Introduction to BART with time-to-event outcomes

Rodney Sparapani
Associate Professor of Biostatistics
Medical College of Wisconsin

September 16, 2025

This research was funded, in part, by the Advancing Healthier Wisconsin Research and Education Program and the NIH National Center for Advancing Translational Sciences

Abstract II: Survival analysis and ensembles

Deficiencies of parametric time-to-event survival analysis lead to semi-parametric methods (Cox 1972; Miller 1976). Due to the discovery of ensembles (Krogh, Sollich 1997) and technological advances like Moore's law, another transition to largely nonparametric methods for survival analysis is underway. Therefore, we choose BART for its relative flexibility, i.e., a nonparametric approach with no precarious parametric nor semi-parametric assumptions such as linearity and/or proportionality. Furthermore, due to its Bayesian nonparametric underpinnings, BART can be naturally extended to interpretable functions as targets of inference along with their measures of uncertainty, e.g., the survival function and its 95% credible intervals.

Outline

- Survival analysis with Cox Proportional Hazards
- Survival analysis with the discrete time approach
- Survival analysis with BART
- Example: advanced lung cancer prognosis
  `demo/lung.surv.bart.R` in the **BART** package
  and `demo/lung.relrisk.R` in the **BART3** package
- Motivation: diabetes and recurrent hospital admissions
  `demo/dm.recur.bart.R`
- Recurrent events with BART
- Motivation: liver transplant waiting list
  `demo/liver.crisk.bart.R`
- Competing risks

Semi-parametric survival analysis with Cox Proportional Hazards

Cox 1972 *JRSS-B*

Data: $(s_i, \delta_i), x_i$ **where** $\delta_i = 0$ for censoring and $\delta_i = 1$ for event

$0 = t_{(0)} < \cdots < t_{(J)} < \infty$: distinct ordered event times of s_i

$$(0, t_{(1)}] \ldots (t_{(J-1)}, t_{(J)}]$$ Conditionally independent intervals

$$\lambda(t|x_i) = \lambda_0(t)\mathrm{e}^{x_i'\beta}$$ Linear and proportional

$$[\beta|\lambda_0(t)] = \prod_i \frac{\mathrm{e}^{x_i'\beta}}{\sum_{j \in R(t_i)} \mathrm{e}^{x_j'\beta}}$$ Partial likelihood

$$\widehat{\Lambda}_0(t) = \sum_{t_i \le t} \frac{\delta_i}{\sum_{j \in R(t_i)} \mathrm{e}^{x_i'\widehat{\beta}}}$$ Cumulative baseline hazard

$$\widehat{S}_0(t) = \mathrm{e}^{-\widehat{\Lambda}_0(t)}$$

$$\widehat{S}(t|x_i) = \widehat{S}_0(t)^{\exp(x_i'\widehat{\beta})}$$

Parametric survival analysis: discrete time approach

Data: $(s_i, \delta_i), x_i(t)$ Time-dependent covariates

$$\mathbf{0} = t_{(0)} < \cdots < t_{(K)} < \infty : \text{distinct ordered times of } s_i$$

$$y_{ij} \overset{\text{ind}}{\sim} \mathbf{B}\big(p(t_{(j)}|x_{ij})\big) \quad \text{where } j = 1, \ldots, J_i = \arg\min_j s_i \leq t_{(j)}$$

$$= \delta_i \, \mathbf{I}(j = J_i)$$

$$[y|p] = \prod_{i=1}^{N} \prod_{j=1}^{J_i} p(t_{(j)}|x_{ij})^{y_{ij}} (1 - p(t_{(j)}|x_{ij}))^{1-y_{ij}} \quad \text{Likelihood}$$

$$S(t_{(j)}|x_{ij}) = \mathbf{P}\big[t > t_{(j)}|x_{ij}\big] = \prod_{j' \leq j} (1 - p(t_{(j')}|x_{ij'}))$$

Discrete time probability $\Rightarrow$ longitudinal dichotomous model

Nonparametric survival analysis with probit BART

Sparapani, Logan et al. 2016 *Statistics in medicine*
Data: $(s_i, \delta_i), x_i(t)$

$$\mathbf{0} = t_{(0)} < \cdots < t_{(K)} < \infty : \text{distinct ordered times, } s_i$$

$$y_{ij} \overset{\text{ind}}{\sim} \mathrm{B}\left(p(t_{(j)}|x_{ij})\right) \text{ where } j = 1, \ldots, J_i = \arg\min_j s_i \leq t_{(j)}$$

$$= \delta_i \, \mathrm{I}(j = J_i)$$

$$\mu = \Phi^{-1}(\bar{y}) \text{ where } \bar{y} = \sum_i \sum_j y_{ij} / \sum_{i'} J_{i'}$$

$$p(t_{(j)}|x_{ij}) = \Phi(f(t_{(j)}, x_{ij})) \text{ where } f \overset{\text{prior}}{\sim} \text{BART}\ (H = 50, \mu)$$

$$S(t_{(j)}|x_{ij}) = \mathrm{P}\left[t > t_{(j)}|x_{ij}\right] = \prod_{j' \leq j} (1 - p(t_{(j')}|x_{ij'}))$$

Discrete time probability model $\Rightarrow$ longitudinal probit BART

Survival analysis with BART and inference

We generate samples of f from the posterior with MCMC

$$\hat{f}(t,x) = M^{-1}\sum_m f_m(t,x) \qquad \text{Estimate } f$$

$$\hat{S}(t|x) = M^{-1}\sum_m S_m(t|x) \qquad \text{Survival function}$$

$$(S_{0.025}(t|x), S_{0.975}(t|x)) \qquad 95\% \text{ Credible Interval}$$

Survival analysis with BART and Friedman's partial dependence function

Friedman 2001 *AnnStat*

$$S(t|x) = S(t|x_S, x_C) \qquad \text{BART function where } x = [x_S, x_C]$$

$$S(t|x_S) = \mathbf{E}_{x_C}\left[S(t|x_S, x_C)|t, x_S\right]$$

$$\approx N^{-1}\sum_i S(t|x_S, x_{iC})$$

$$S_m(t|x_S) \equiv N^{-1}\sum_i S_m(t|x_S, x_{iC})$$

$$\hat{S}(t|x_S) \equiv M^{-1}\sum_m S_m(t|x_S)$$

Relative Risk with Friedman's partial dependence function

$$RR_m(t|x_n, x_d, x_C) = \frac{p_m(t|x_n, x_C)}{p_m(t|x_d, x_C)} \quad \text{Relative Risk}$$

$$= \frac{\Phi(f_m(t, x_n, x_C))}{\Phi(f_m(t, x_d, x_C))}$$

$$RR_m(t|x_n, x_d) \equiv N^{-1} \sum_i \frac{p_m(t|x_n, x_{iC})}{p_m(t|x_d, x_{iC})}$$

$$RR_m(.|x_n, x_d) = K^{-1} \sum_j RR_m(t_{(j)}|x_n, x_d) \quad \text{Assuming Proportionality}$$

surv.bart and mc.surv.bart input and output: part 1

```
post=surv.bart(x.train, times=times, delta=delta,
    ..., ndpost=M, ntree=50, keepevery=10) or
post=mc.surv.bart(x.train, times=times, delta=delta,
...,ndpost=M,ntree=50,keepevery=10,mc.cores=2,seed=99)
```

Input vector `times` with K distinct values and `x.train`: x_i

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

Output `post`, of type `survbart` which is essentially a list of matrices including: `post$prob.train`: $\hat{p}_m(t_{(j)}|x_i)$

$$\begin{bmatrix} \hat{p}_1(t_{(1)}|x_1) & \dots & \hat{p}_1(t_{(J_1)}|x_1) & \dots & \hat{p}_1(t_{(1)}|x_N) & \dots & \hat{p}_1(t_{(J_N)}|x_N) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{p}_M(t_{(1)}|x_1) & \dots & \hat{p}_M(t_{(J_1)}|x_1) & \dots & \hat{p}_M(t_{(1)}|x_N) & \dots & \hat{p}_M(t_{(J_N)}|x_N) \end{bmatrix}$$

surv.bart and mc.surv.bart input and output: part 2

```
post=surv.bart(x.train, times=times, delta=delta,
  x.test=x.train, ..., ndpost=M) or
post=mc.surv.bart(x.train, times=times, delta=delta,
  x.test=x.train, ..., ndpost=M, mc.cores=2, seed=99)
```

Input vector `times` with K distinct values and `x.train`: x_i

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

Output `post`, of type `survbart` which is essentially a list of matrices including: `post$surv.test`: $\hat{S}_m(t_{(j)}|x_i)$

$$\begin{bmatrix} \hat{S}_1(t_{(1)}|x_1) & \dots & \hat{S}_1(t_{(K)}|x_1) & \dots & \hat{S}_1(t_{(1)}|x_N) & \dots & \hat{S}_1(t_{(K)}|x_N) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{S}_M(t_{(1)}|x_1) & \dots & \hat{S}_M(t_{(K)}|x_1) & \dots & \hat{S}_M(t_{(1)}|x_N) & \dots & \hat{S}_M(t_{(K)}|x_N) \end{bmatrix}$$

`surv.pre.bart` input and output: part 1

```
pre <- surv.pre.bart(times, delta, x.train)
```

Output a list containing the data transformed such as

matrix `pre$tx.train` and vector `pre$y.train`:

$$\begin{bmatrix} t_{(1)} & x_1 \\ \vdots & \vdots \\ t_{(J_1)} & x_1 \\ \vdots & \vdots \\ t_{(1)} & x_N \\ \vdots & \vdots \\ t_{(J_N)} & x_N \end{bmatrix} \quad \begin{bmatrix} y_{11} = 0 \\ \vdots \\ y_{1J_1} = \delta_1 \\ \vdots \\ y_{N1} = 0 \\ \vdots \\ y_{NJ_N} = \delta_N \end{bmatrix}$$

surv.pre.bart input and output: part 2

```
pre <- surv.pre.bart(times, delta, x.train,
    x.test=x.train)
```

Output a list containing the data transformed such as matrix pre$tx.test:

$$\begin{bmatrix} t_{(1)} & x_1 \\ \vdots & \vdots \\ t_{(K)} & x_1 \\ \vdots & \vdots \\ t_{(1)} & x_N \\ \vdots & \vdots \\ t_{(K)} & x_N \end{bmatrix}$$

`predict.survbart` input and output

```
pred <- predict(post, pre$tx.test, mc.cores=1, ...)
```

Input matrices: `x.test`: x_i

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_Q \end{bmatrix}$$

Output `pred` of type `survbart` with `pred$surv.test`: $\hat{S}_m(t_{(j)}|x_i)$

$$\begin{bmatrix} \hat{S}_1(t_{(1)}|x_1) & \dots & \hat{S}_1(t_{(K)}|x_1) & \dots & \hat{S}_1(t_{(1)}|x_Q) & \dots & \hat{S}_1(t_{(K)}|x_Q) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{S}_M(t_{(1)}|x_1) & \dots & \hat{S}_M(t_{(K)}|x_1) & \dots & \hat{S}_M(t_{(1)}|x_Q) & \dots & \hat{S}_M(t_{(K)}|x_Q) \end{bmatrix}$$

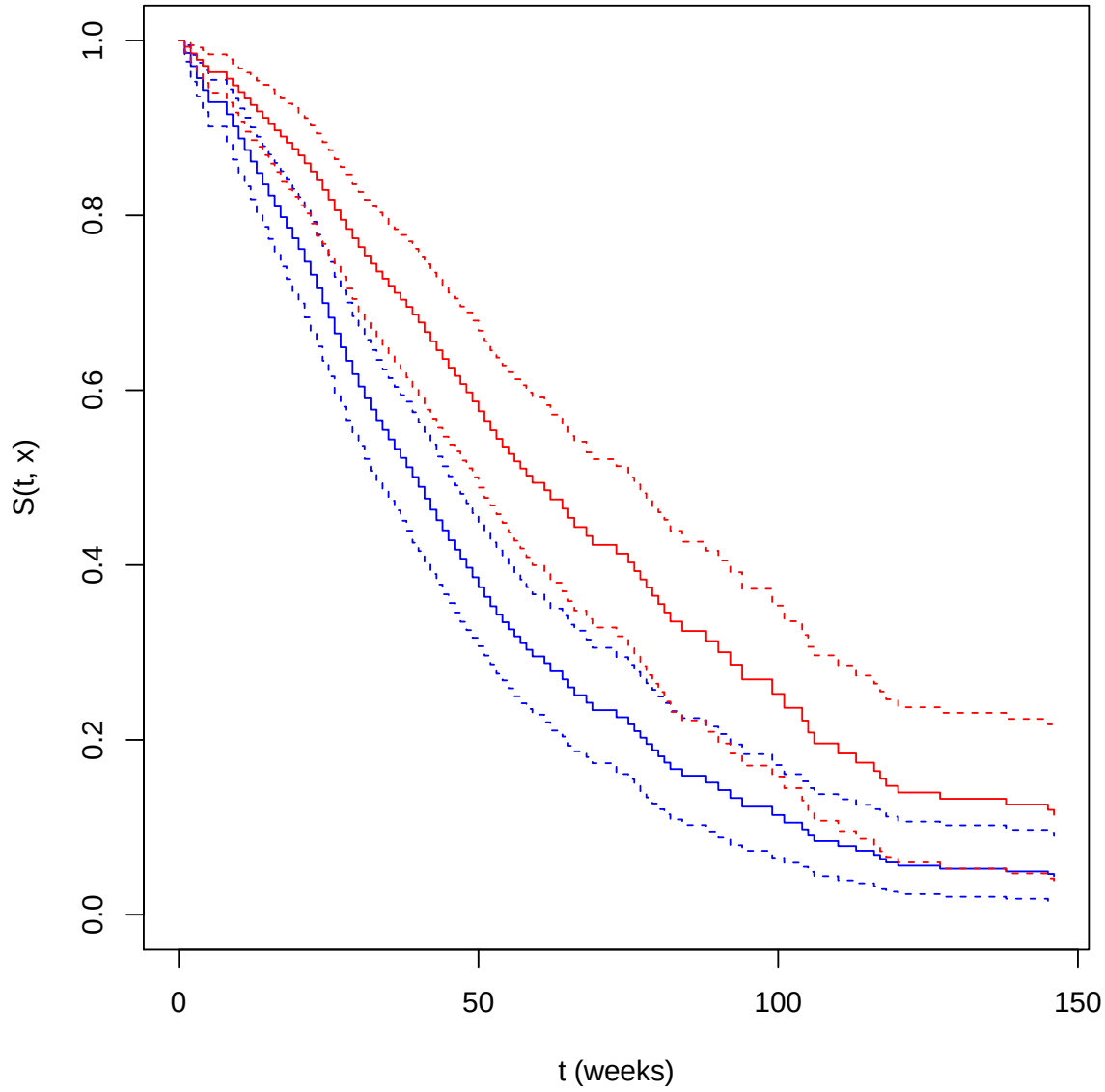
Survival analysis: advanced lung cancer prognosis

Loprinzi et al. 1994 *JCO*

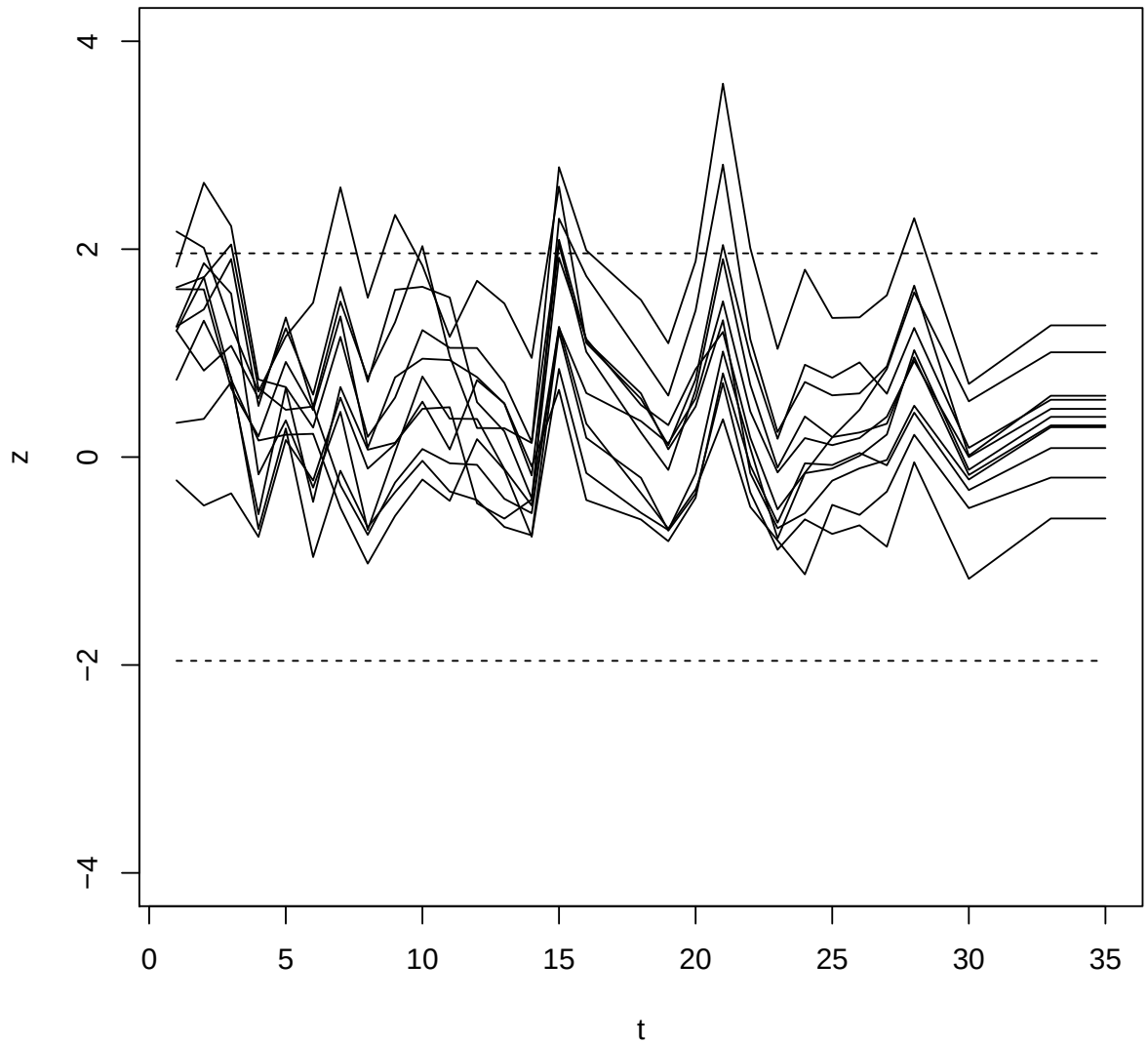
- The North Central Cancer Treatment Group surveyed 228 advanced lung cancer patients
- Study focused on prognostic variables
- Patient responses paired with some clinical variables
- We control for age, gender and Karnofsky performance score as rated by the physician
- We will compare males to females with Friedman's partial dependence function
- `lung` data set in the **BART** R package

```
system.file('demo/lung.surv.bart.R', package='BART')
system.file('demo/geweke.lung.surv.bart.R',
package='BART')
```

Friedman's partial dependence function with 95% credible intervals: M (blue) vs. F (red)



Geweke convergence diagnostics: Advanced lung cancer example



Diabetes and recurrent hospital admissions

- We have IRB approval to study a cohort of newly diagnosed diabetes patients from a single health care system
- We have the electronic health records (EHR) for these patients from 2007-2012: prior records may, or may not, be available
- EHR are an omnibus of digital health care information
- We focus on 82 covariates: patient demographics, health insurance, health care charges, diagnoses, procedures, anti-diabetic therapy, laboratory values and vital signs
- By its nature, EHR data is fundamentally time-varying
- EHR covariates are occasionally missing at time zero even when carrying the last value forward
- Imputed 15 continuous variables with Sequential BART (Xu, Daniels & Winterstein 2016 *Biostatistics*)

Diabetes and recurrent hospital admissions

- 488 patients followed 5 years from 2008-2012
  the survival rate was high 0.939 (noninformative censoring)
  yet experienced a high rate of hospital admissions: 525 total
- For diabetes, which covariates increase the risk of admission? What about the number of previous admissions or an acutely recent admission?
- What are the functional forms of the covariates i.e. linear, quadratic, logarithm, etc.? Are the covariate effects additive or multiplicative?
- Are there interactions? Are these effects constant with respect to time, i.e., proportionality assumption?
- We want to avoid precarious restrictive assumptions hence we chose to use Bayesian Additive Regression Trees (BART)

Recurrent event analysis with BART

Sparapani, Rein et al. 2018 *Biostatistics*

Data: $(s_i, t_{i1}, \ldots, t_{iN_i}), x_i(t)$

$(0, t_{(1)}] \ldots (t_{(K-1)}, t_{(K)}]$: grid of distinct ordered times for s_i and t_{ik}

$$y_{ij} \overset{\text{ind}}{\sim} \mathbf{B}\big(p(t_{(j)}|x_{ij})\big) \qquad j = 1, \ldots, J_i$$
$$= \max_{k=1,\ldots,N_i} \mathbf{I}\big(t_{ik} = t_{(j)}\big)$$

$$p(t_{(j)}|x_{ij}) = \Phi(f(t_{(j)}, x_{ij})) \qquad f \overset{\text{prior}}{\sim} \mathbf{BART}\ (\mu)$$

$$[y|f] = \prod_{i=1}^{N} \prod_{j=1}^{J_i} p(t_{(j)}|x_{ij})^{y_{ij}} (1 - p(t_{(j)}|x_{ij}))^{1-y_{ij}} \qquad \text{Likelihood}$$

$$\Lambda(t_{(j)}|x_{ij}) = \int_0^{t_{(j)}} \mathbf{d}\Lambda(t|x_i(t)) = \sum_{j'=1}^{j} p(t_{(j')}|x_{ij'})$$

Discrete time intensity model $\Rightarrow$ longitudinal probit BART

Semi-Markov process and conditional independence

- Note that $(t_{i1}, \ldots, t_{iN_i})$ are not independent
  we assume that they are conditionally independent given $x_i(t)$
  and the event history which we summarize by $N_i(t)$ and $v_i(t)$
- $N_i(t)$ is the number of events process and $N_i \equiv N_i(s_i)$
  When $N_i = 0$, then $t_{iN_i} = t_{i0} \equiv 0$
- Semi-Markov process, i.e., condition on summaries of
  the event history just prior to time t which is denoted by $t-$

$$N_i(t-)$$ Number of events at $t-$

$$v_i(t) \equiv t - t_{iN_i(t-)}$$ Sojourn time from the last event

$$\widetilde{x}_{ij} = \left[v_i(t_{(j)}), N_i(t_{(j-1)}), x_{ij}\right]$$

$$y_{ij} \overset{\text{ind}}{\sim} \mathbf{B}\left(p(t_{(j)}|\widetilde{x}_{ij})\right)$$

$$p(t_{(j)}|\widetilde{x}_{ij}) = \Phi(f(t_{(j)}, \widetilde{x}_{ij}))$$

Diabetes and recurrent hospital admissions

	Patients		Admissions	
Number of Admissions	488		525	
0	308	(63.0)	0	
1	79	(16.2)	79	(15.0)
2-3	50	(10.3)	115	(21.9)
4-16	51	(10.5)	331	(63.1)

Diabetes and recurrent hospital admissions

- We focus on 82 covariates: patient demographics, health insurance, health care charges, diagnoses, procedures, anti-diabetic therapy, laboratory values and vital signs
- These covariates are inherently time-dependent and occasionally missing at time zero even when carrying the last value forward
- Imputed 15 continuous variables with Sequential BART
  8 lab values and 7 vital signs
  Xu, Daniels & Winterstein 2016 *Biostatistics*
- Variable selection: Decoupling Shrinkage and Selection (DSS)
  Hahn & Carvalho 2015 *JASA*; McCulloch et al. 2015 JSM
- Divided the cohort at random into training and validation sets
- Risk agonists: insulin treatment, peripheral vascular disease (PVD) and the number of previous admissions, $N_i(t-)$

Diabetes and recurrent hospital admissions

	Patients		Admissions	
Gender	488		525	
M	216	(44.3)	228	(43.4)
F	272	(55.7)	297	(56.6)
Race	488		525	
Black	174	(35.7)	265	(50.5)
White	314	(64.3)	260	(49.5)
Age	488		525	
Mean, SD	60.9	15.0	60.3	15.7
ZIP3 area	488		525	
532/urban	378	(77.5)	454	(86.5)
530/suburb	110	(22.5)	71	(13.5)
Insurance and Age	488		525	
Government 65+	191	(39.1)	224	(42.7)
Government <65	138	(28.3)	208	(39.6)
Commercial <65	143	(29.3)	71	(13.5)
Other <65	16	(3.3)	22	(4.2)

Diabetes and recurrent hospital admissions

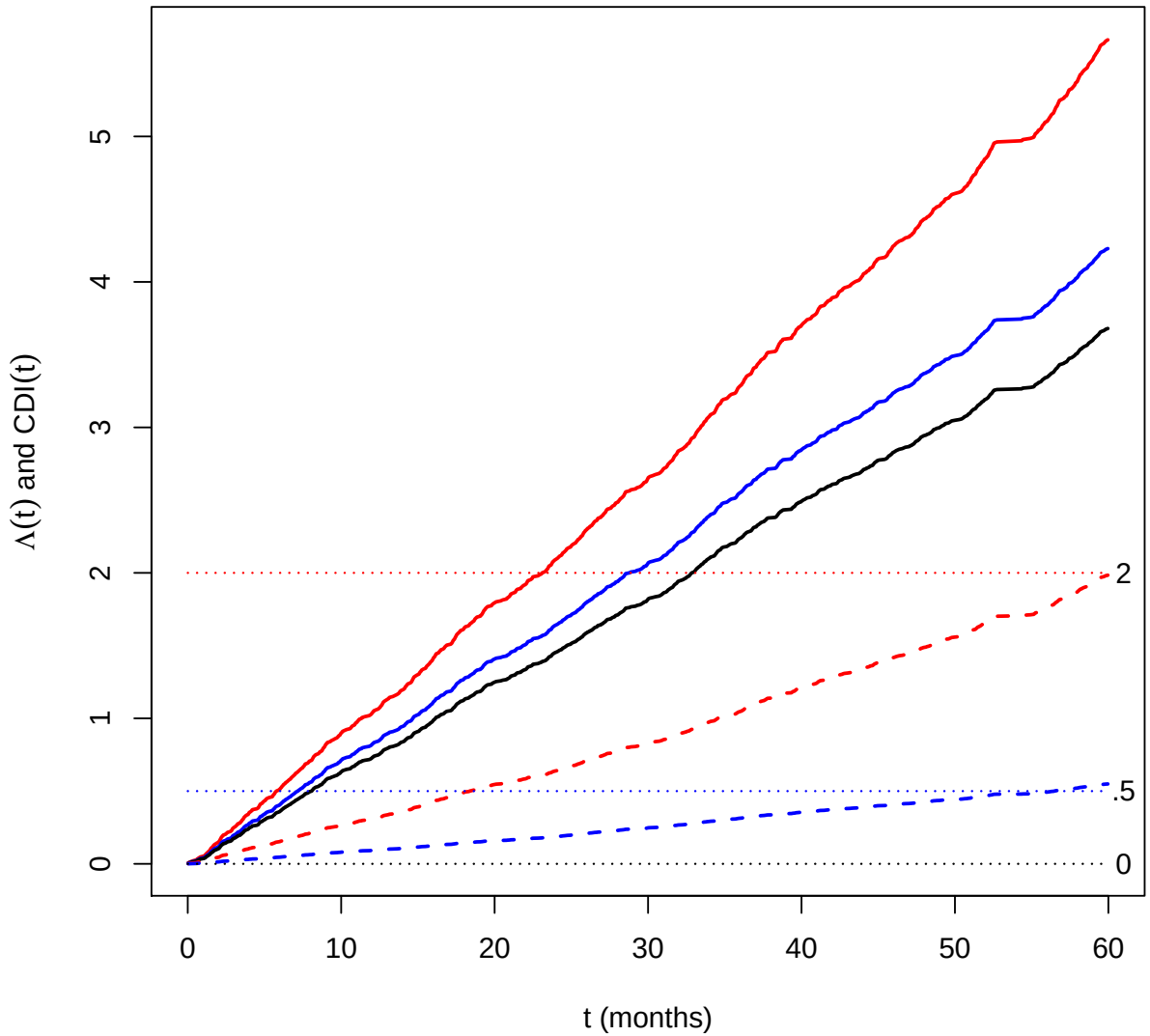
	Patients		Admissions		Relative Intensity	95% Credible Interval
Insulin	488		525		2.39	1.56, 3.25
Yes	206	(42.2)	391	(74.5)		
No	282	(57.8)	134	(25.5)		
PVD	488		525		2.90	2.00, 3.89
Yes	272	(55.7)	488	(93.0)		
No	216	(44.3)	37	(7.0)		

partial dependence function

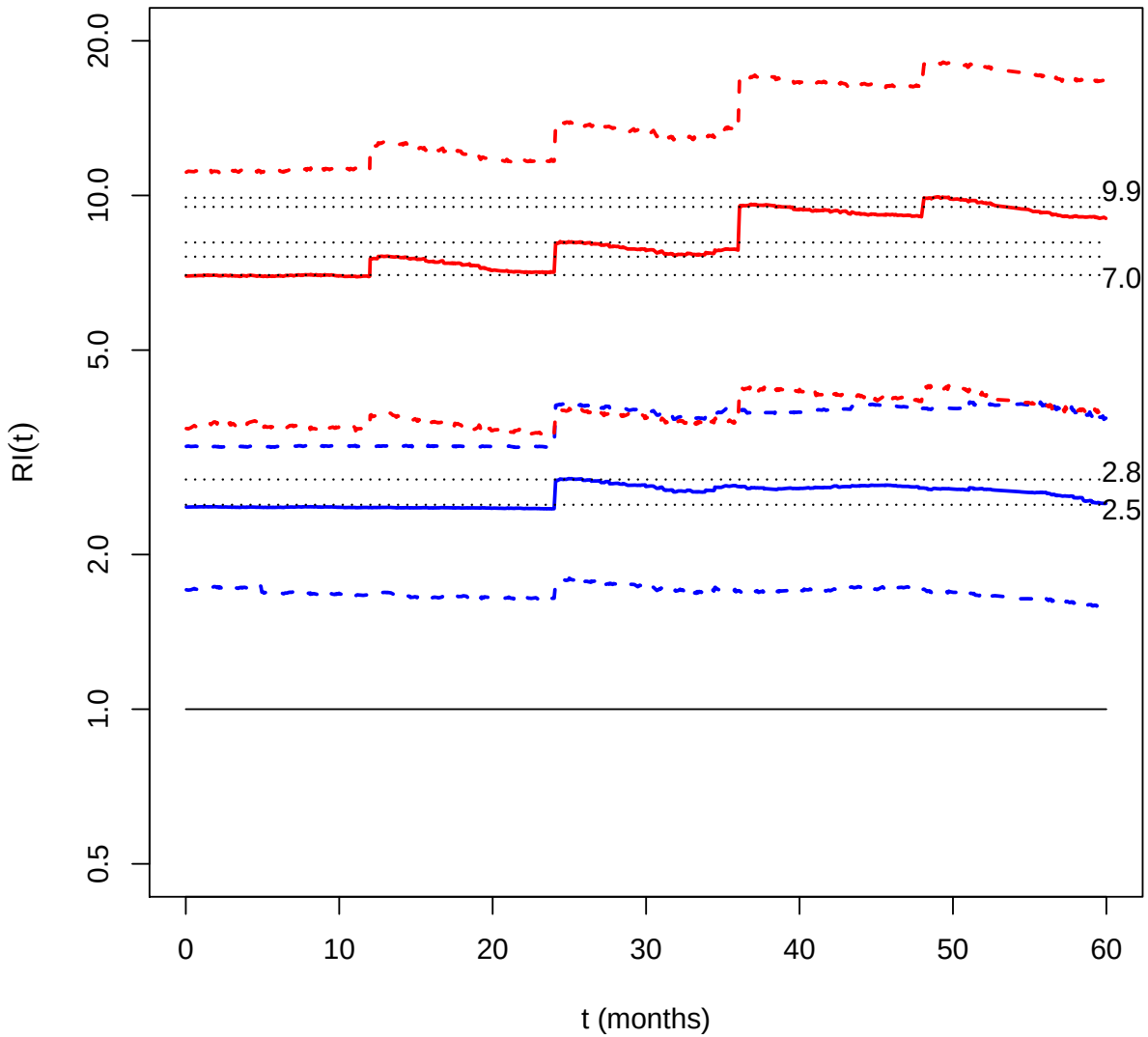
Hospital admission risk profiles

Risk	Insulin	PVD	$N_i(t)$ with time in months 0	12	24	36	48	60
Low	0	0	0	0	0	0	0	0
Medium	1	0	0	0	1	1	1	1
High	1	1	0	1	2	3	4	4

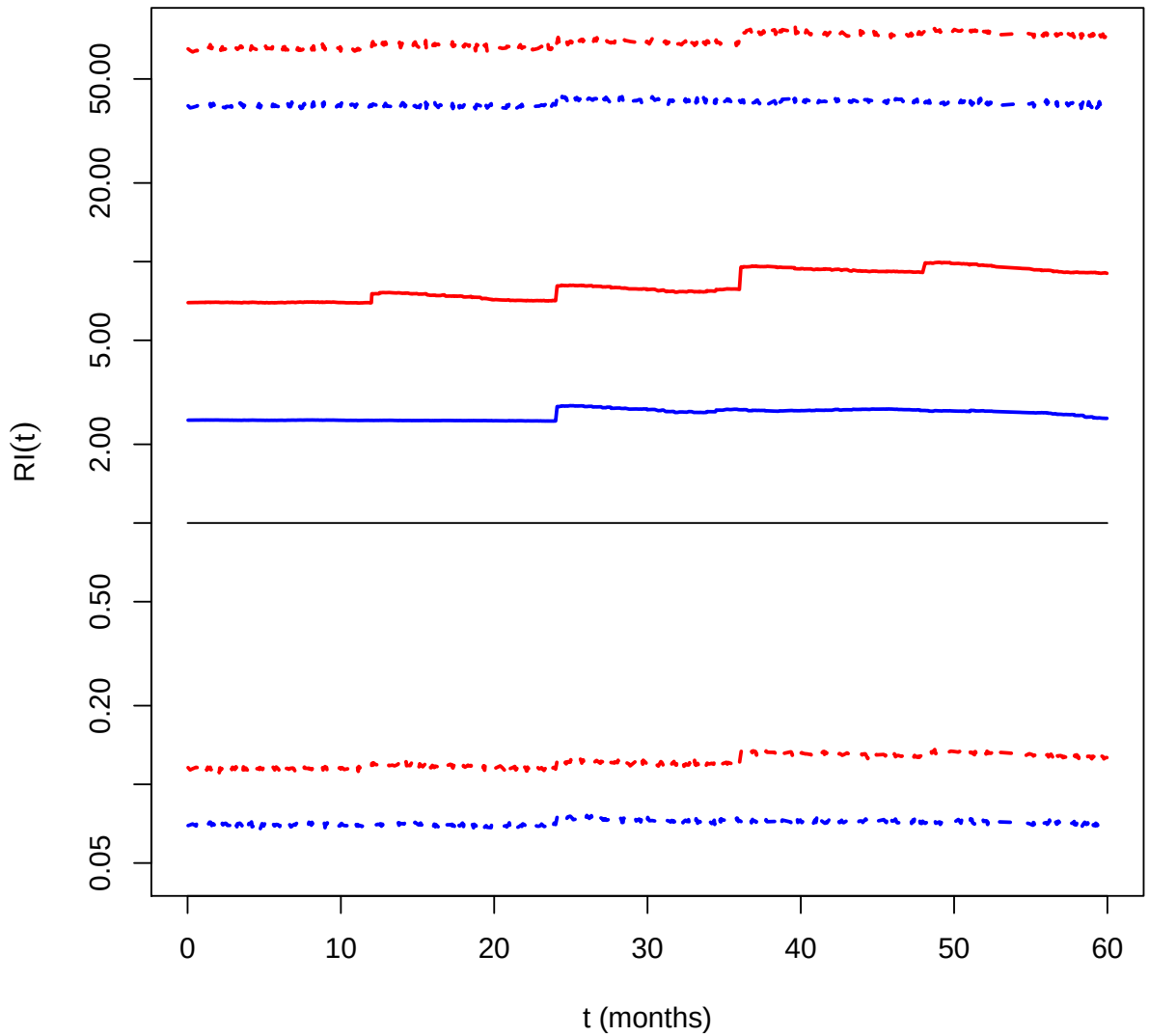
Risk profiles: Cumulative Intensity partial dependence function



Risk profiles: Relative Intensity partial dependence function



Risk profiles: Relative Intensity & 95% Prediction Intervals partial dependence function



Diabetes and hospital admission risk

- Some diabetes patients are at high risk for hospital admission
  - diagnosed with PVD
  - prescribed insulin therapy
  - with a recent hospital admission
  - and/or several previous hospital admissions
- Health policy implications: Diabetic patients' health care post-discharge should be carefully orchestrated to ensure the delivery of quality clinical care which maximizes healthy outcomes while preventing adverse events and costly unnecessary hospital admissions
- **BART** package contains a roughly 20% random sample
  50 patients from training: `ydm20.train` & `xdm20.train`
  50 patients from validation: `xdm20.test`
- See example: `system.file('demo/dm.recur.bart.R', package='BART')`
- The complete data set is available in the **BART3** package

Competing risks: More demanding `crisk.bart`

Data: $(t_i, \delta_i), x_i(t)$ where $\delta_i \in \{0, 1, 2\}$

$0 = t_{(0)} < \cdots < t_{(K)} < \infty$: distinct ordered, t_i, times

$$
\begin{aligned}
y_{1ij} &= \mathrm{I}(\delta_i = 1)\,\mathrm{I}(j = J_i)\,,\ j = 1, \ldots, J_i \\
&\sim \mathrm{B}(p_{1ij}) \\
p_{1ij} &= \Phi(f_1(t_{(j)}, x_{ij})) \text{ where } f_1 \overset{\text{prior}}{\sim} \text{BART}\,(\mu_1) \\
y_{2ij} &= \mathrm{I}(\delta_i = 2)\,\mathrm{I}(j = J_i)\,,\ j = 1, \ldots, K_i \\
&\quad \text{where } K_i = J_i - \mathrm{I}(\delta_i = 1) \\
&\sim \mathrm{B}(p_{2ij}) \\
p_{2ij} &= \Phi(f_2(t_{(j)}, x_{ij})) \text{ where } f_2 \overset{\text{prior}}{\sim} \text{BART}\,(\mu_2)
\end{aligned}
$$

$$
\begin{aligned}
[y|p] = \prod_{i=1}^{N} &\left(\prod_{j=1}^{J_i} p_{1ij}^{y_{1ij}} (1 - p_{1ij})^{1 - y_{1ij}} \right) \\
&\times \left(\prod_{j=1}^{K_i} p_{2ij}^{y_{2ij}} (1 - p_{2ij})^{1 - y_{2ij}} \right)
\end{aligned}
$$

Competing risks: More demanding `crisk.bart`

$$S(t|x_i(t)) = 1 - F(t|x_i(t)) = \prod_{j=1}^{k}(1-p_{1ij})(1-p_{2ij})$$

$$\text{where } k = \arg\max_j \left[t_{(j)} \leq t\right]$$

$$F_1(t|x_i(t)) = \int_0^t S(u-|x_i(u-))\lambda_1(u|x_i(u))\mathrm{d}u$$

$$= \sum_{j=1}^{k} S(t_{(j-1)}|x_i(t_{(j-1)}))p_{1ij}$$

$$F_2(t|x_i(t)) = \int_0^t S(u-|x_i(u-))\lambda_2(u|x_i(u))\mathrm{d}u$$

$$= \sum_{j=1}^{k} S(t_{(j-1)}|x_i(t_{(j-1)}))(1-p_{1ij})p_{2ij}$$

Competing risks: Less demanding `crisk2.bart`

Sparapani, Logan et al. 2019 SMMR (SparLoga19)

Data: $(t_i, \delta_i), x_i(t)$ **where** $\delta_i \in \{0, 1, 2\}$

$0 = t_{(0)} < \cdots < t_{(K)} < \infty$: distinct ordered, t_i, times

$$y_{1ij} = \mathbf{I}(\delta_i > 0)\,\mathbf{I}(j = J_i)\,,\ j = 1, \ldots, J_i$$

$$y_{1ij}|p_{1ij} \sim \mathbf{B}(p_{1ij})$$

$$p_{1ij} = \Phi(f_1(t_{(j)}, x_{ij})) \text{ where } f_1 \overset{\text{prior}}{\sim} \text{BART}\ (\mu_1)$$

$$y_{2i} = \mathbf{I}(\delta_i = 1|\delta_i > 0)$$

$$y_{2i}|p_{2i} \sim \mathbf{B}(p_{2i})$$

$$p_{2i} = \Phi(f_2(t_i, x_{iJ_i})) \text{ where } f_2 \overset{\text{prior}}{\sim} \text{BART}\ (\mu_2)$$

$$[y|p] = \left(\prod_{i=1}^{N}\prod_{j=1}^{J_i} p_{1ij}^{y_{1ij}}(1-p_{1ij})^{1-y_{1ij}}\right)\left(\prod_{i:\delta_i>0} p_{2i}^{y_{2i}}(1-p_{2i})^{1-y_{2i}}\right)$$

Competing risks: Less demanding `crisk2.bart`

$$S(t|x_i(t)) = 1 - F(t|x_i(t)) = \prod_{j=1}^{k}(1 - p_{1ij})$$

$$\text{where } k = \arg\max_j \left[t_{(j)} \leq t\right]$$

$$F_1(t|x_i(t)) = \int_0^t S(u - |x_i(u-))\lambda_1(u|x_i(u))\mathrm{d}u$$

$$= \sum_{j=1}^{k} S(t_{(j-1)}|x_i(t_{(j-1)}))p_{1ij}\Phi(f_2(t_{(j)}, x_{ij}))$$

$$F_2(t|x_i(t)) = \int_0^t S(u - |x_i(u-))\lambda_2(u|x_i(u))\mathrm{d}u$$

$$= \sum_{j=1}^{k} S(t_{(j-1)}|x_i(t_{(j-1)}))p_{1ij}\left[1 - \Phi(f_2(t_{(j)}, x_{ij}))\right]$$

Competing risks with 3 causes: More demanding

`crisk3.bart`

$$
\begin{aligned}
y_{1ij} &= \mathbf{I}(\delta_i = 1)\,\mathbf{I}(j = J_i)\,,\ j = 1, \ldots, J_i \\
&\sim \mathbf{B}(p_{1ij}) \\
p_{1ij} &= \Phi(f_1(t_{(j)}, x_{ij})) \text{ where } f_1 \overset{\text{prior}}{\sim} \text{BART}\ (\mu_1) \\
y_{2ij} &= \mathbf{I}(\delta_i = 2)\,\mathbf{I}(j = J_i)\,,\ j = 1, \ldots, K_i = J_i - \mathbf{I}(\delta_i = 1) \\
&\sim \mathbf{B}(p_{2ij}) \\
p_{2ij} &= \Phi(f_2(t_{(j)}, x_{ij})) \text{ where } f_2 \overset{\text{prior}}{\sim} \text{BART}\ (\mu_2) \\
y_{3ij} &= \mathbf{I}(\delta_i = 3)\,\mathbf{I}(j = J_i)\,,\ j = 1, \ldots, L_i = J_i - \mathbf{I}(\delta_i \in \{1, 2\}) \\
&\sim \mathbf{B}(p_{3ij}) \\
p_{3ij} &= \Phi(f_3(t_{(j)}, x_{ij})) \text{ where } f_3 \overset{\text{prior}}{\sim} \text{BART}\ (\mu_3)
\end{aligned}
$$

Competing risks with 3 causes: More demanding

`crisk3.bart`

$$S(t|x_i(t)) = 1 - F(t|x_i(t)) = \prod_{j=1}^{k}(1-p_{1ij})(1-p_{2ij})(1-p_{3ij})$$

$$\text{where } k = \arg\max_j \left[t_{(j)} \leq t\right]$$

$$F_1(t|x_i(t)) = \sum_{j=1}^{k} S(t_{(j-1)}|x_i(t_{(j-1)}))p_{1ij}$$

$$F_2(t|x_i(t)) = \sum_{j=1}^{k} S(t_{(j-1)}|x_i(t_{(j-1)}))(1-p_{1ij})p_{2ij}$$

$$F_3(t|x_i(t)) = \sum_{j=1}^{k} S(t_{(j-1)}|x_i(t_{(j-1)}))(1-p_{1ij})(1-p_{2ij})p_{3ij}$$

Competing risks with more than two causes: Less demanding

Data: $(t_i, \delta_i), x_i(t)$ where $\delta_i \in \{0, 1, \ldots, K\}$
$0 = t_{(0)} < \cdots < t_{(J)} < \infty$: distinct ordered, t_i, times

$$
\begin{aligned}
y_{ij} &= \mathrm{I}(\delta_i > 0)\,\mathrm{I}(j = J_i)\,,\ j = 1, \ldots, J_i \\
&\sim \mathrm{B}(p_{ij}) \\
p_{ij} &= \Phi(f_0(t_{(j)}, x_{ij})) \text{ where } f_0 \overset{\text{prior}}{\sim} \text{BART}\ (\mu_0) \\
\psi_{ik} &= \mathrm{I}(\delta_i = k | \delta_i > 0) \ \text{ where } \psi_{i0} = \mathrm{I}(\delta_i > 0) \\
&\sim \text{Multinomial}\ (1,\ \pi_i) \ \text{ where } \pi_{ik}(t_i | x_{iJ_i}) \text{ (Categorical BART)}
\end{aligned}
$$

$$
[y, \psi | p, \pi] = \prod_{i=1}^{N} \left(\prod_{j=1}^{J_i} p_{ij}^{y_{ij}} (1 - p_{ij})^{1 - y_{ij}} \right) \prod_{k=1}^{K} \left(\pi_{ik}^{\psi_{ik}} (1 - \pi_{ik})^{1 - \psi_{ik}} \right)^{\psi_{i0}}
$$

Competing risks with more than two causes: Less demanding

$$S(t|x_i(t)) = 1 - F(t|x_i(t)) = \prod_{j=1}^{j'}(1 - p_{ij})$$

$$\text{where } j' = \arg\max_j \left[t_{(j)} \leq t\right]$$

$$F_k(t|x_i(t)) = \int_0^t S(u - |x_i(u-))\lambda_k(u|x_i(u))\mathrm{d}u$$

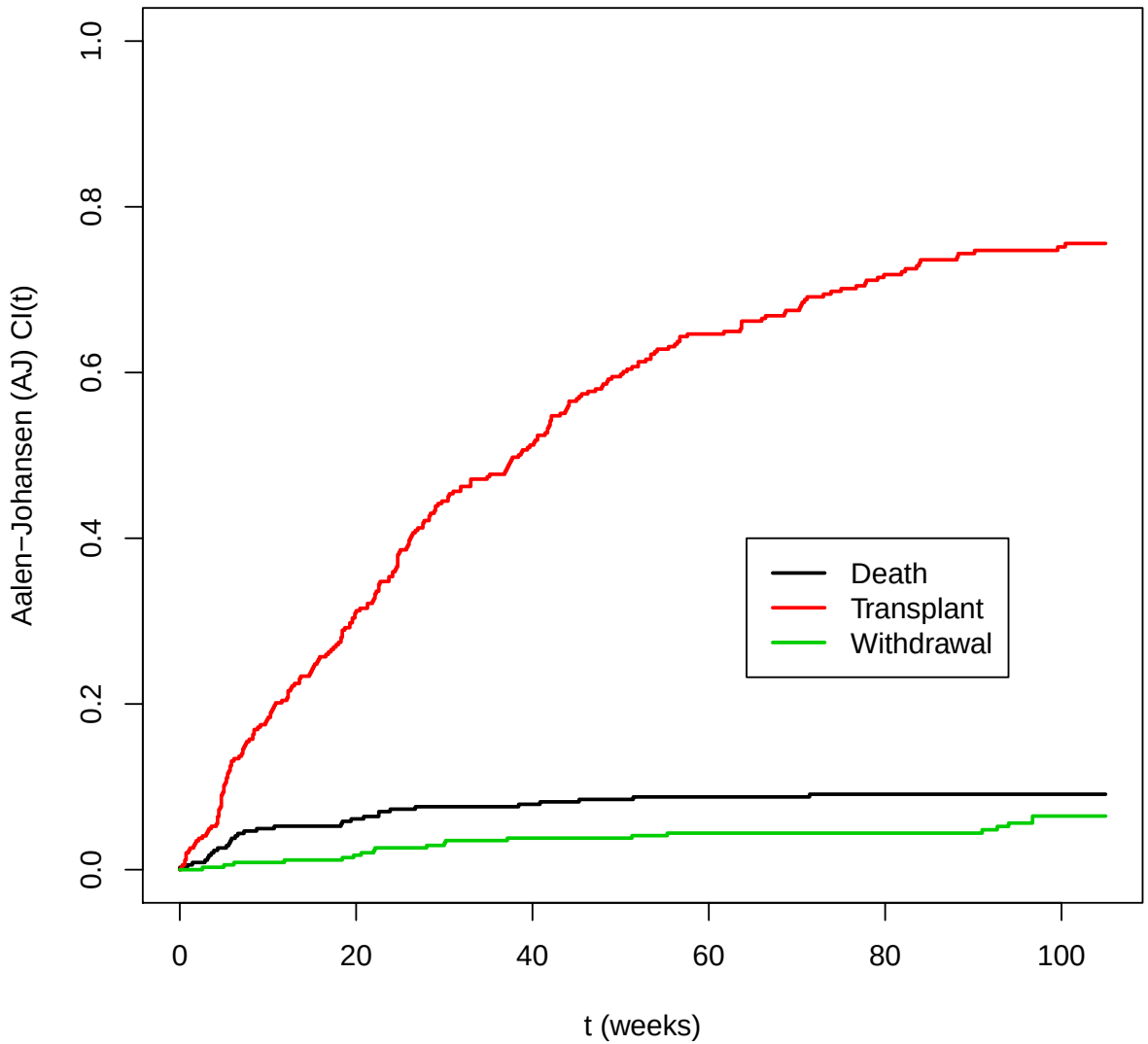
$$= \sum_{j=1}^{j'} S(t_{(j-1)}|x_i(t_{(j-1)}))p_{ij}\pi_{ik}(t_{(j)}|x_{ij})$$

Liver transplant

Kim et al. 2006 *Hepatology*

- Mayo Clinic Liver transplant waiting list data from 1990-1999
- During this period, liver allocation policy was flawed
- Donor livers from subjects with blood type O can be used by patients with A, B, AB or O blood types, whereas an A, B, AB liver can only be used by an A, B, AB recipient respectively
- Type O subjects on the waiting list were at a disadvantage since the pool of competitors was larger for type O donor livers
- Current policies have evolved and now depend on each individual patient's risk and need which are assessed and updated regularly while a patient is on the waiting list
- However, the overall donor liver shortage remains acute today
- `transplant` data set in **BART**/**BART3** packages: $N = 815$
- `system.file('demo/liver.crisk.bart.R', package='BART')`

Liver transplant Competing Risks for Type O patients
Aalen-Johansen estimator



Liver transplant Competing Risks for Type O patients
Aalen-Johansen vs. BART