Prognostic accuracy measures for survival models

C. Jason Liang and Patrick J. Heagerty

NIH | National Institute of Allergy and Infectious Diseases

June 7, 2022

Prognostic accuracy for binary outcomes

Calibration: how close are predicted risks and observed event rates?

- Plotting predicted risks against observed event rates
- Hosmer-Lemeshow statistic

Discrimination: how well are cases and controls separated?

- Area under the ROC curve (AUC)
- Discrimination slope (DS)

Mayo PBC example

MAYO CLINIC

Search Mayo Clinic

Request an Appointment
Find a Doctor
Find a Job
Give Now

Log in to Profession Account
Translated Content

PATIENT CARE & HEALTH INFO | DEPARTMENTS & CENTERS | RESEARCH | EDUCATION | FOR MEDICAL PROFESSIONALS | PRODUCTS & SERVICES | GIVING TO MAYO CLINIC

The Mayo Natural History Model for Primary Biliary Cirrhosis

In the following model, survival probability of a patient with primary biliary cirrhosis without treatment is estimated based on the following variables. Please enter data in the corresponding boxes.

How old is the patient? 59 (years)

What is the bilirubin? 2.7 (mg/dl)

What is the albumin? 2.5 (g/dl)

What is the prothrombin time? 2.5 (seconds)

Does the patient have peripheral edema? ○ No ◉ Yes

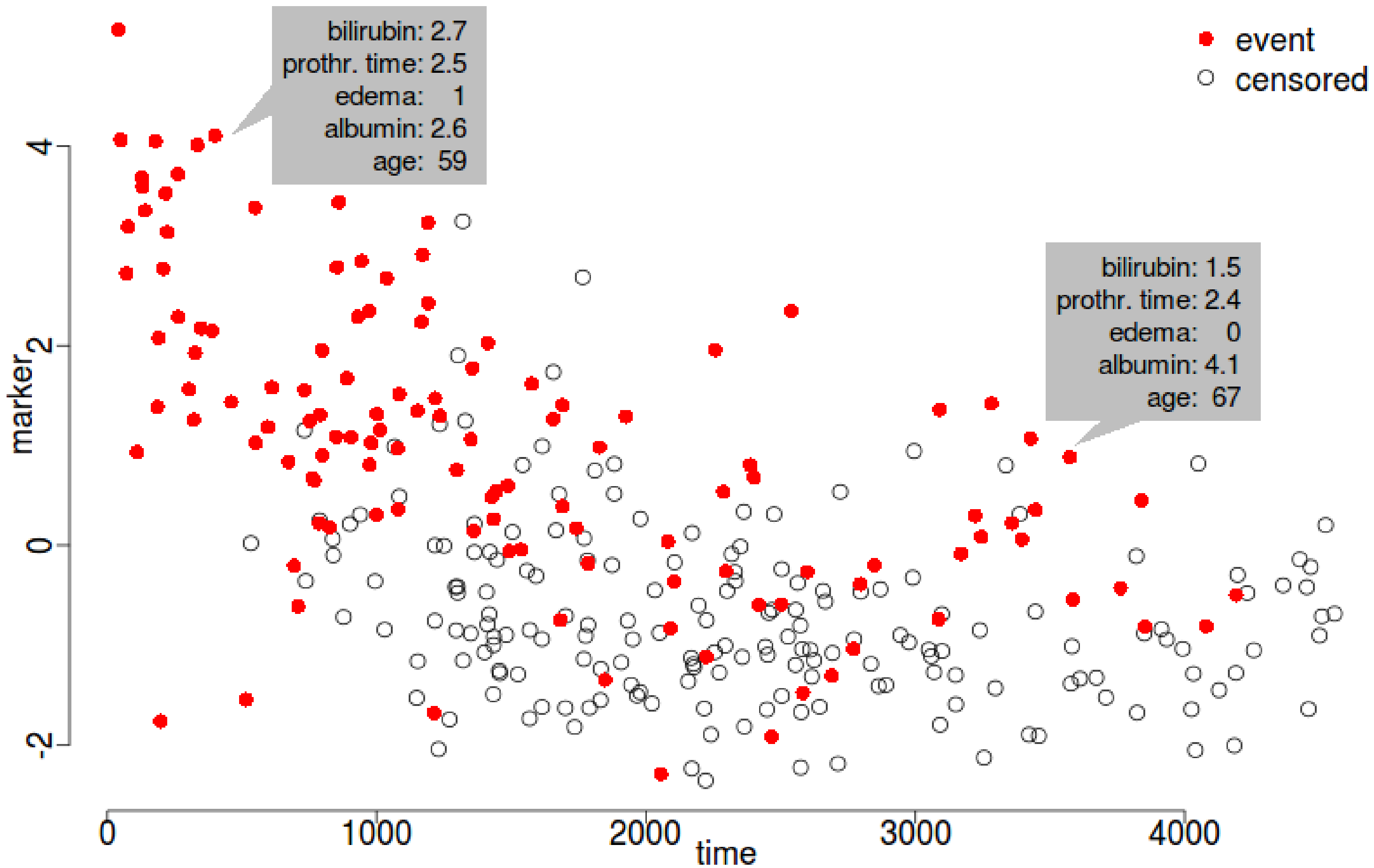
Is the patient on diuretic therapy? ◉ No ○ Yes

Compute

Risk score: 3.45817ł

Estimated Probability of Survival (%)

PBC Mayo data



Time-varying AUC curves

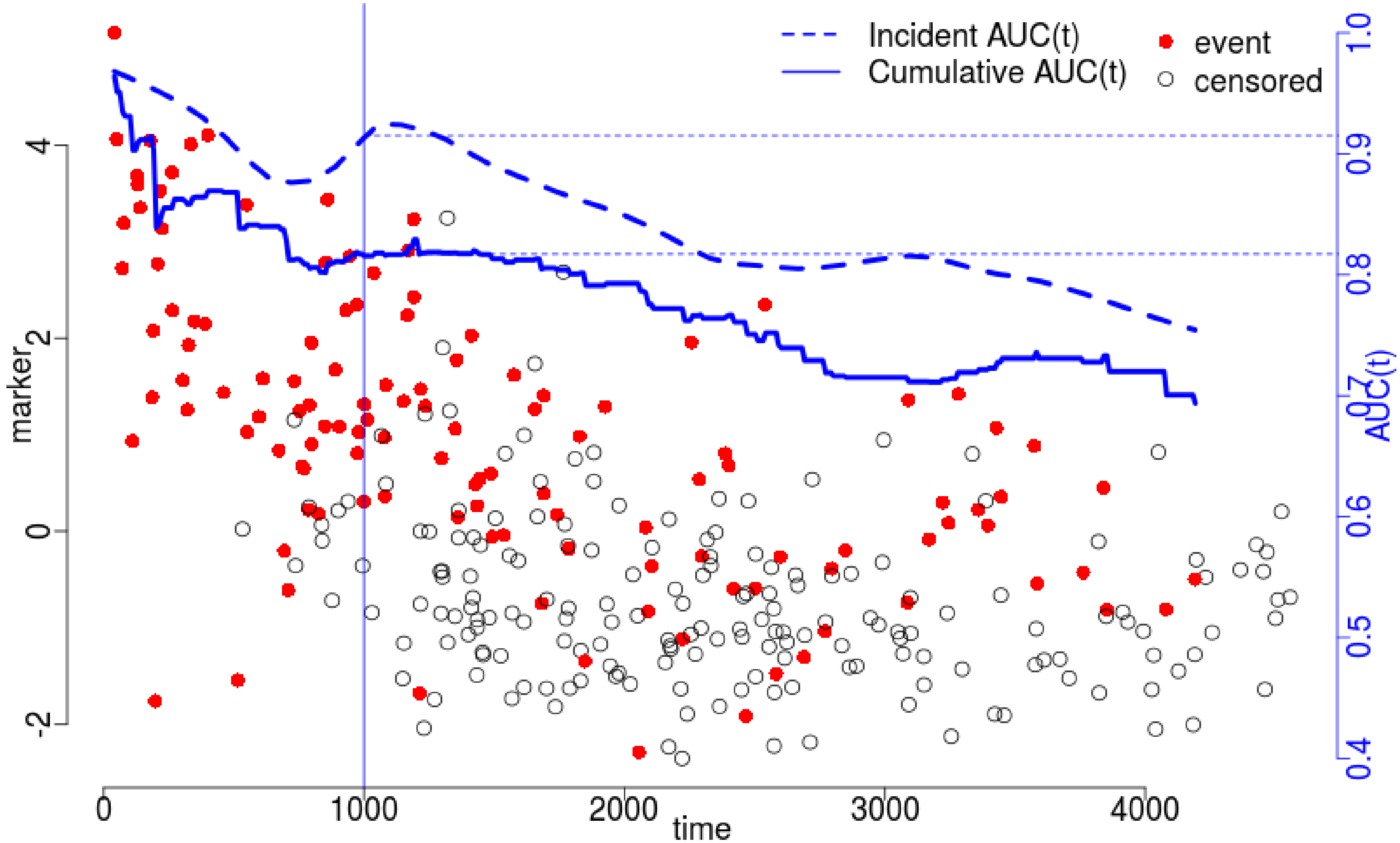
Cumulative cases:

$$AUC^C(t) = \mathrm{P}(M_i > M_j \,|\, T_i \leq t, T_j > t)$$

Incident cases:

$$AUC^I(t) = \mathrm{P}(M_i > M_j \,|\, T_i = t, T_j > t)$$

Time-varying AUC curves



Discrimination slope

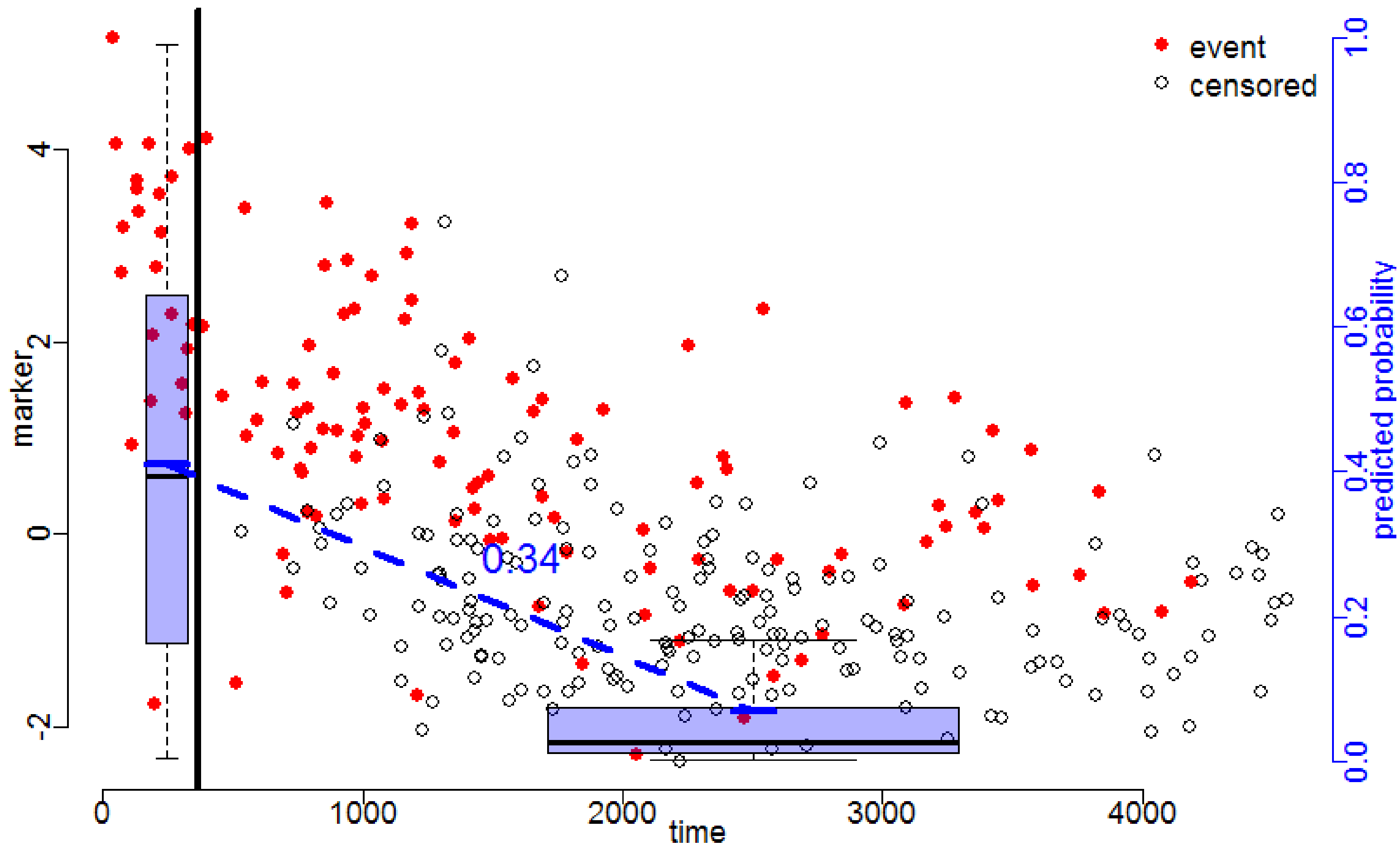
For binary outcomes:

$$DS = \textbf{mean case risk} - \textbf{mean control risk}$$

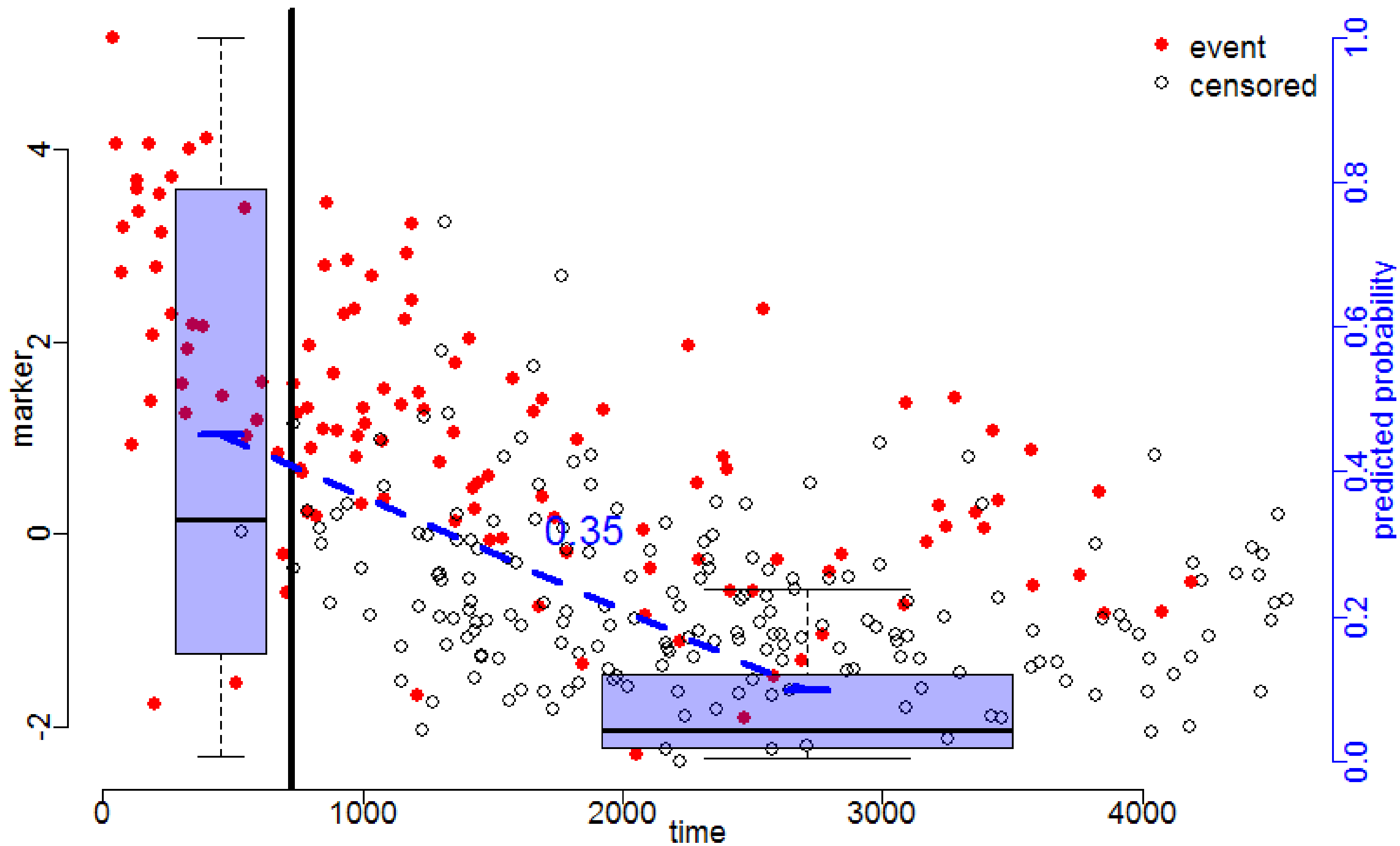
To generalize to survival outcomes:

$$\begin{aligned} DS(t) &= \textbf{mean case(t) risk} - \textbf{mean control(t) risk} \\ &= \mathbf{E}\left[\mathbf{F}\left(t \mid M\right) \mid T \leq t\right] - \mathbf{E}\left[\mathbf{F}\left(t \mid M\right) \mid T > t\right] \end{aligned}$$

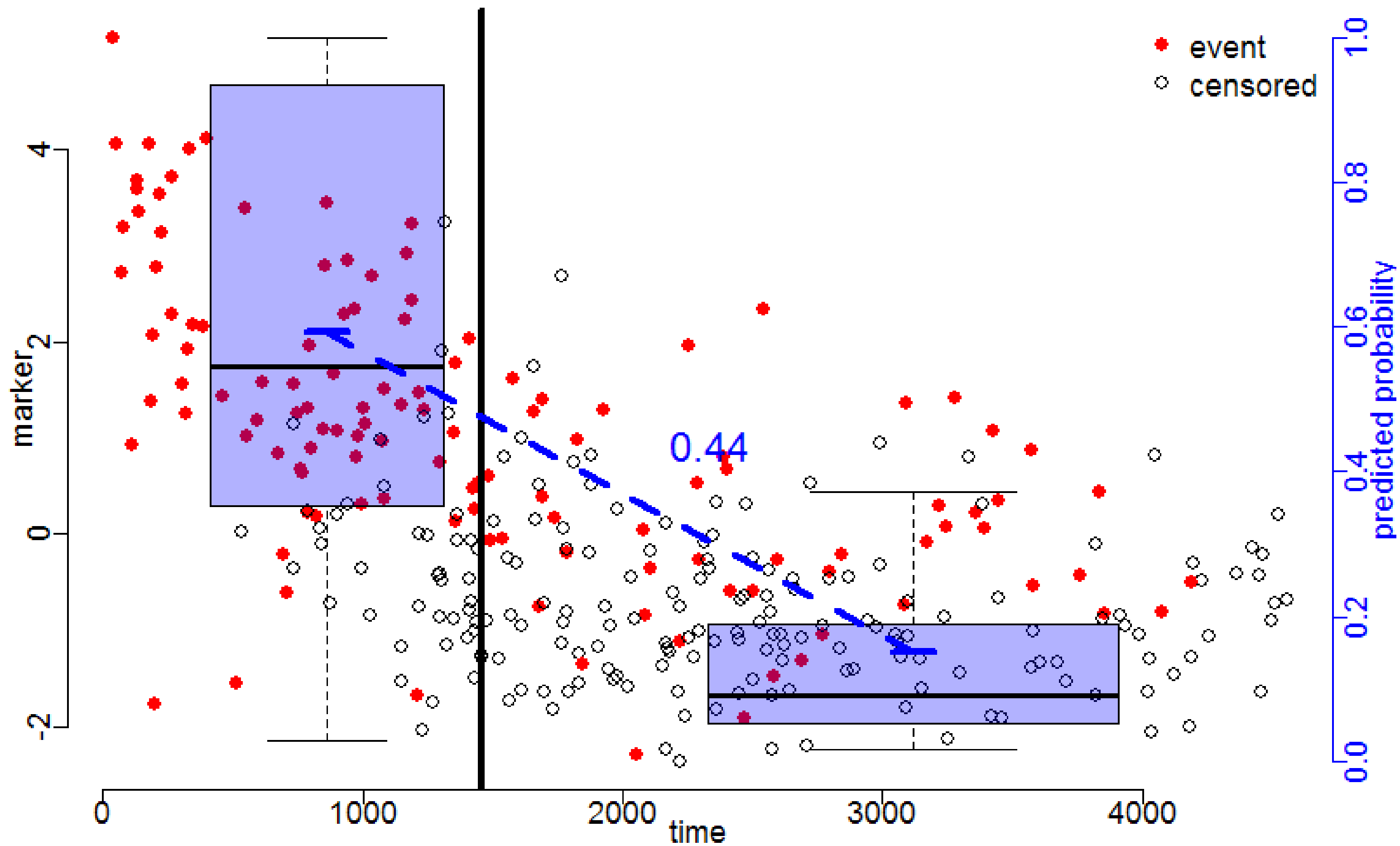
Discrimination slope (cumulative)



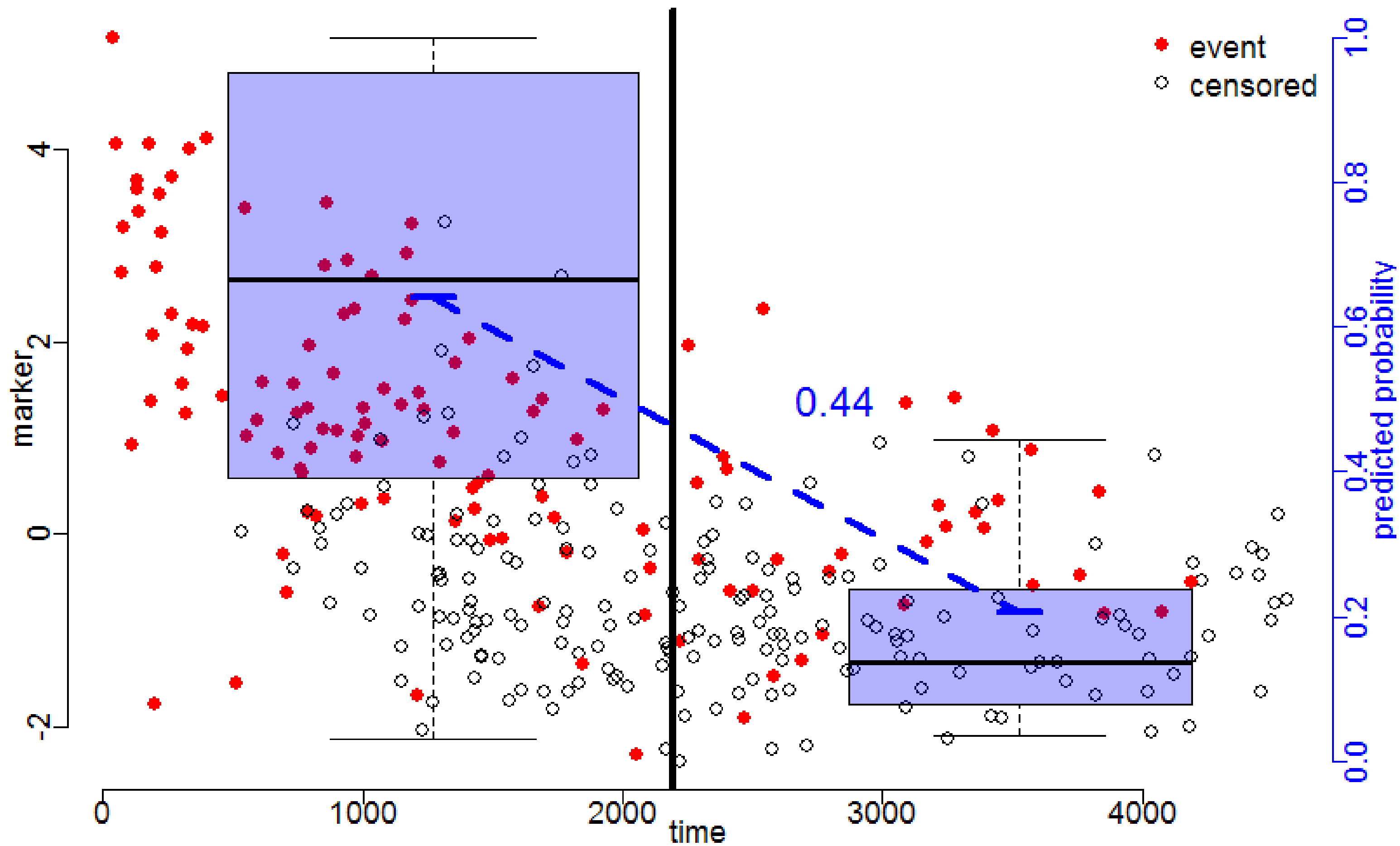
Discrimination slope (cumulative)



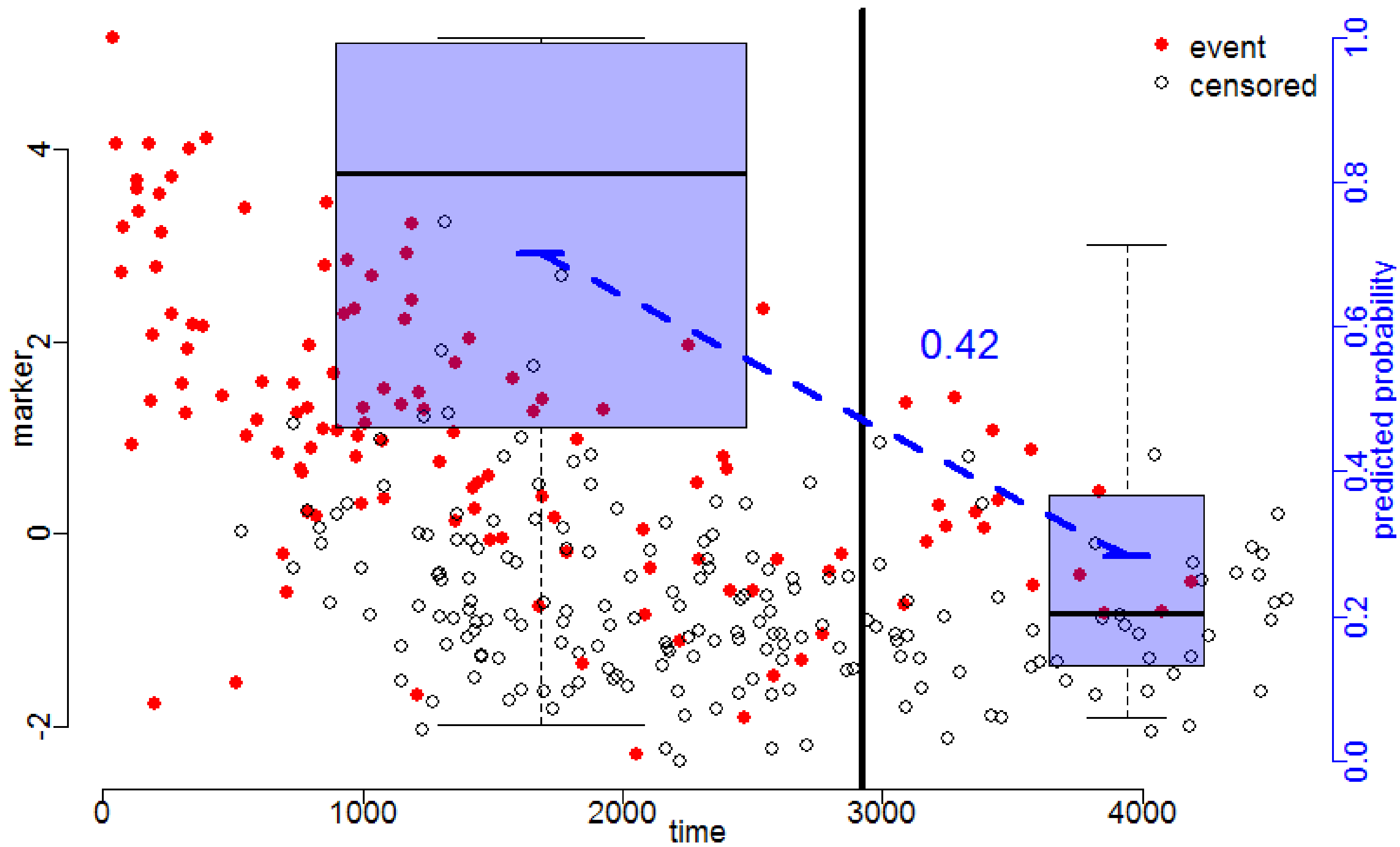
Discrimination slope (cumulative)



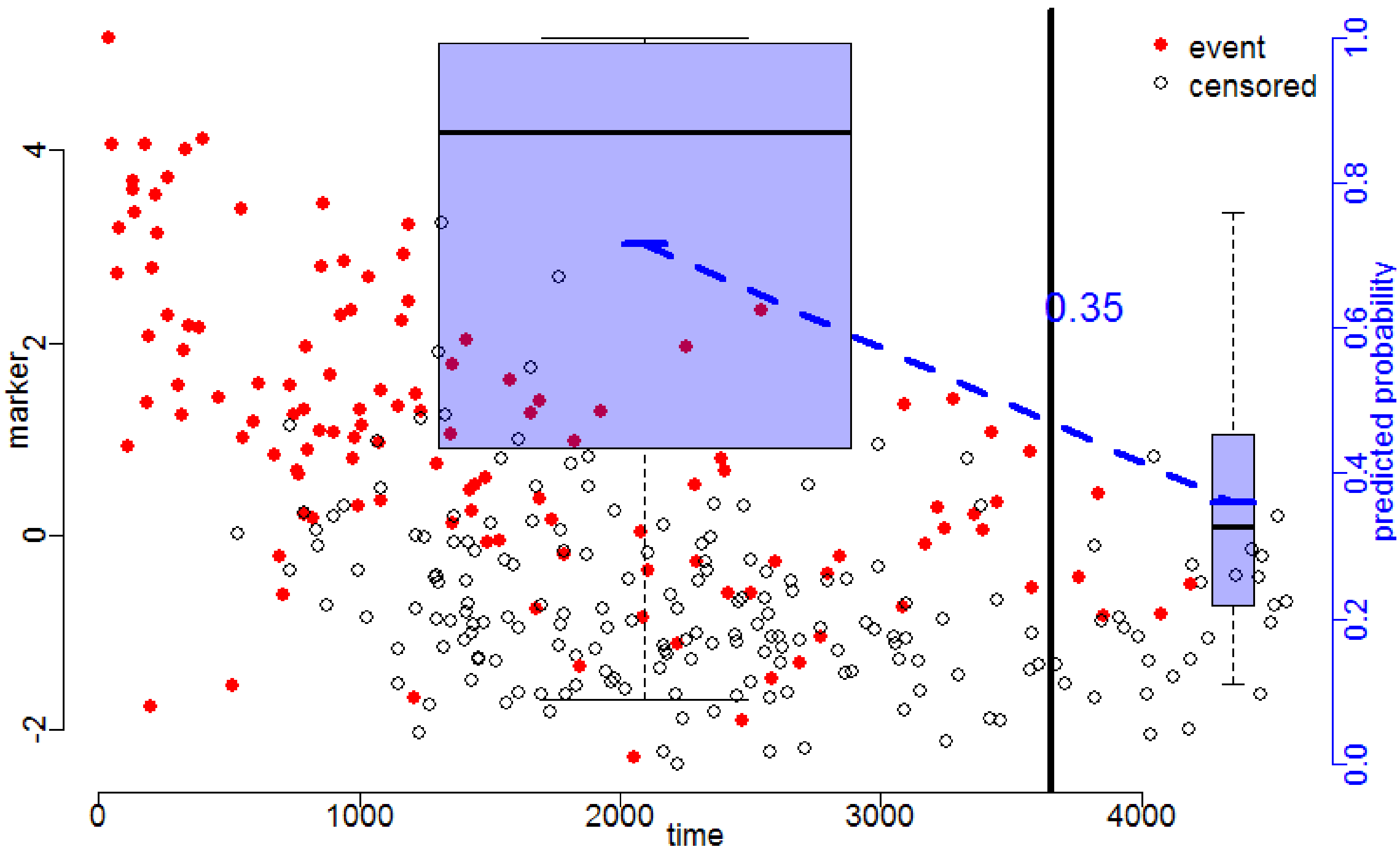
Discrimination slope (cumulative)



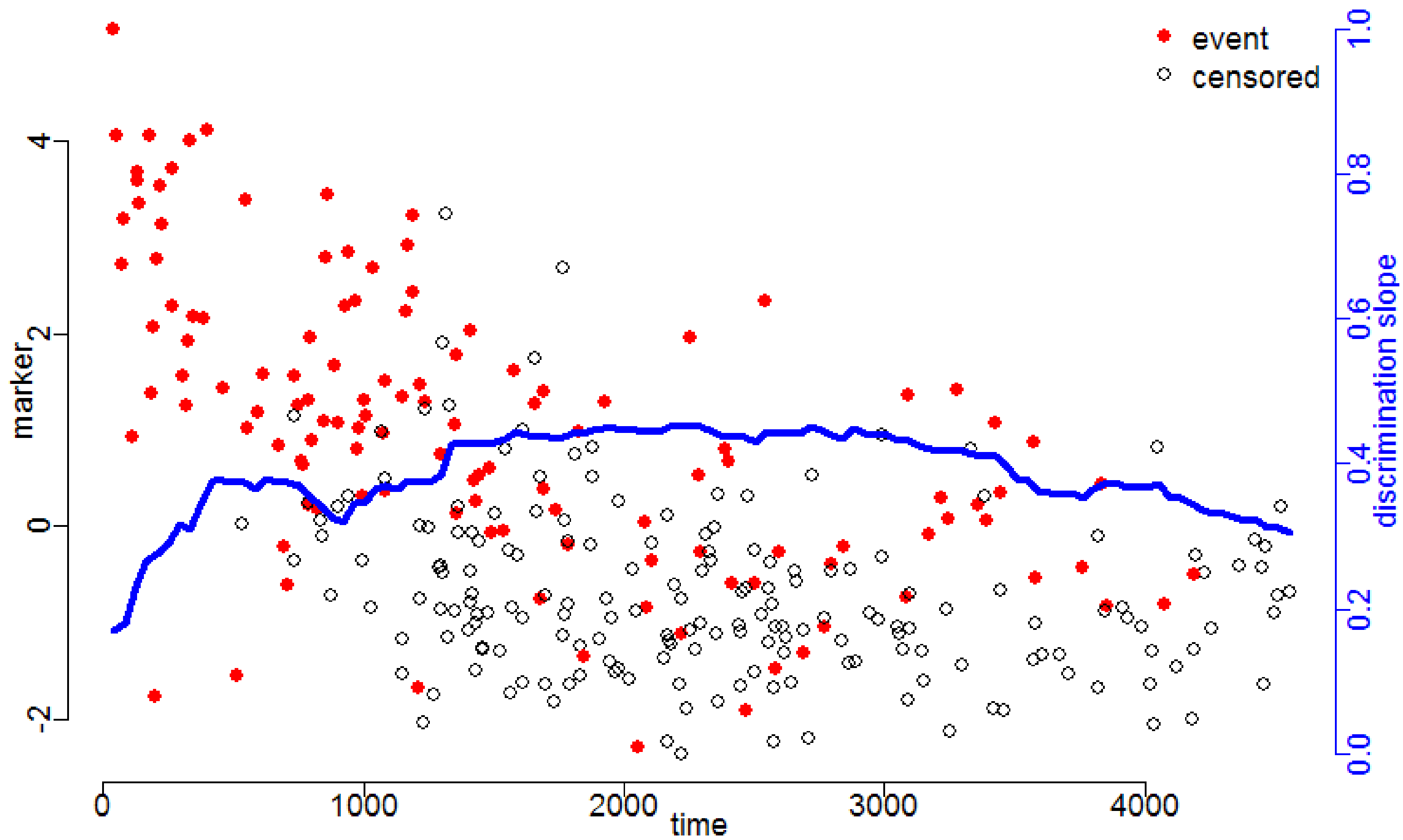
Discrimination slope (cumulative)



Discrimination slope (cumulative)



Discrimination slope (cumulative)



Measures of time-varying discrimination

	Cumulative cases	Incident cases
Rank based	• $AUC^C(t)$ • Heagerty, Lumley, and Pepe (2000)	• $AUC^I(t)$ • Heagerty and Zheng (2005); Saha and Heagerty, (2013)
Risk based	• $DS(t)$ • Yates (1982); Pencina *et al.* (2008); Uno *et al.*, (2012)	• ? • Liang & Heagerty (2017)

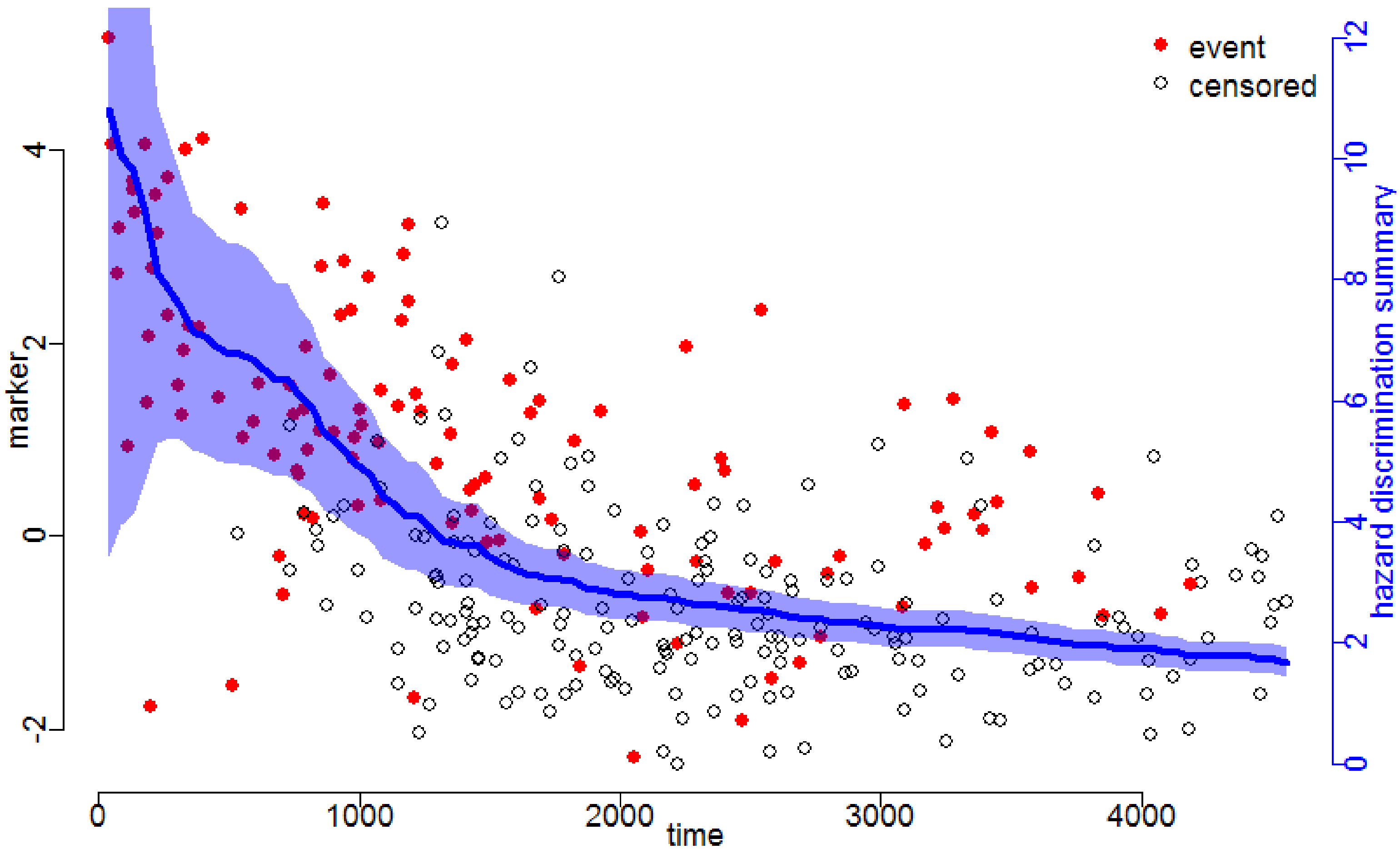
Parameter of Interest

Hazard Discrimination Summary
(*incident* cases)

$$HDS(t) = \frac{\mathrm{E}\left[\lambda(t|M)\,|\,T = t\right]}{\mathrm{E}\left[\lambda(t|M)\,|\,T > t\right]}$$

How well separated are incident cases from controls?

Illustration of $HDS(t)$



Estimation

Under the **Cox model** we can re-write $HDS(t)$

$$
\begin{aligned}
HDS(t) &= \frac{\mathrm{E}\left[\lambda(t|M) \mid T = t\right]}{\mathrm{E}\left[\lambda(t|M) \mid T > t\right]} \\
&= \frac{\int \exp\left(2\beta m - e^{\beta m}\Lambda_0(t)\right) dF_m \int \exp\left(-e^{\beta m}\Lambda_0(t)\right) dF_m}{\left[\int \exp(\beta m - e^{\beta m}\Lambda_0(t)) dF_m\right]^2}
\end{aligned}
$$

Estimation

Under the **Cox model** we can re-write $HDS(t)$

$$
\begin{aligned}
HDS(t) &= \frac{\mathrm{E}\left[\lambda(t|M) \mid T = t\right]}{\mathrm{E}\left[\lambda(t|M) \mid T > t\right]} \\
&= \frac{\int \exp\big(2\beta m - e^{\beta m}\Lambda_0(t)\big) dF_m \int \exp\big(-e^{\beta m}\Lambda_0(t)\big) dF_m}{\left[\int \exp(\beta m - e^{\beta m}\Lambda_0(t)) dF_m\right]^2}
\end{aligned}
$$

which lends itself to a straightforward plug-in estimator under the Cox model.

$$
\widehat{HDS}(t) = \frac{\sum \exp\left(2\hat{\beta} m_i - e^{\hat{\beta} m_i}\hat{\Lambda}_0(t)\right) \sum \exp\left(-e^{\hat{\beta} m_i}\hat{\Lambda}_0(t)\right)}{\left[\sum \exp\left(\hat{\beta} m_i - e^{\hat{\beta} m_i}\hat{\Lambda}_0(t)\right)\right]^2}
$$

Standard Errors

Estimator decomposes into two asymptotically uncorrelated parts: one dominated by the **empirical marker distribution** and one by $\hat{\beta}$, $\hat{\Lambda}_0(t)$

$$
\begin{aligned}
\widehat{HDS}(t) - HDS(t) \approx \\
& HDS(t; \hat{F}_M, \beta, \Lambda) - HDS(t) \\
& + HDS(t; F_M, \hat{\beta}, \hat{\Lambda}) - HDS(t)
\end{aligned}
$$

Applying Tsiatis (1981), van der Vaart & Wellner (2007) gives **pointwise asymptotic normality** with **analytic, consistent SE estimators**.

Relaxing the proportional hazards assumption

Instead of assuming

$$\lambda(t|M) = \lambda_0(t)\exp(\beta M)$$

Assume

$$\lambda(t|M) = \lambda_0(t)\exp(\beta(t) M)$$

Can we still estimate $HDS(t)$?

Relaxing the proportional hazards assumption

$$\widehat{HDS}^{LC}(t) = \frac{\sum \exp\left(2\hat{\beta}(t)m_i - e^{\hat{\beta}(t)m_i}\hat{\Lambda}_0(t)\right) \sum \exp\left(-e^{\hat{\beta}(t)m_i}\hat{\Lambda}_0(t)\right)}{\left[\sum \exp\left(\hat{\beta}(t)m_i - e^{\hat{\beta}(t)m_i}\hat{\Lambda}_0(t)\right)\right]^2}$$

$\hat{\beta}(t)$: Cai and Sun (2003)

$\hat{\Lambda}_0(t)$: Tian, Zucker, & Wei (2005)

Requires choosing a time-bandwidth $h = O(n^{-v})$, where $1/4 < v < 1/2$

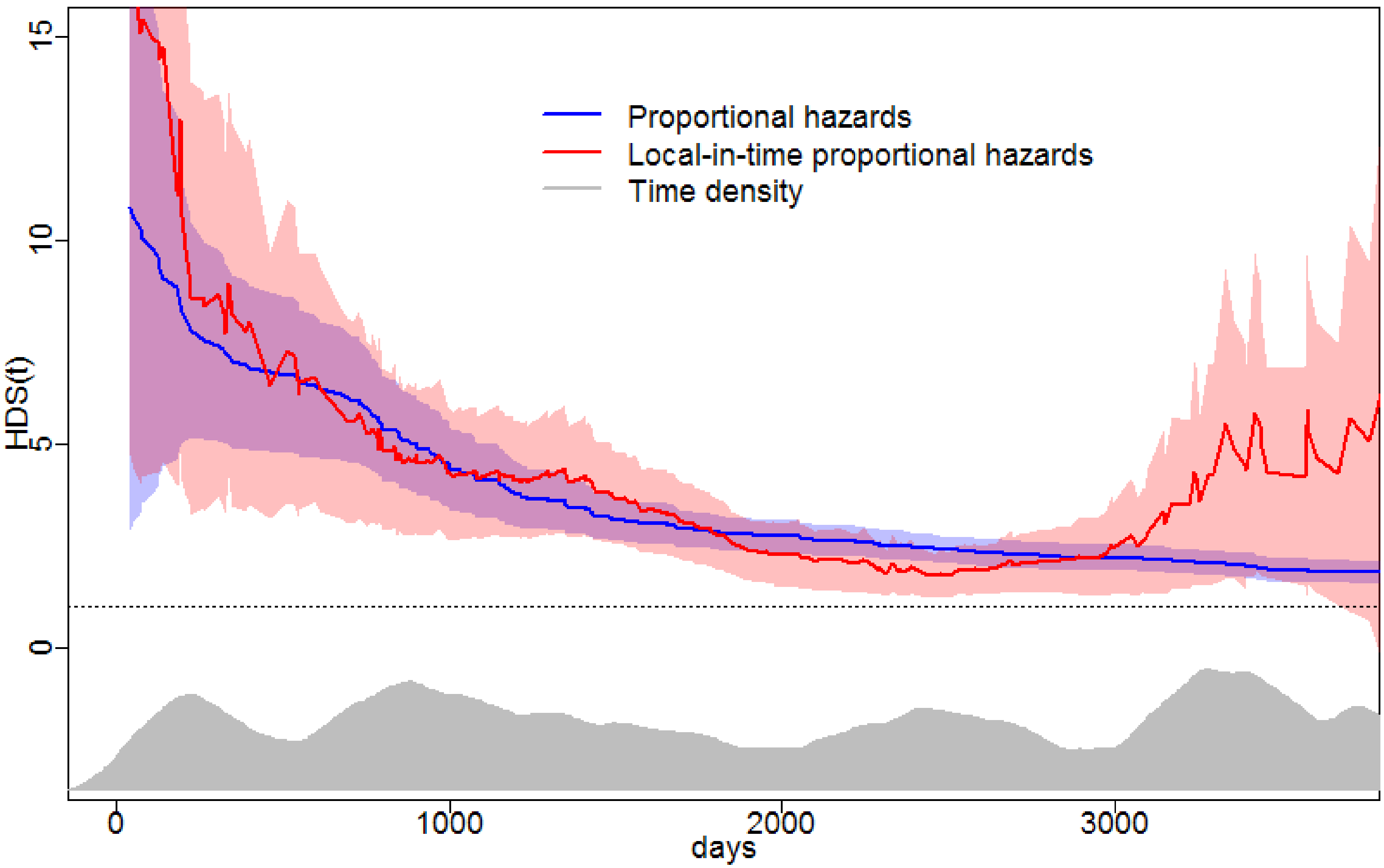
Standard Errors

Estimator is dominated by *just* $\hat{\beta}(t)$

$$\begin{aligned}\widehat{HDS}^{LC}(t) - HDS(t) \approx & \\ & HDS(t; \hat{F}_M, \beta(t), \Lambda) - HDS(t) \\ & + HDS(t; F_M, \hat{\beta}(t), \hat{\Lambda}) - HDS(t)\end{aligned}$$

Applying Cai and Sun (2003), Tian Zucker & Wei (2005), van der Vaart & Wellner (2007) gives **pointwise asymptotic normality** ($\sqrt{nh}$-rate) with **analytic, consistent SE estimators**.

Comparing two estimators



Summary

- Different types of time-varying measures of discrimination
- $HDS(t)$, an incident extension of discrimination slope
- Risk-based measure of time-varying discrimination
- Complements existing measures such as $AUC^C(t)$ and $AUC^I(t)$
- Handles ties more gracefully
- Can be more flexible measure of discrimination

Summary

- Connection between $HDS(t)$ and the partial likelihood
- Liang CJ, Heagerty PJ (2017 Biometrics): A risk-based measure of time-varying prognostic discrimination for survival models (with discussion).
- https://github.com/liangcj/hds and also on CRAN
- More work: time-varying covariates