 months with adjuvant imatinib
- Shape of the individual hazard function depends on the surgery time covariates and time dependent covariates



Left: Individual patient hazard functions for GIST recurrence. Right: The hazard of GIST recurrence with time for six prototype patients treated with three years of adjuvant imatinib based on tumour site and mitotic count. The hazard is low during adjuvant imatinib, but higher after completion of adjuvant treatment.

- Predicted hazard functions for new patients can be used to optimize timing of computed tomography examinations to reduce the time from recurrence to observation without increasing the radiation dose and imaging costs



Left: Timing of CT scans according to the National Comprehensive Cancer Network of the US guideline recommendations5 (red diamonds, CT carried out at 6 or 3 month intervals for 3 or 5 years) and optimised timing of CT scans according to the model (blue circles). Right: The estimated time intervals between the date of GIST recurrence and its detection by CT with different imaging schedules.

Conclusion

- The numbers of CT scans can be reduced approximately 30% during the first six years of follow-up since initiation of treatment compared with the current follow-up recommendations without jeopardizing early detection of recurrence
- The method may be applicable to the follow-up of other types of human cancer to facilitate early detection of recurrence or to reduce the radiation hazards associated with CT scans

References

- Joensuu, H., Reichardt, P., Eriksson, M., Hall, K.S., Vehtari, A. (2013). Timing of computed tomography scans based on the hazard of recurrence in the follow-up of patients with gastrointestinal stromal tumour treated with adjuvant imatinib. Submitted.
- Joensuu, H., Vehtari, A., Riihimäki, J., Nishida, T., Steigen, S.E., Brabec, P., Plank, L., Nilsson, B., Cirilli, C., Braconi, C., Bordoni, A., Magnusson, M.K., Linke, Z., Sufliarsky, J., Massimo, F., Jonasson, J.G., Dei Tos, A.P. and Rutkowski, P. (2012). Risk of gastrointestinal stromal tumour recurrence after surgery: an analysis of pooled population-based cohorts. The Lancet Oncology, 13(3):265-274.
- Vanhatalo, J., Riihimäki, J., Hartikainen, J., Jylänki, P., Tolvanen, V. and Vehtari, A. (2013). GPstuff: Bayesian Modeling with Gaussian Processes. Journal of Machine Learning Research, 14(Apr):1175-1179.