Bayesian sparse survival analysis for detecting subgroup effects: with application to comparing first-line hypertension treatments

Aki Nishimura[1], Martijn J. Schuemie[1,2], Seng Chan You[3], Marc A. Suchard[1]

[1]University of California - Los Angeles, [2]Janssen Research and Development, [3]Ajou University

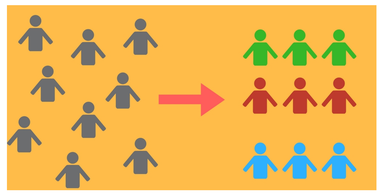
Beyond Population-level Drug Effect

Current OHDSI (and most other) observational studies quantify effectiveness and safety **averaged** over cohorts of patients.

"Average" patient???

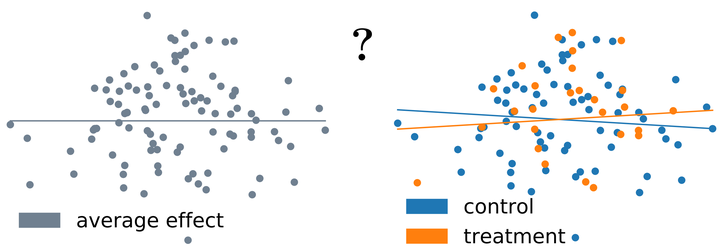
To generate further scientific and clinical insights, we would like to **identify subgroups** who benefit most from particular treatments.

Tailor treatment based on patient characteristics.



Detecting Subgroup via Feature Selection Method

Is the feature x_j relevant?
i.e. does treatment effect vary as a function of x_j?



In a statistics language, identifying subgroups amounts to deciding

$$\gamma_j = 0 \text{ or } \gamma_j \neq 0 \text{ in the model: (hazard rate)} = \sum_{j=1}^{p} (\beta_j + \gamma_j z_j) x_j,$$

where $z_j \in \{0, 1\}$ indicates treatment assignment.

Challenge — high-dimensional features & low incidence:

- the number of covariates p — 10,000+.
- the number of observed events — 100s $\sim$ 1,000s.

Bayesian Sparse Regression / Feature Selection

We look for γ_j's distinguishable from 0 via Bayesian sparse regression:

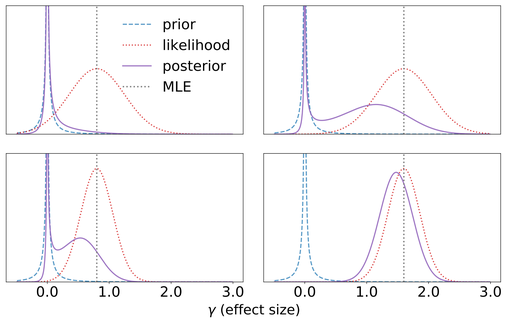
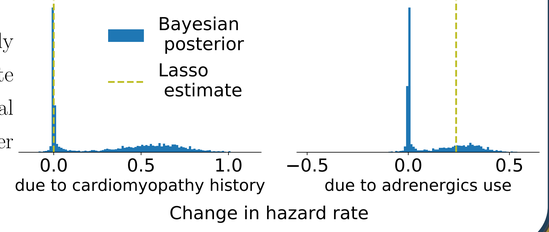


Illustration of the posterior distribution of a coefficient γ_j under different effect sizes and levels of uncertainty in data.

Compare to alternatives, Bayesian method has the advantages of

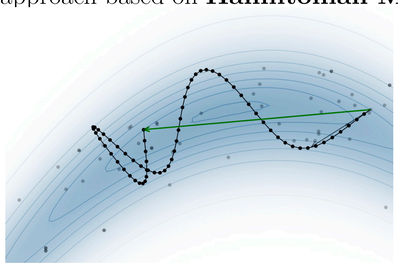
- better separation of the significant coefficients from the rest.
- quantified uncertainty in the estimate and decision $\gamma_j \neq 0$.

Widely-used Lasso only provides a point estimate despite the substantial uncertainty in whether $\gamma_j = 0$ or $\gamma_j \neq 0$.



State-of-the-art Computational Techniques

Bayesian sparse regression had **previously** been **computationally intractable** at the scale of OHDSI studies. We develop a new approach based on **Hamiltonian Monte Carlo** (HMC) algorithm.



Originating from computational physics, HMC exploits the properties of Hamiltonian dynamics to efficiently explore the parameter space.

HMC's performance is sensitive to its tuning parameters; to achieve the algorithm's full potential, we rely on

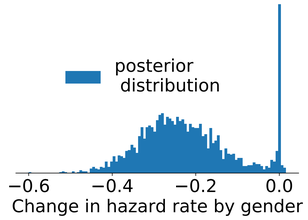
- theory of prior-preconditioning by Nishimura and Suchard [1].
- Lanczos iteration from numerical linear algebra to determine the curvature (largest eigenvalue of Hessian) of the posterior log-density.

Application: ACE inhibitor & thiazide comparison

Goal: Compare effectiveness of the two most common hypertension treatment in preventing major cardiovascular events.

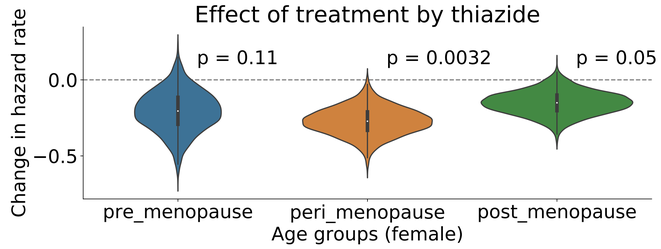
Data: 1,065,745 patients with 7,884 clinical covariates, among whom 5,054 events are observed (0.5% incidence rate).

Result: Bayesian sparse survival analysis identifies gender as a significant source of heterogeneity in the treatment effect:



Regression coefficient estimate for the thiazide-female interaction in the survival model.

Motivated by the above finding and pathophysiology of hypertension among women, we investigated whether the effect varies by age:



Our result suggests that **women in their peri-menopausal stage** benefit most from treatment by thiazide over ACE.

Conclusion & References

Our Bayesian method identifies a statistically significant subgroup effect **among 7,884 possibilities** in the hypertension data. Software is under development to conduct a further study at larger scale.

[1] Nishimura A and Suchard MA (2018). Prior-preconditioned conjugate gradient for accelerated Gibbs sampling in" large n & large p" sparse Bayesian logistic regression models. *arXiv:1810.12437*.

[2] Nishimura A, Schuemie MJ, and Suchard MA (2019+). Scalable Bayesian sparse generalized linear models and survival analysis via curvature-adaptive Hamiltonian Monte Carlo for high-dimensional log-concave distributions. *In preparation*.