Bayesian Nonparametric Statistics for Fostering Innovation and Discovery in Biomedical Research

Yanxun Xu
Department of Applied Math and Statistics
Mathematics Institute for Data Science
Division of Biostatistics and Bioinformatics
The Sidney Kimmel Comprehensive Cancer Center
Johns Hopkins University

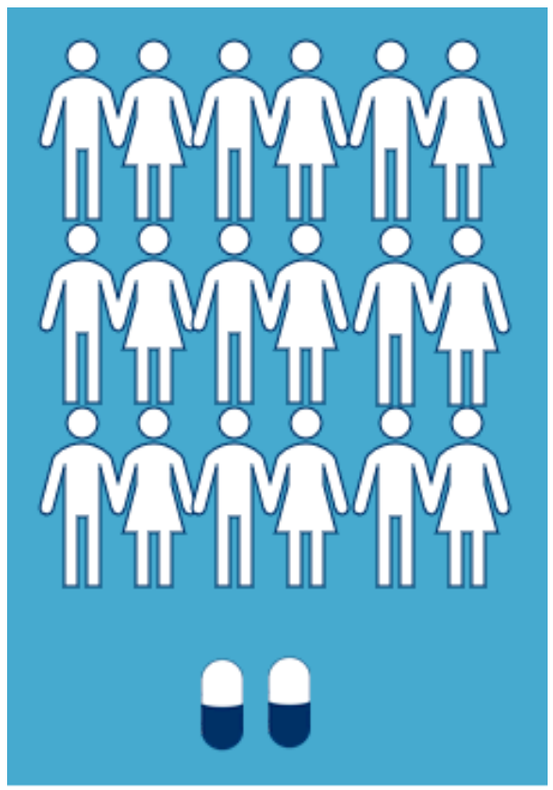
April, BNP 2022

Outline

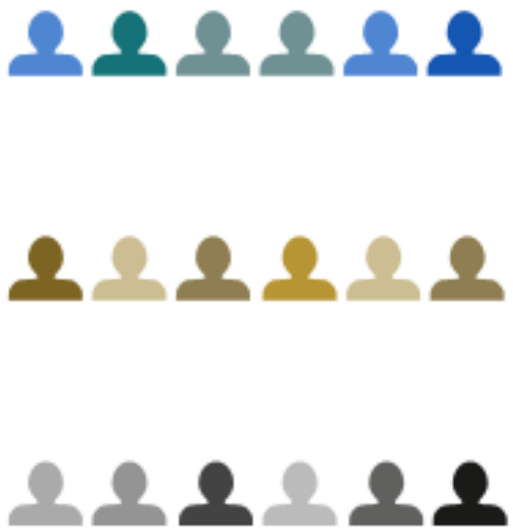
- **Part 1:** *Monday*
  - **Density estimation** for efficient clinical trial designs
  - **Regression** for precision dosing
- **Part 2:** *Wednesday*
  - **Clustering** for subgroup finding
  - **Latent feature models** for tumor heterogeneity
- **Part 3:** *Friday*
  - Estimating treatment effects from observational data

Treatment Effect

Clinical Trials



Biomarker-based Trials



Observational Data

Population → Subgroup → Personalization

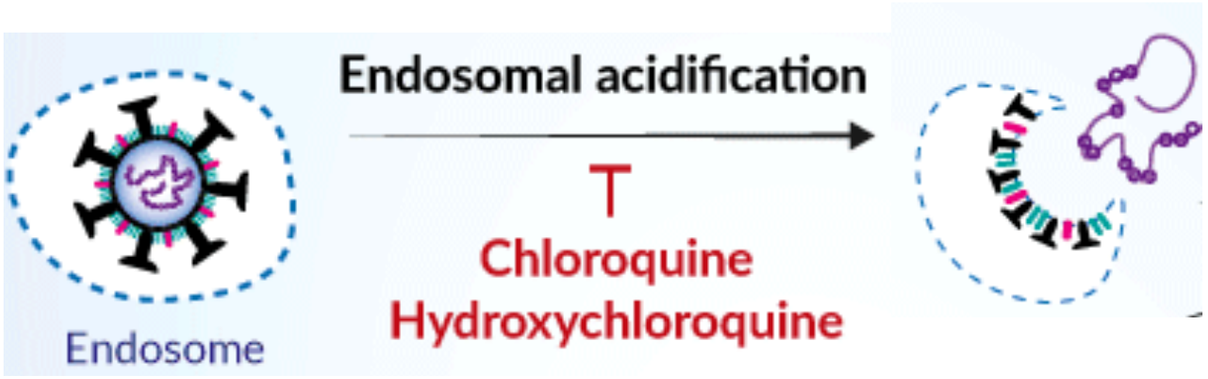
Outline

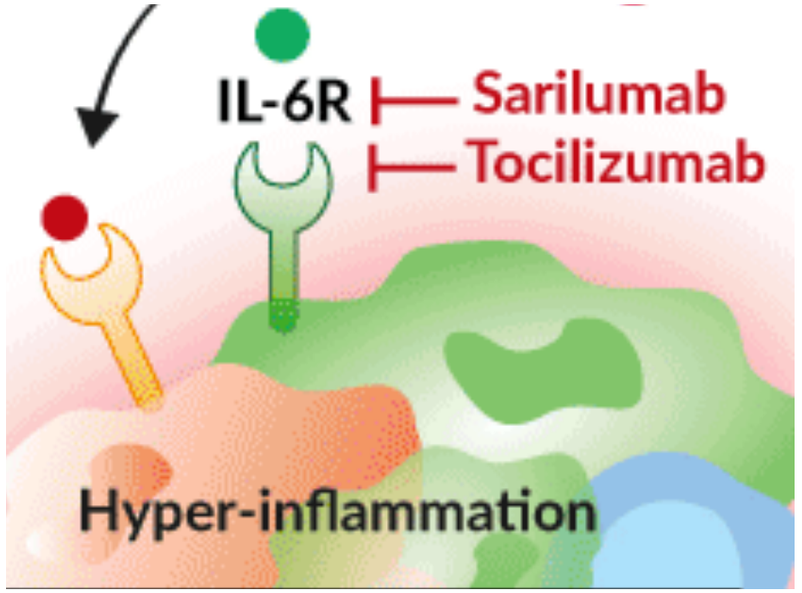
Part 3: Estimating treatment effects from observational data

- Single stage treatment
- Dynamic treatment regimens (multiple stage treatments)
- Treatments in continuous time
- Connection to offline reinforcement learning


Wearable Technology
Healthcare

Repurposed Drugs

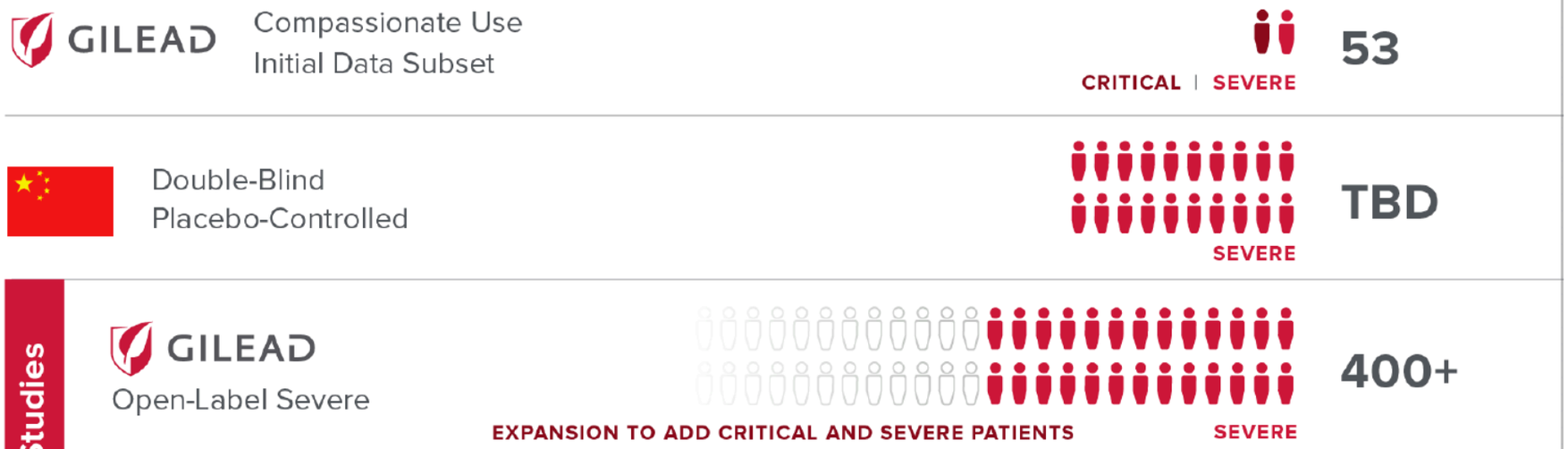




Corticosteroids (e.g., dexamethasone)

Clinical Trials for Remdesivir

DATA SOURCE		NUMBER OF PATIENTS
GILEAD	Compassionate Use Initial Data Subset	53
	Double-Blind Placebo-Controlled	TBD
Simple Studies: GILEAD	Open-Label Severe	400+
Simple Studies: GILEAD	Open-Label Moderate	600+
NIH NIAID	Double-Blind Placebo-Controlled	800+
Other Trials: World Health Organization, Inserm		10,000+



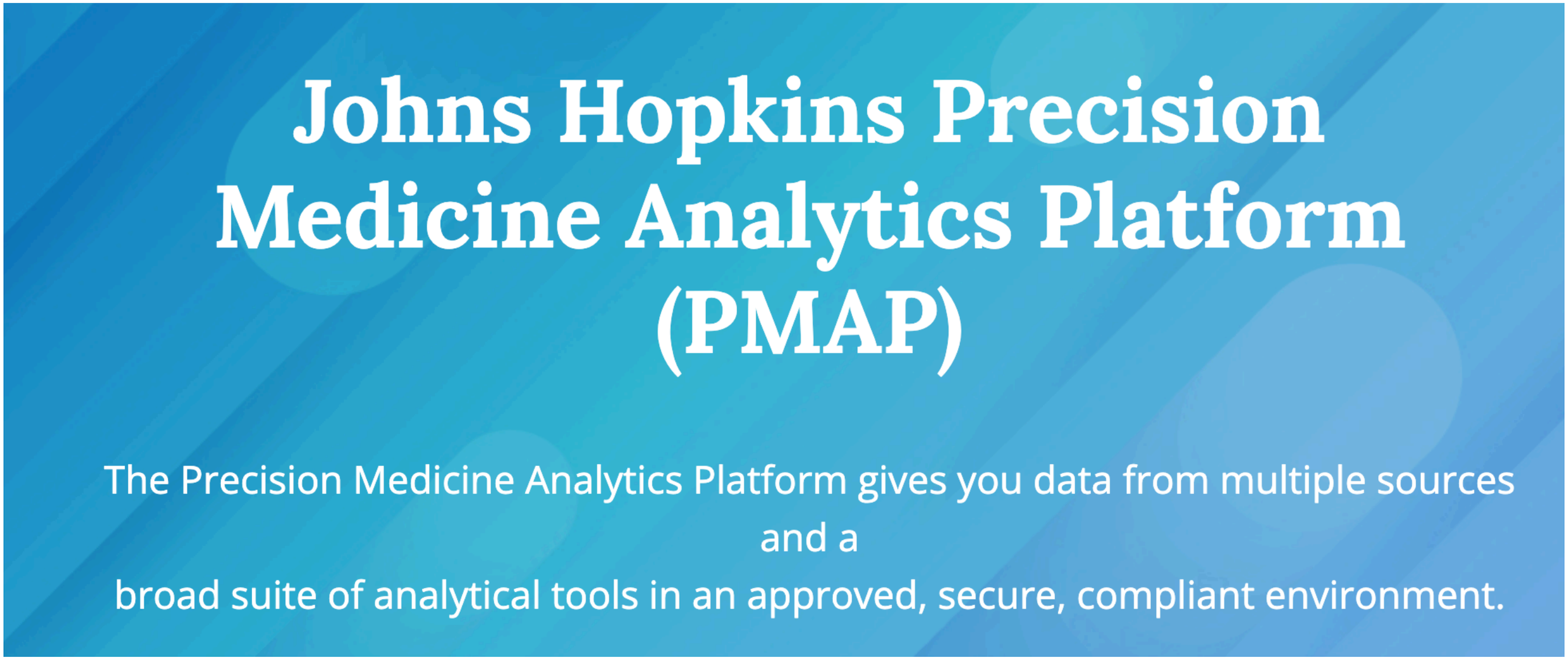
Critical Intubated

Severe Requiring Oxygen

Moderate Hospitalized Patients

No Definitive Conclusions!

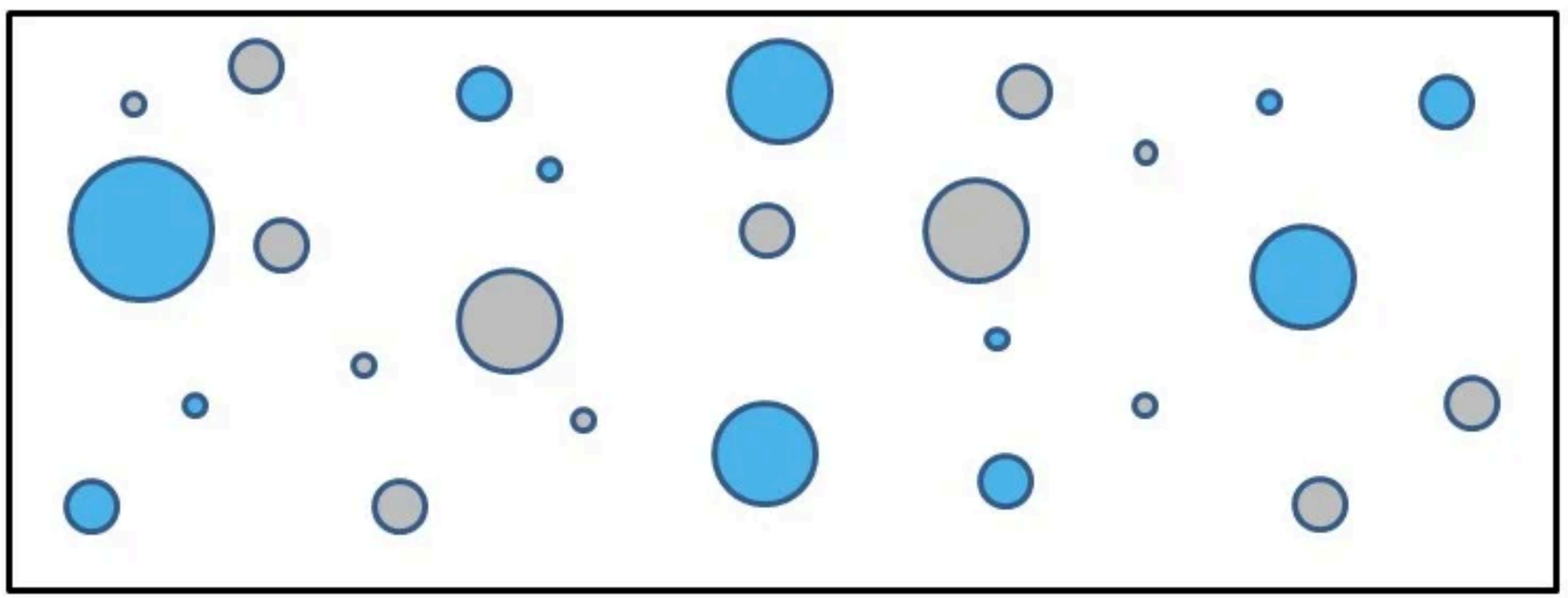
Estimating the Effect of Remdesivir from Real World Data



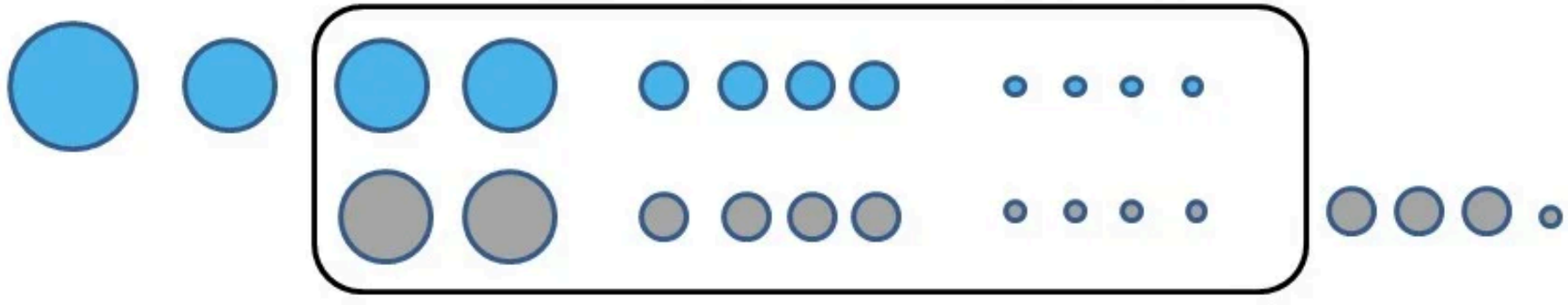
	All Remdesivir (n = 342)	All Control (n = 1957)
Demographics:		
Male	189 (55.3%)	1004 (51.3%)
Race Black	124 (36.3%)	715 (36.5%)
Race Latinx	114 (33.3%)	519 (26.5%)
Race White	66 (19.3%)	534 (27.3%)
Race Others	38 (11.1%)	189 (9.7%)
Age, Median (IQR)	60 (11.5)	60 (15)
BMI, Median (IQR)	30.1 (5.2)	28.2 (4.5)
DNR/DNI, no. (%)	61 (17.8%)	435 (22.2%)
O2 Devices, no. (%):		
No Supplemental Oxygen	16 (4.7%)	907 (46.3%)
Nasal Cannula or Face Mask	210 (61.4%)	819 (41.8%)
High Flow Nasal Cannula	60 (17.5%)	79 (4%)
Noninvasive Positive-Pressure Ventilation	5 (1.5%)	34 (1.7%)
Mechanical Ventilator	51 (14.9%)	105 (5.4%)

Correct for Assignment Bias

Population with varying characteristics



Treatment Control



Causal Estimands

Average treatment effect:

$$\Delta = \mathbb{E}[Y(1) - Y(0)]$$

$Y(z)$ is the potential outcome under z, $z = 0,1$

Stable Unit Treatment Value Assumption (SUTVA)

$$Y_i(Z_i) \perp\!\!\!\perp Z_j$$

Assumptions

- **Consistency**

$$Y = ZY(Z) + (1 - Z)Y(1 - Z)$$

- **Positivity**

$$0 < \Pr(Z = 1 \mid X, Y(0), Y(1)) < 1$$

- **No unmeasured confounders assumption (NUCA)**

$$\Pr(Z = 1 \mid X, Y(0), Y(1)) = \Pr(Z = 1 \mid X)$$

Average treatment effect:

$$\Delta = E\{\mu_1(X) - \mu_0(X)\}$$

$$= E\{\frac{ZY}{e(X)} - \frac{(1-Z)Y}{1-e(X)}\}$$

$$\mu_z(X) = E(Y \mid Z = z, X)$$

Propensity score: $e(x) = \Pr(Z = 1 \mid X = x)$

Estimators

- **Outcome model**

$$\hat{\Delta}_O = \frac{1}{n}\sum_{i=1}^{n}\{\hat{\mu}_1(X_i) - \hat{\mu}_0(X_i)\}$$

- **Inverse probability weighting (IPW)**

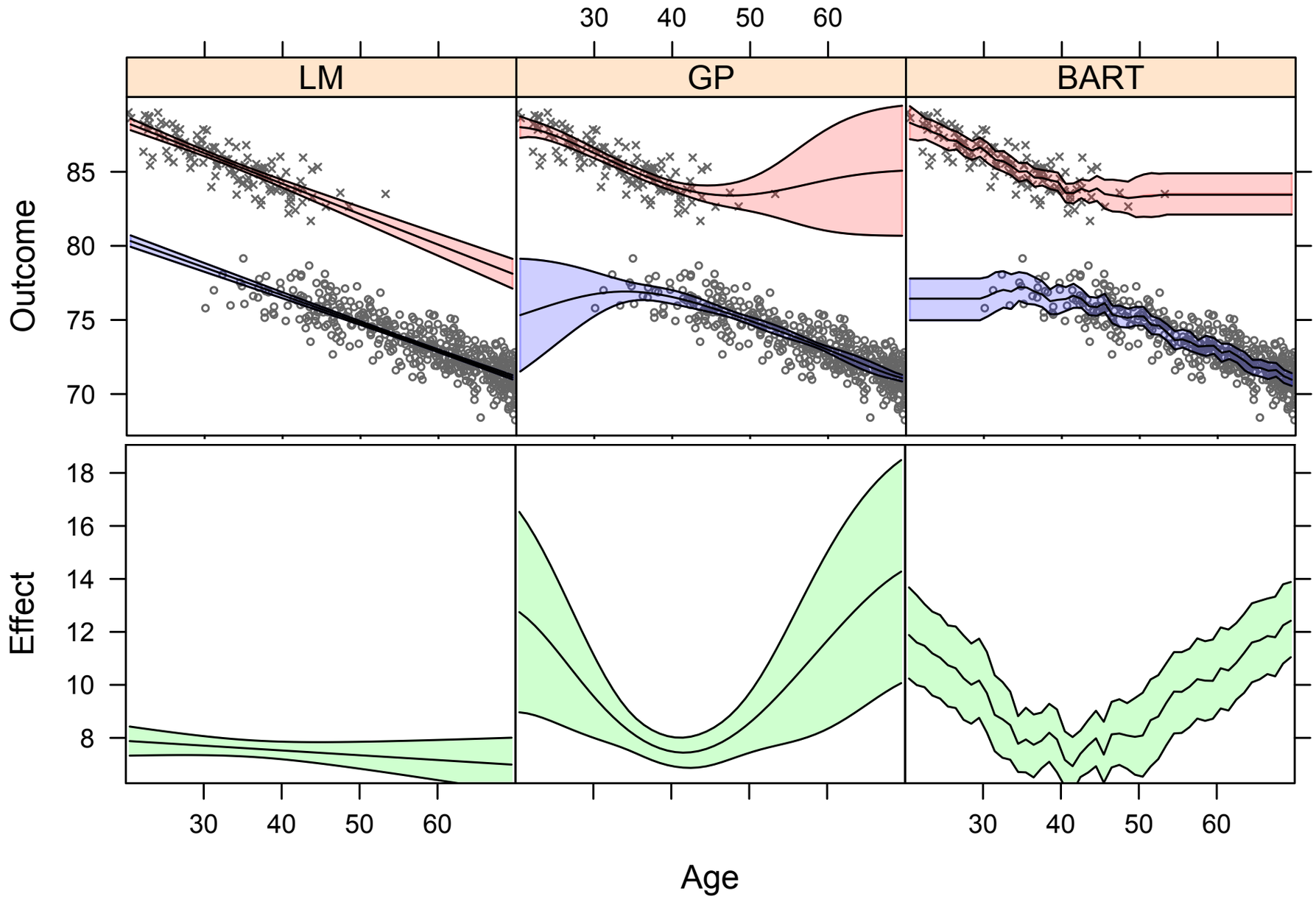
$$\hat{\Delta}_{ipw} = \frac{\sum_{i=1}^{n} Z_i Y_i/\hat{e}(X_i)}{\sum_{i=1}^{n} Z_i} - \frac{\sum_{i=1}^{n}(1-Z_i)Y_i/(1-\hat{e}(X_i))}{\sum_{i=1}^{n}(1-Z_i)}$$

BNP Methods

- **Outcome model**
  - BART (Hill, 2011)
  - Dirichlet process mixture (Kim et al, 2017)
  - Gaussian process (Roy et al.)
- **Inverse probability weighting (IPW)**
  - Pitman-Yor process pior (Karabatsos and Walker, 2011)
  - BART (Hahn et al., 2020)

Choice of Prior

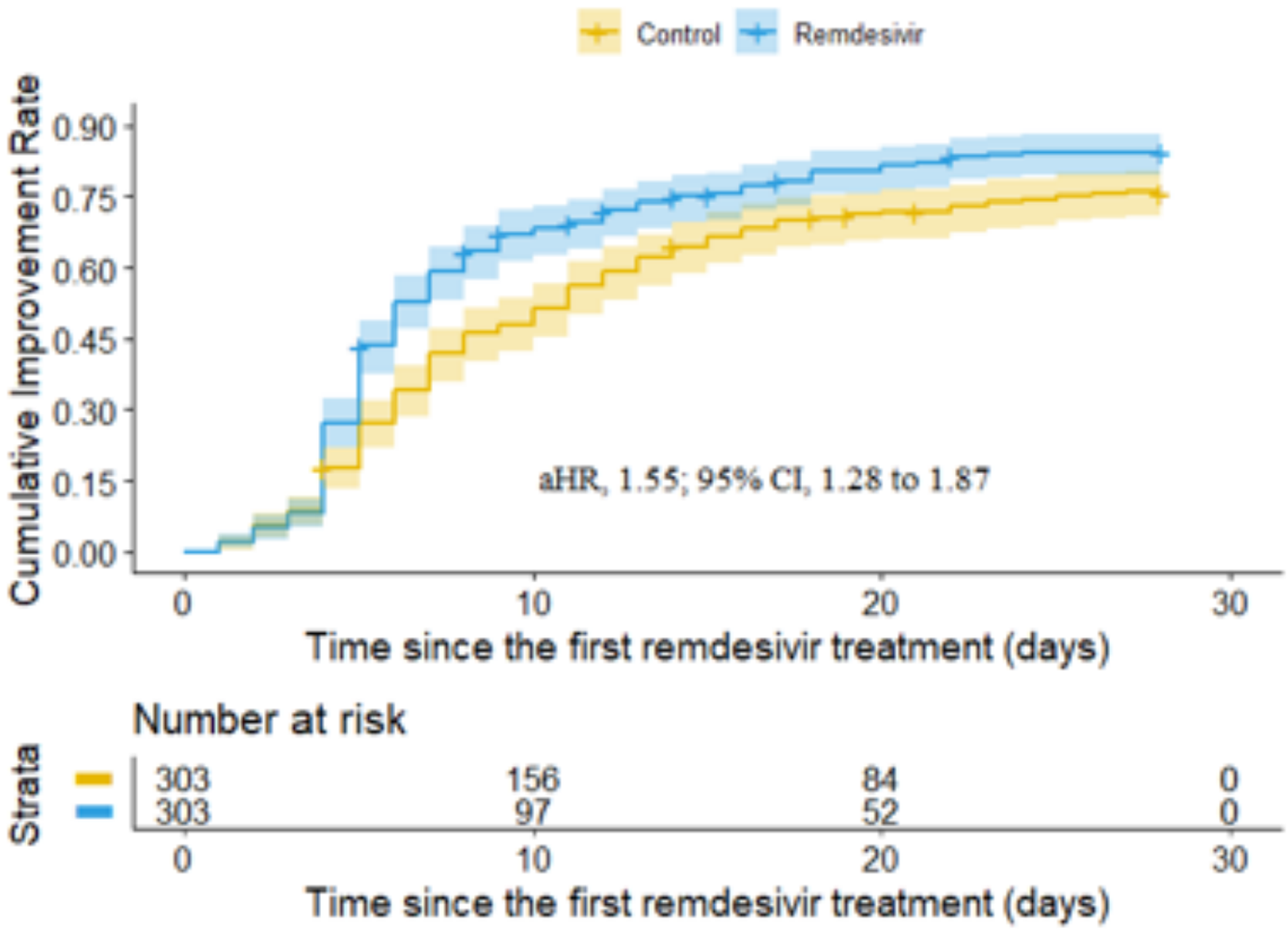
Papadogeorgou and Li, 2020



For causal inference (or anything), being Bayesian should be a tool, not a goal. —Fan Li

Effectiveness results

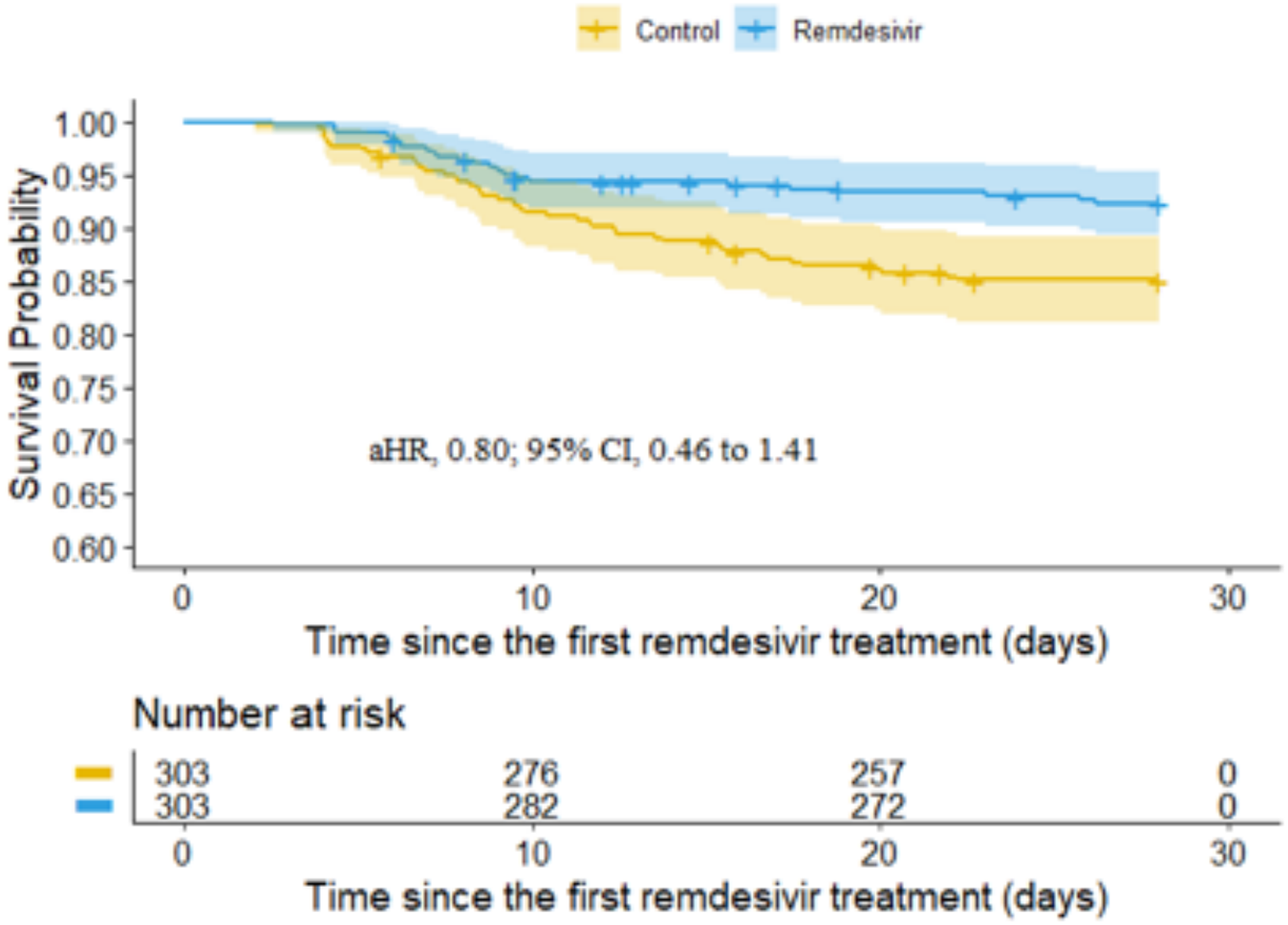
- **Primary outcome:** Time to clinical improvement
- **Result:** Remdesivir **had benefits** in time to clinical improvement with aHR=1.55, p<1e-05, 95% CI: 1.28-1.87



Garibaldi et al., JAMA Network Open, 2021

Effectiveness results

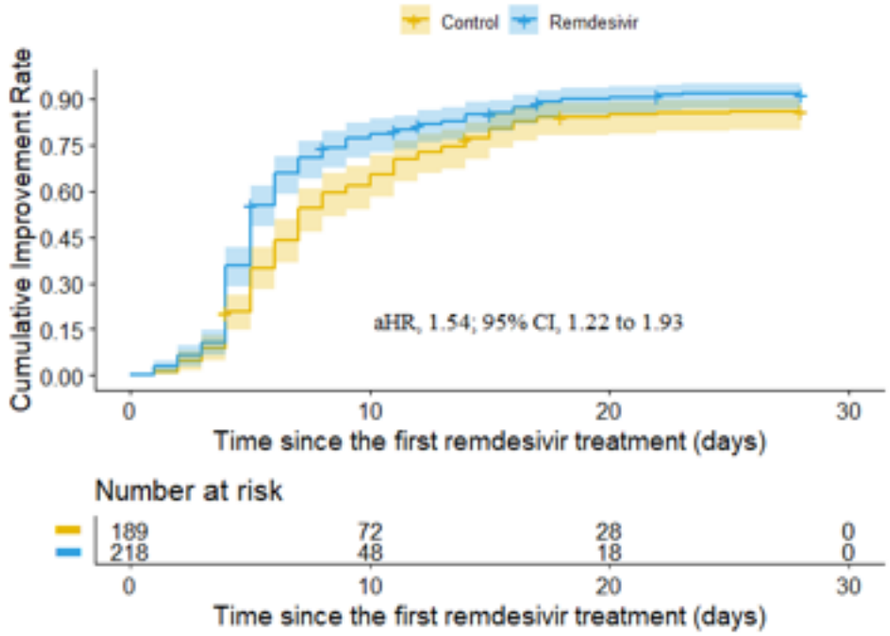
- **Secondary outcome:** Time to death
- **Results:** not statistically significant with aHR=0.8, p=0.44, 95% CI: 0.46-1.41



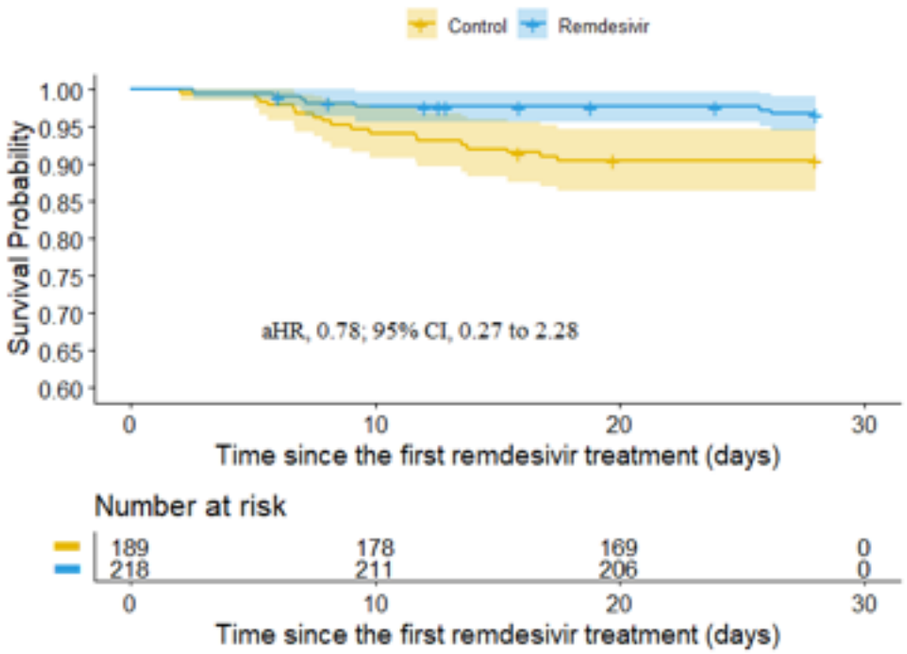
Subgroup analysis stratified by severity

Mild/Moderate

Time to clinical improvement:
aHR 1.54, 95% CI: 1.22-1.93



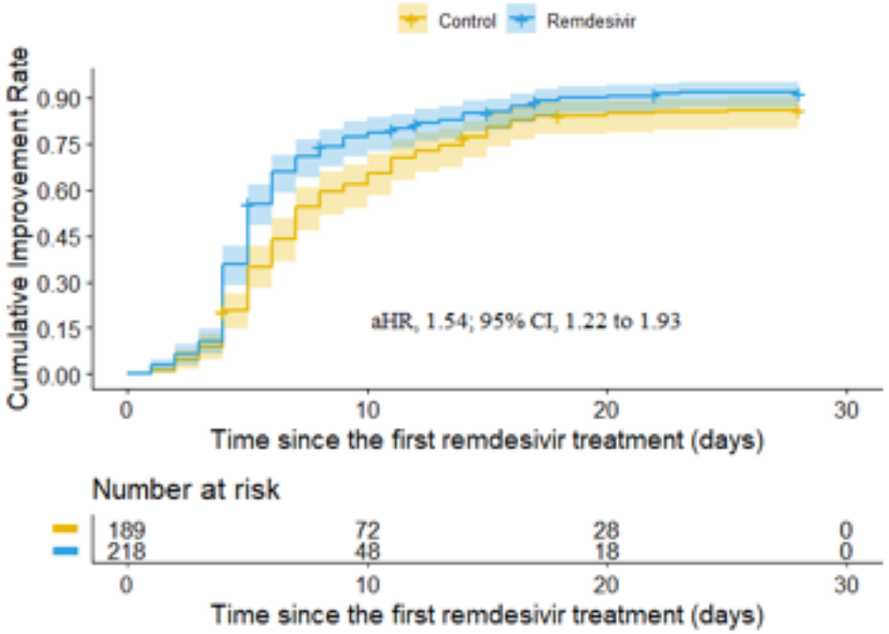
Time to death: aHR 0.78, 95% CI: 0.27-2.28



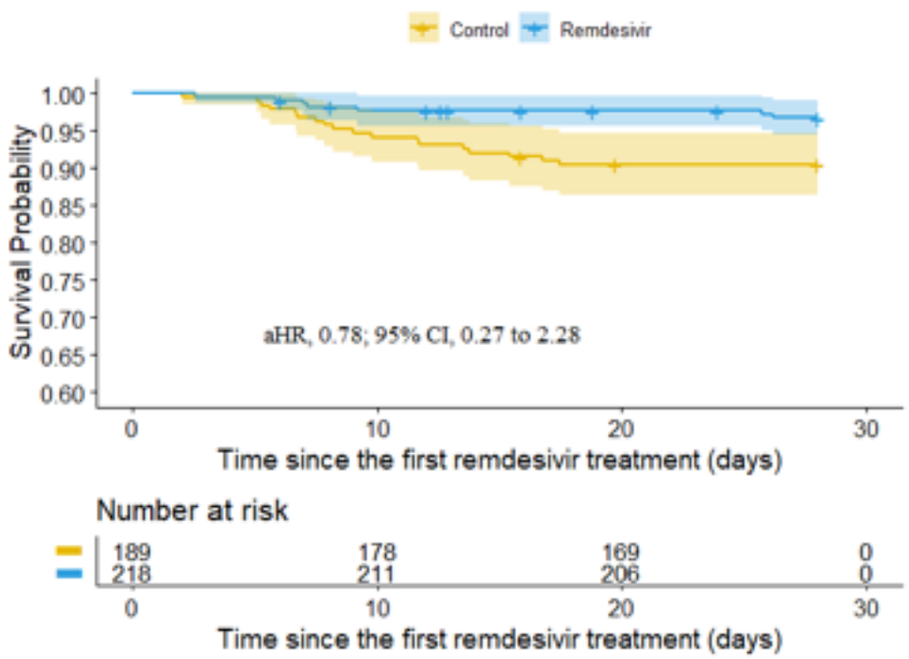
Treat early!!

Severe

Time to clinical improvement:
aHR 1.39, 95% CI: 0.91-2.11



Time to death: aHR 0.94, 95% CI: 0.43-2.03



Outline

Part 3: Estimating treatment effects from observational data

- Single stage treatment
- Dynamic treatment regimens (multiple stage treatments)
- Treatments in continuous time
- Connection to offline reinforcement learning

Motivation: Acute Leukemia Trial

Frontline: “Remission Induction”

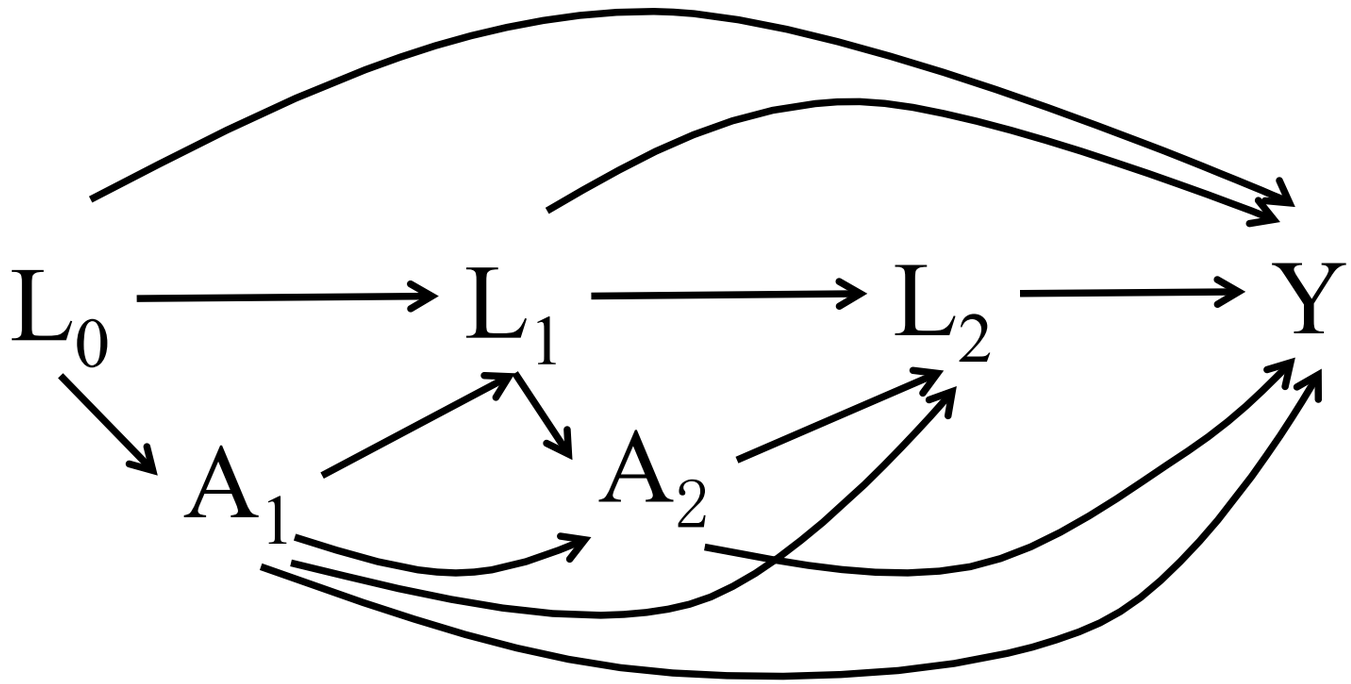
- At the start: chemotherapy, to achieve CR.
  - Less than 5% blastic blood cells, and none with leukemic phenotype
  - Platelet count > $10^5 / \mu L$
  - WBC count > $10^3 / \mu L$
- Patients may 1)die while in Induction, 2) resistant to frontline, or 3) relapse after CR.

Salvage

Dynamic Treatment Regimens

K stages for one individual

$$L_0, A_1, L_1, \ldots, A_K, L_K, Y$$



Time-varying confounding: doctors use the measurement of a variable (L_{k-1}) to **determine whether or not to treat** (A_k) which **affects** the variable's value (L_k) **at a subsequent time**.

Dynamic Treatment Regimens

Denote $H_j = (L_0, A_1, L_1, \ldots, A_j, L_j)$

- The dynamic treatment regimen is the sequence of decision rules:

$$d_1(H_0), d_2(H_1), \ldots, d_K(H_{K-1})$$

- Give a dynamic treatment regimen, we can employ the actions determined by decision rules

$$a_1 = d_1(H_0), a_2 = d_2(H_1), \ldots, a_K = d_K(H_{K-1})$$

Goal: find decision rules that maximize the expected cumulative reward.

Assumptions

- **Consistency**

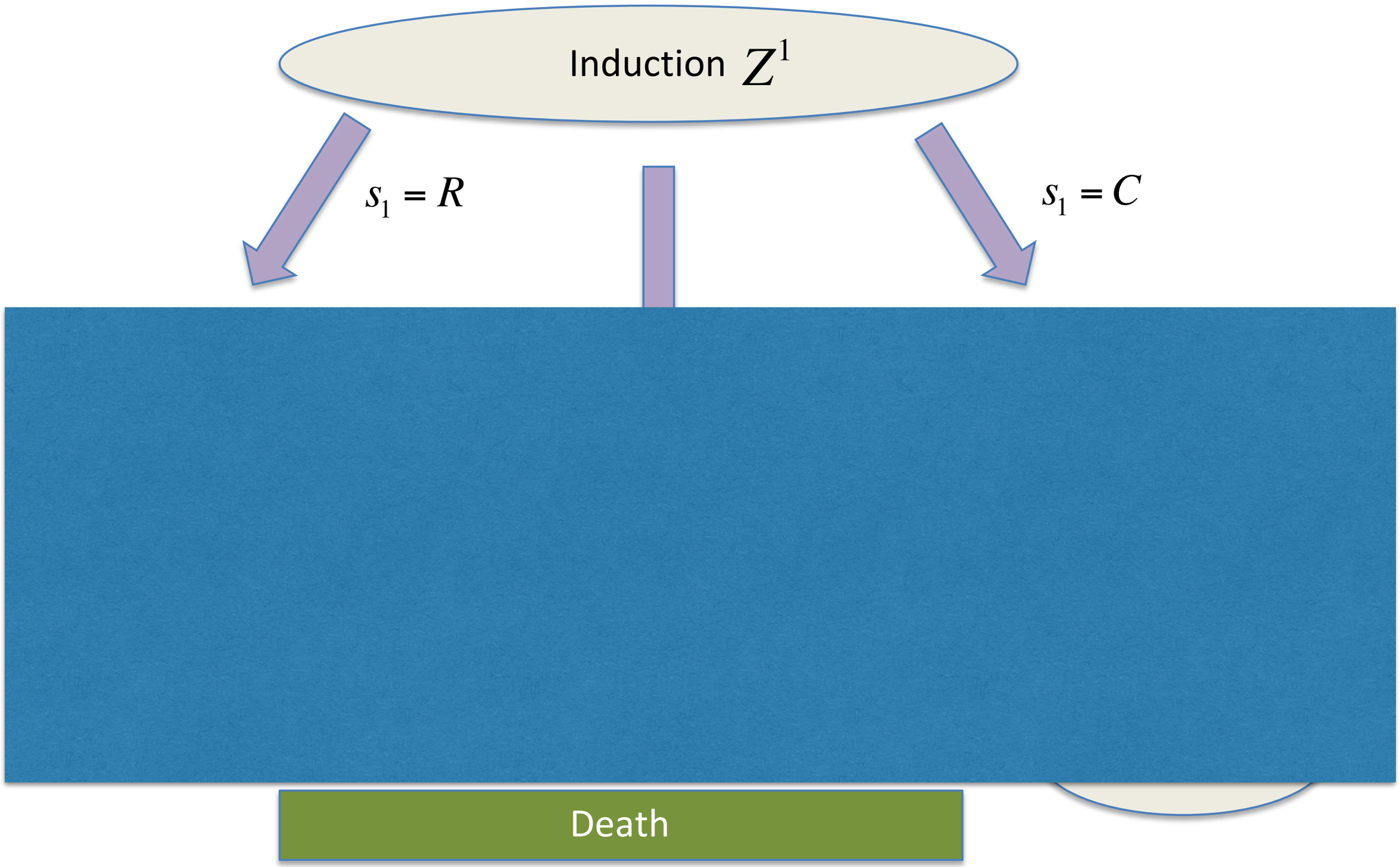
$$L_j = \sum_{\bar{a}_{j-1} \in \bar{\mathcal{A}}_{j-1}} L_j^*(\bar{a}_{j-1}) I(\bar{A}_{j-1} = \bar{a}_{j-1}), j = 1, \ldots, K$$

$$Y = \sum_{\bar{a}_K \in \bar{\mathcal{A}}_K} Y^*(\bar{a}) I(\bar{A} = \bar{a})$$

- **Positivity**

- **No unmeasured confounders assumption (NUCA)**

- **Dynamic treatment regimens:** G-computation (Robins, 1986), G-estimation of structural nested models (Robins, 2004), IPTW (van der Laan and Petersen, 2007), doubly robust IPTW (Tsiatis, 2007; Zhao et al., 2015).
- **BNP:**
  - DDP-GP in the context of G-computation (Xu et al., 2017)
  - DP mixture in the context of policy search (Quan et al., 2020) and in the context of G-computation,
  - BART in the context of Q learning (Murray et al., 2017)


Induction Z^1
$s_1 = R$
$s_1 = C$
Death

Dynamic Treatment Regimens

Regime (A, B_1, B_2)

- Treat with induction therapy A
- If the disease is resistant to A then give salvage B_1
- If relapse occurs after achieving CR then give salvage B_2.

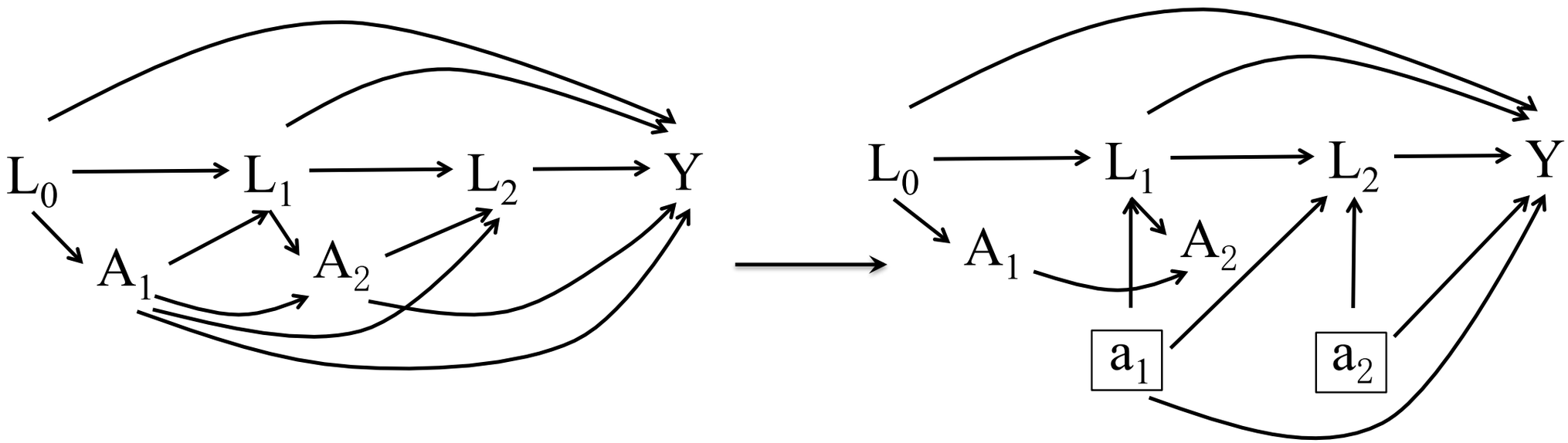
Regimes in the AML/MDS Trial

A total of 16 treatment regimes (a, b_1, b_2)
Induction: $a \in \{\text{FAI}, \ \text{FAI+G}, \ \text{FAI+ATRA}, \ \text{FAI+G+ATRA}\}$
Salvage: $b_1, b_2 \in \{\text{HDAC}, \ \text{OTHER}\}$

The 16 Actual Dynamic Treatment Regimes in the AML/MDS Trial

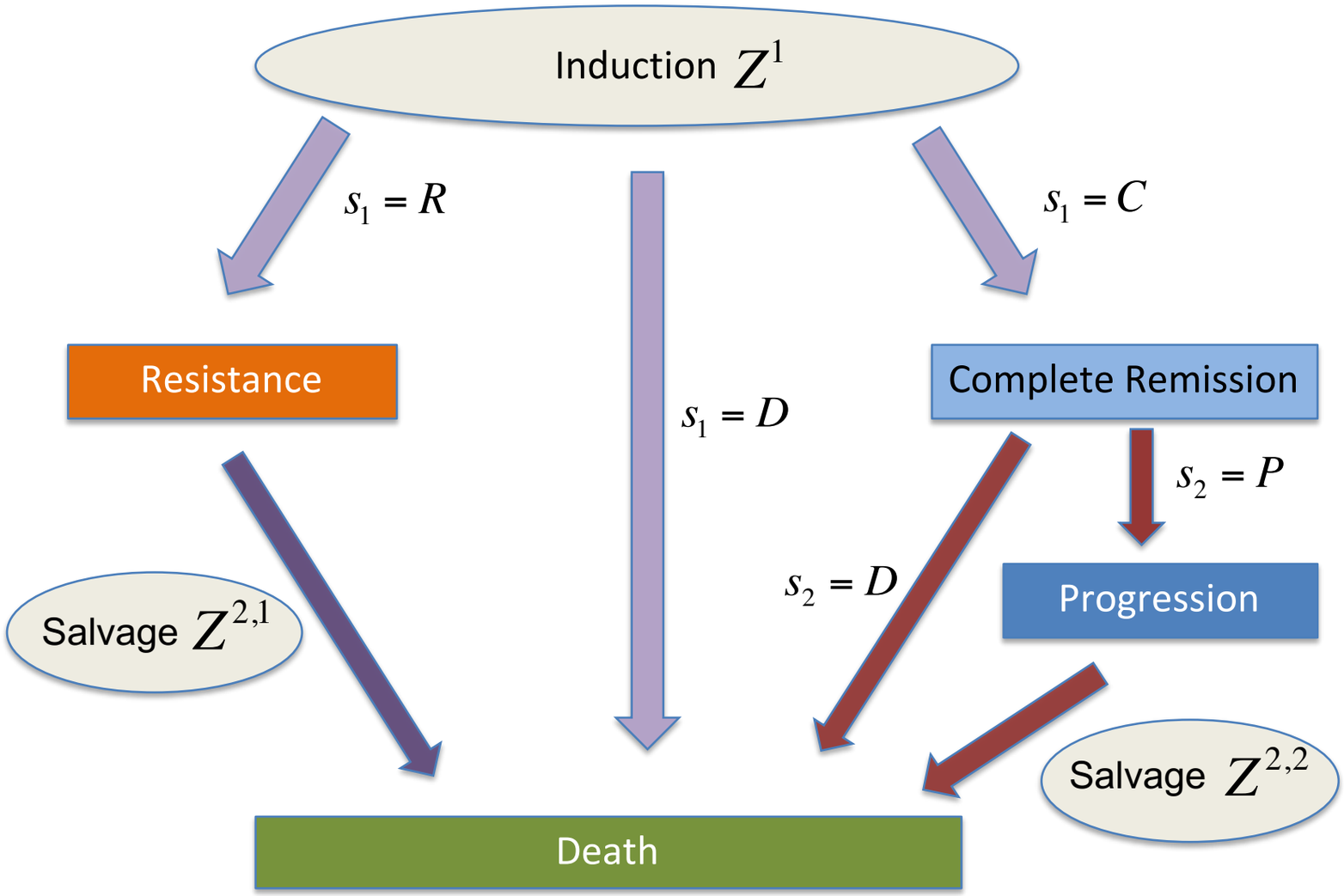
Induction	Salvage for Resistant Disease	Salvage after Progression
FAI	HDAC	HDAC
FAI	HDAC	Other
FAI	Other	HDAC
FAI	Other	Other
FAI + ATRA	HDAC	HDAC
FAI + ATRA	HDAC	Other
FAI + ATRA	Other	HDAC
FAI + ATRA	Other	Other
FAI + G-CSF	HDAC	HDAC
FAI + G-CSF	HDAC	Other
FAI + G-CSF	Other	HDAC
FAI + G-CSF	Other	Other
FAI + G-CSF + ATRA	HDAC	HDAC
FAI + G-CSF + ATRA	HDAC	Other
FAI + G-CSF + ATRA	Other	HDAC
FAI + G-CSF + ATRA	Other	Other

To address this, we use **G-computation formula** (Robins, 1986).



$$p(Y(a_1, a_2)|L_0, L_1, L_2) = p(Y(a_1, a_2)|A_1, A_2, L_0, L_1, L_2)$$
$$= p(Y|A_1 = a_1, A_2 = a_2, L_0, L_1, L_2)$$

Potential outcome

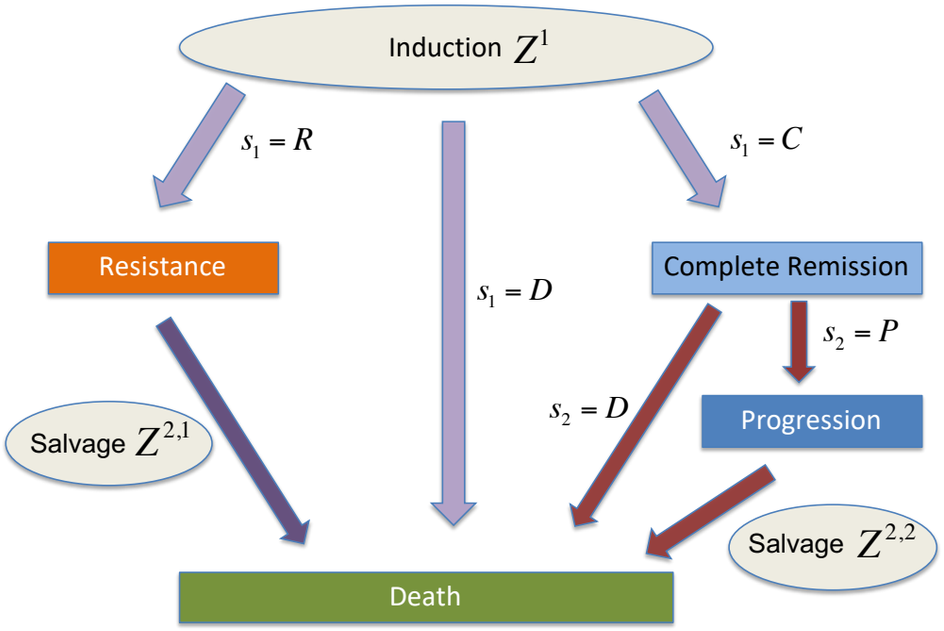


Survival Time =

T^D	if death during induction
$T^R + T^{RD}$	if death after salvage for resistant disease
$T^C + T^{CP} + T^{PD}$	if death after salvage for progression after CR
$T^C + T^{CD}$	if death in CR

Xu, Yanxun, Peter Müller, Abdus S. Wahed, and Peter F. Thall. "Bayesian nonparametric estimation for dynamic treatment regimes with sequential transition times." *Journal of the American Statistical Association* 111, no. 515 (2016): 921-950.

DDP-GP

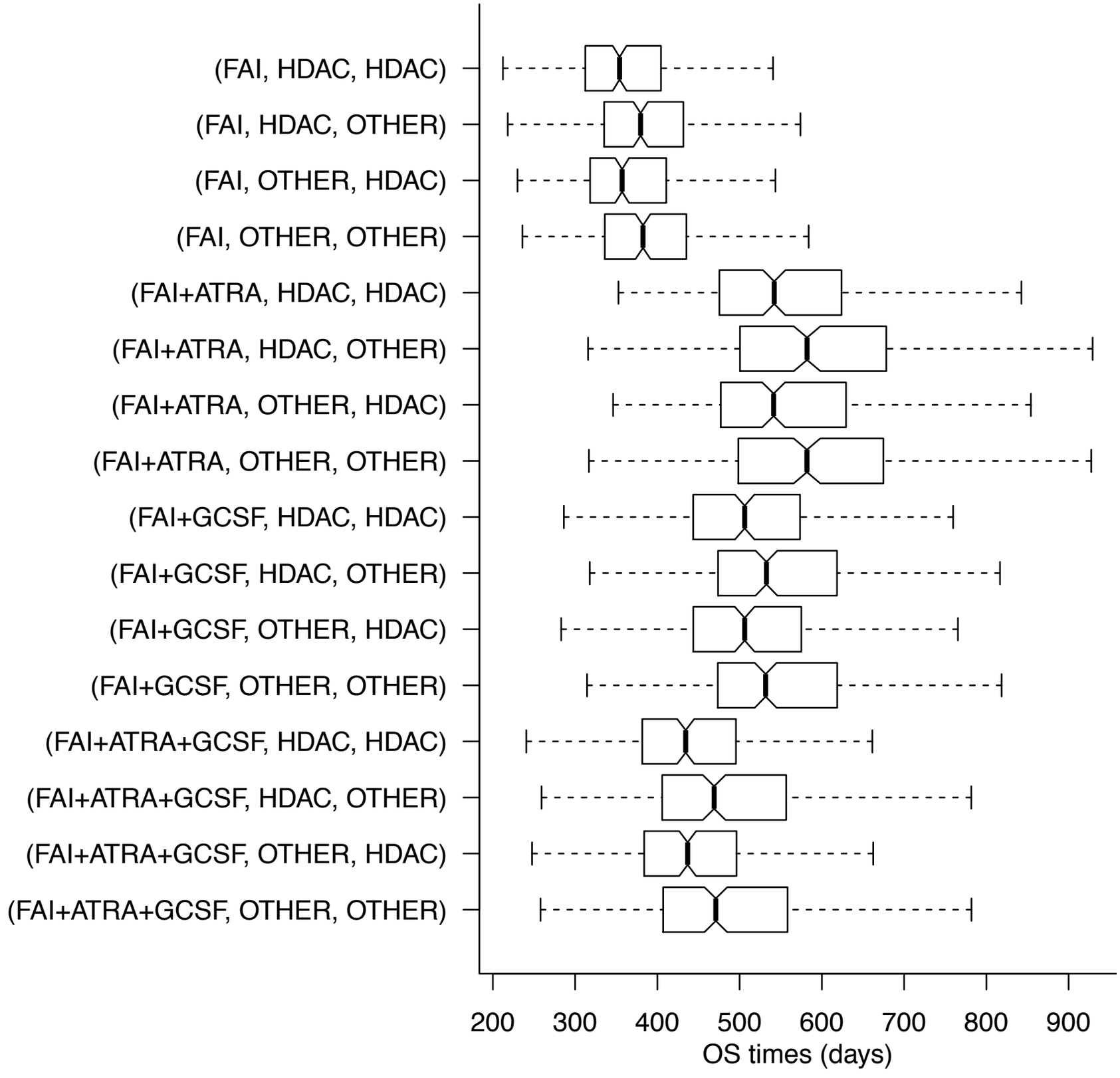


Overall Mean Survival Under Regime (A, B_1, B_2):

$$
\begin{aligned}
\theta(A, B_1, B_2) \;=\; & \int \Big\{ Pr(Z_1 = 0|A, X)\theta^D(A, X) + Pr(Z_1 = 1|A, X) \Big[\theta^R(A, X) \\
+ \; & \int \theta^{RD}(A, B_1, X, X^{(R)}) d\mu(X^{(R)}) \Big] \\
+ \; & Pr(Z_1 = 2|A, X) \Big\{ \theta^C(A, X) + \int \Big[Pr(Z_2 = 0|Z_1 = 2, A, X, X^{(C)}) \\
\times \; & \theta^{CD}(A, X, X^{(C)}) + Pr(Z_2 = 1|Z_1 = 2, A, X, X^{(C)}) \Big(\theta^{CP}(A, X, X^{(C)}) \\
+ \; & \int \theta^{PD}(A, B_2, X, X^{(C)}, X^{(P)}) d\mu(X^{(P)}) \Big) \Big] d\mu(X^{(C)}) \Big\} \Big\} d\mu(X)
\end{aligned}
$$

Xu, Yanxun, Peter Müller, Abdus S. Wahed, and Peter F. Thall. "Bayesian nonparametric estimation for dynamic treatment regimes with sequential transition times." *Journal of the American Statistical Association* 111, no. 515 (2016): 921-950.

Trial Data Analysis



Conclusion

- FAI + ATRA followed by non-HDAC *at disease progression* after CR seems promising

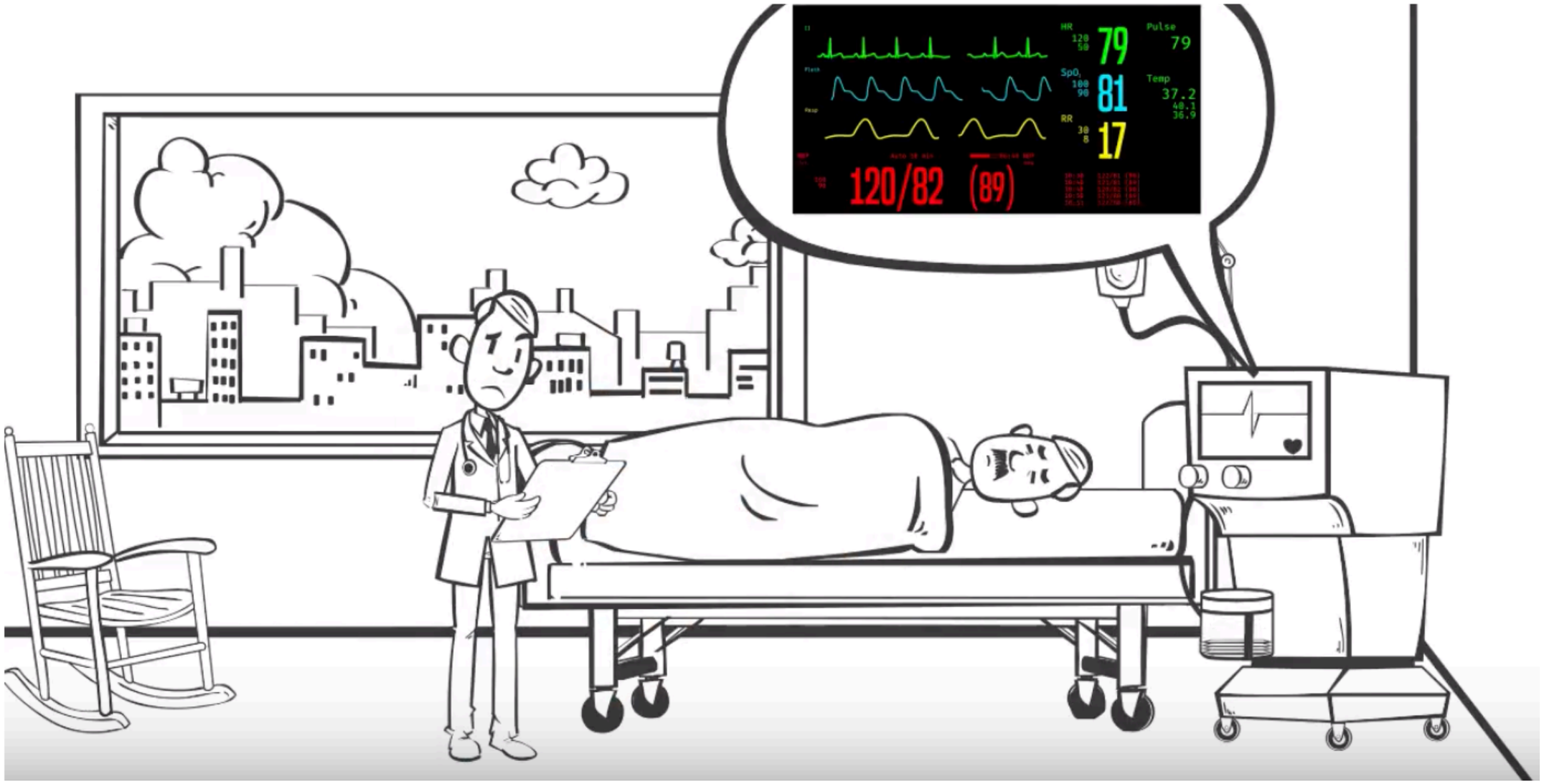
- If we had done this analysis before, ATRA might have been studied further

Outline

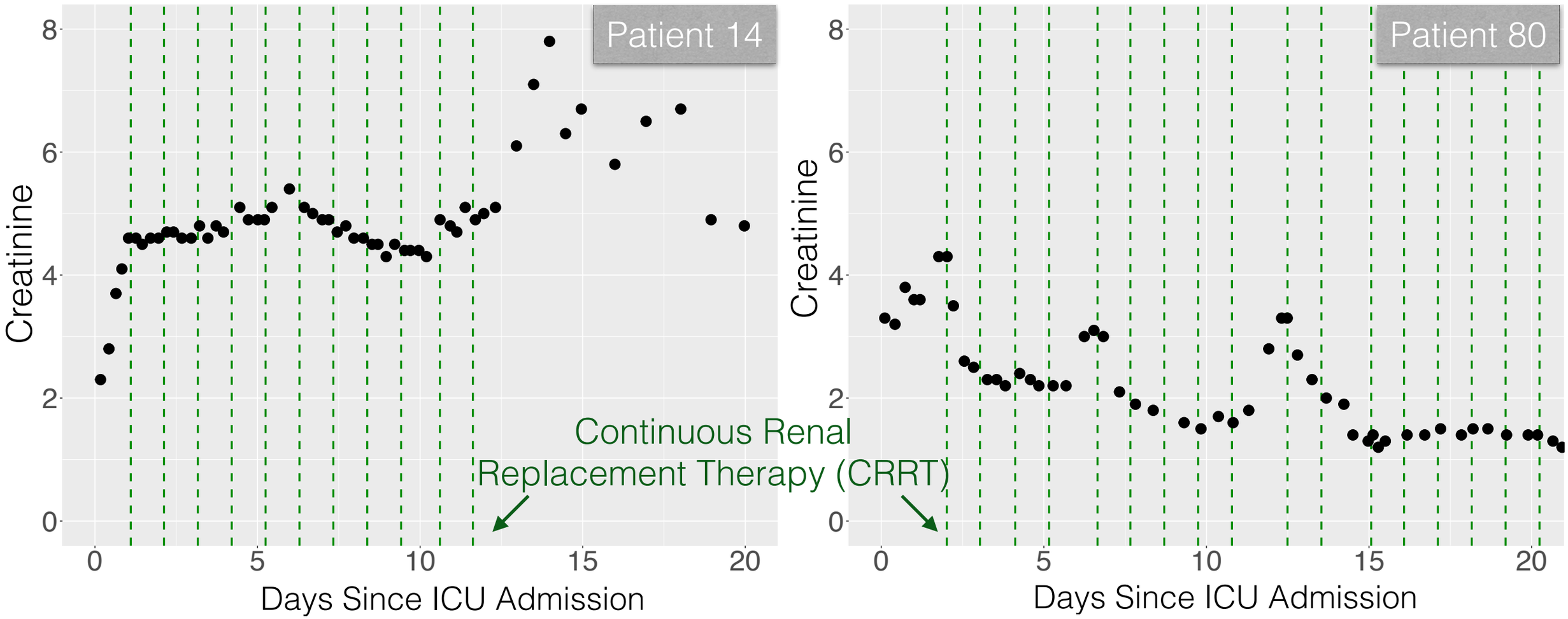
Part 3: Estimating treatment effects from observational data

- Single stage treatment
- Dynamic treatment regimens (multiple stage treatments)
- Treatments in continuous time
- Connection to offline reinforcement learning

Treatment in Continuous Time



Acute Kidney Injury



Our **goal** is to estimate **individual's response over time** from **Electronic Health Record (EHR)** data.

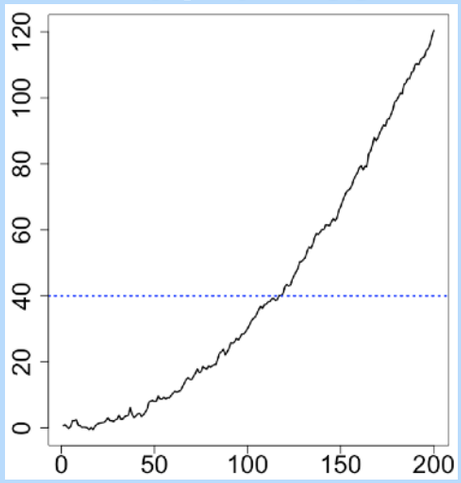
Approach: Individualized treatment response (ITR) model

$$Y_{ij}|\boldsymbol{X}_i, \mathcal{H}_{ij} = \underbrace{b(\boldsymbol{X}_i) + \boldsymbol{u}_i(t_{ij})}_{\text{baseline progression}} + \underbrace{\boldsymbol{f}_i(t_{ij}; \mathcal{H}_{ij})}_{\text{treatment response}} + \underbrace{\boldsymbol{\epsilon}_i(t_{ij}; \mathcal{H}_{ij})}_{\text{noise}}, \; j = 1, ..., J_i.$$

Approach: Individualized treatment response (ITR) model

$$Y_{ij}|\boldsymbol{X}_i, \mathcal{H}_{ij} = \underbrace{b(\boldsymbol{X}_i) + \boldsymbol{u}_i(t_{ij})}_{\text{baseline progression}} + \underbrace{\boldsymbol{f}_i(t_{ij}; \mathcal{H}_{ij})}_{\text{treatment response}} + \underbrace{\boldsymbol{\epsilon}_i(t_{ij}; \mathcal{H}_{ij})}_{\text{noise}}, \ j = 1, ..., J_i.$$

Longitudinal normal dynamics



$$b(X_{ij}; \boldsymbol{\beta}_i) = X_{ij}^T \boldsymbol{\beta}_i = X_{i0}^T \boldsymbol{\beta}_{i0} + X_{i1}(t_{ij})^T \boldsymbol{\beta}_{i1}$$

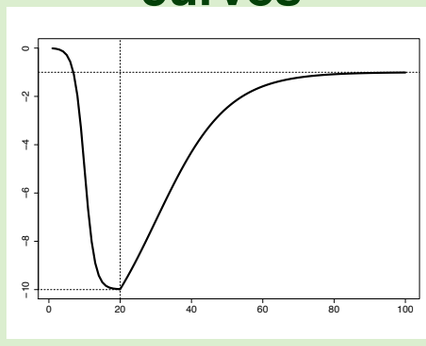
$$\mathbf{u}_i = GP(0, \ \mathcal{K}_{ui})$$

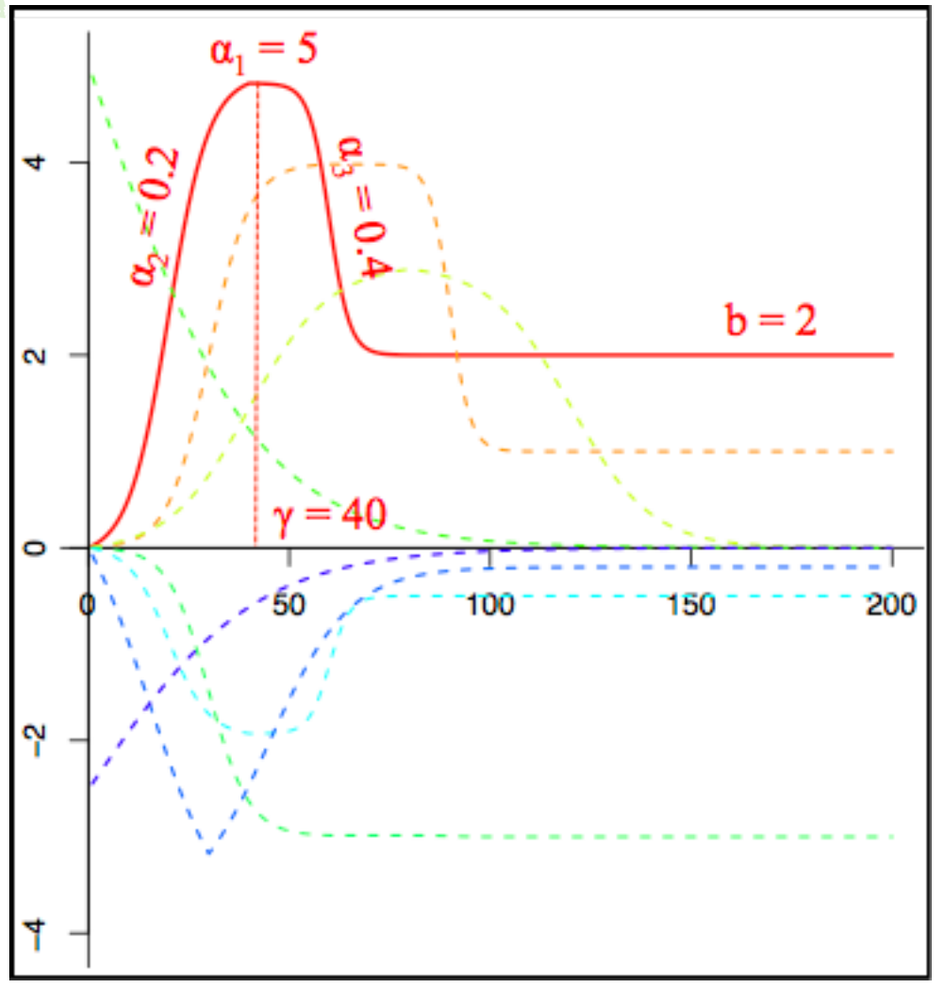
$$\mathcal{K}_{ui}(\sigma_{ui}^2, \rho_{ui}) = Cov(\boldsymbol{u}_i(t_{ij}), \boldsymbol{u}_i(t_{ij'})) = \sigma_{ui}^2 \rho_{ui}^{|t_{ij} - t_{ij'}|}.$$

Approach: Individualized treatment response (ITR) model

$$Y_{ij}|\boldsymbol{X}_i, \mathcal{H}_{ij} = \underbrace{b(\boldsymbol{X}_i) + \boldsymbol{u}_i(t_{ij})}_{\text{baseline progression}} + \underbrace{\boldsymbol{f}_i(t_{ij}; \mathcal{H}_{ij})}_{\text{treatment response}} + \underbrace{\boldsymbol{\epsilon}_i(t_{ij}; \mathcal{H}_{ij})}_{\text{noise}}, \ j = 1, ..., J_i.$$

Treatment-response curves





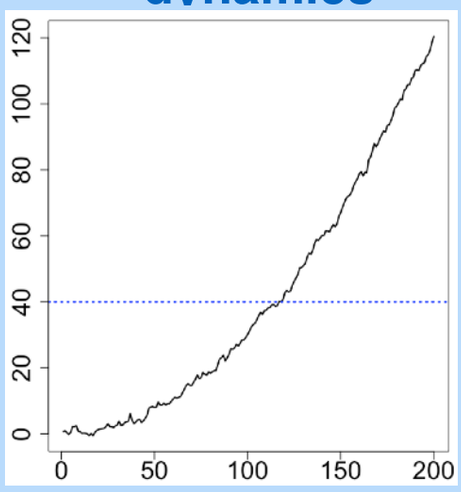
$$\boldsymbol{f}_i(t_{ij}; A_{i,<t_{ij}}) = \sum_{l:\tau_{il}<t_{ij}} g_{i,A_{il}}(t_{ij} - \tau_{il})$$

$$g_{id}(t) = \begin{cases} b_0 + \alpha_{1_{id}}/[1 + \exp(-\alpha_{2_{id}}(t - \gamma_{id}/2))], & \text{if } 0 \le t < \gamma_{id} \\ b_{id} + \alpha_0/[1 + \exp(\alpha_{3_{id}}(t - 3\gamma_{id}/2))], & \text{if } t \ge \gamma_{id}, \end{cases}$$

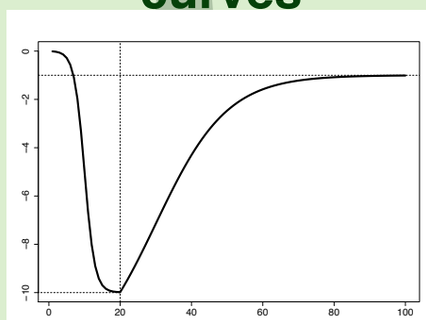
Approach: Individualized treatment response (ITR) model

$$Y_{ij}|\boldsymbol{X}_i, \mathcal{H}_{ij} = \underbrace{b(\boldsymbol{X}_i) + \boldsymbol{u}_i(t_{ij})}_{\text{baseline progression}} + \underbrace{\boldsymbol{f}_i(t_{ij}; \mathcal{H}_{ij})}_{\text{treatment response}} + \underbrace{\boldsymbol{\epsilon}_i(t_{ij}; \mathcal{H}_{ij})}_{\text{noise}}, \ j = 1, ..., J_i.$$

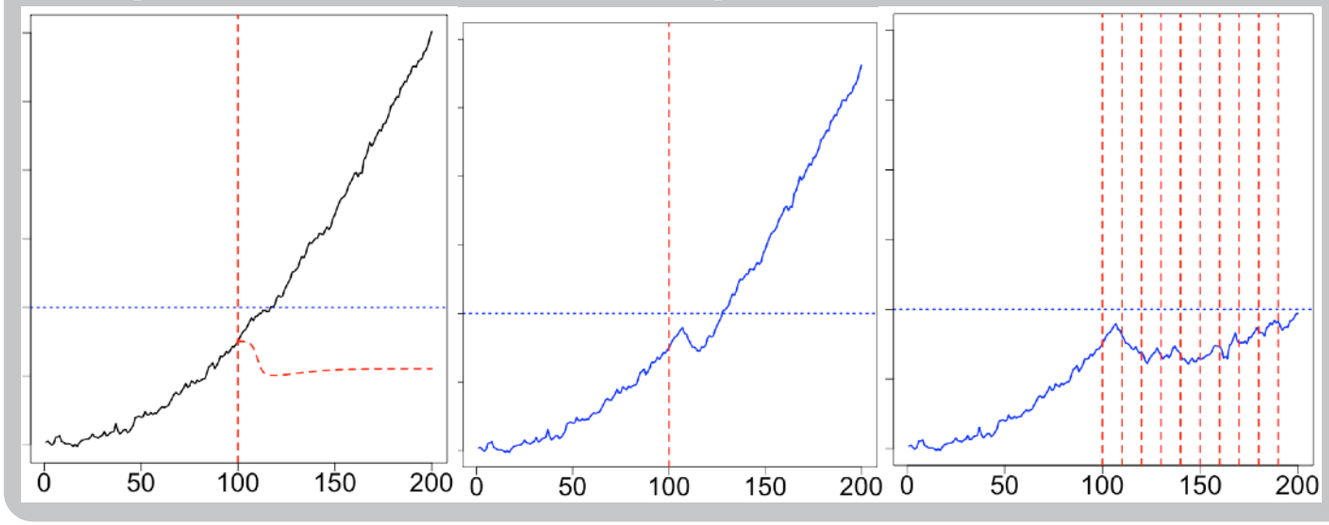
Longitudinal normal dynamics



Treatment-response curves



Response curves from multiple treatments are added

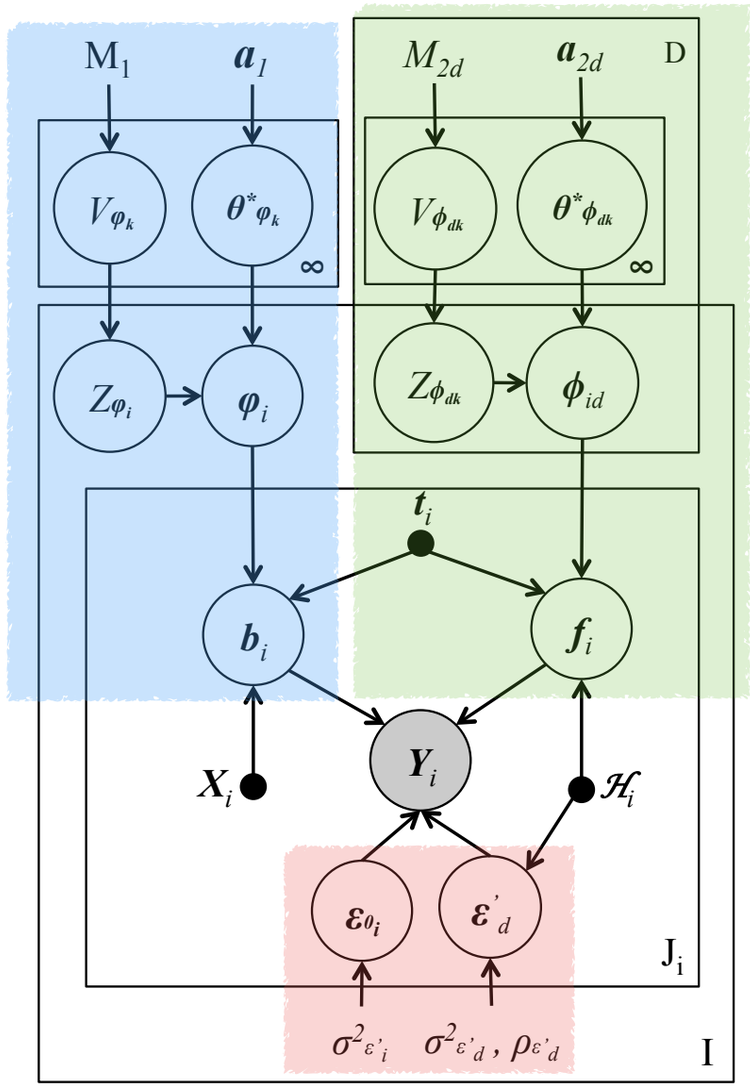


Approach: Individualized treatment response (ITR) model

$$Y_{ij}|\boldsymbol{X}_i, \mathcal{H}_{ij} = \underbrace{b(\boldsymbol{X}_i) + \boldsymbol{u}_i(t_{ij})}_{\text{baseline progression}} + \underbrace{\boldsymbol{f}_i(t_{ij}; \mathcal{H}_{ij})}_{\text{treatment response}} + \underbrace{\boldsymbol{\epsilon}_i(t_{ij}; \mathcal{H}_{ij})}_{\text{noise}}, \ j = 1, ..., J_i.$$

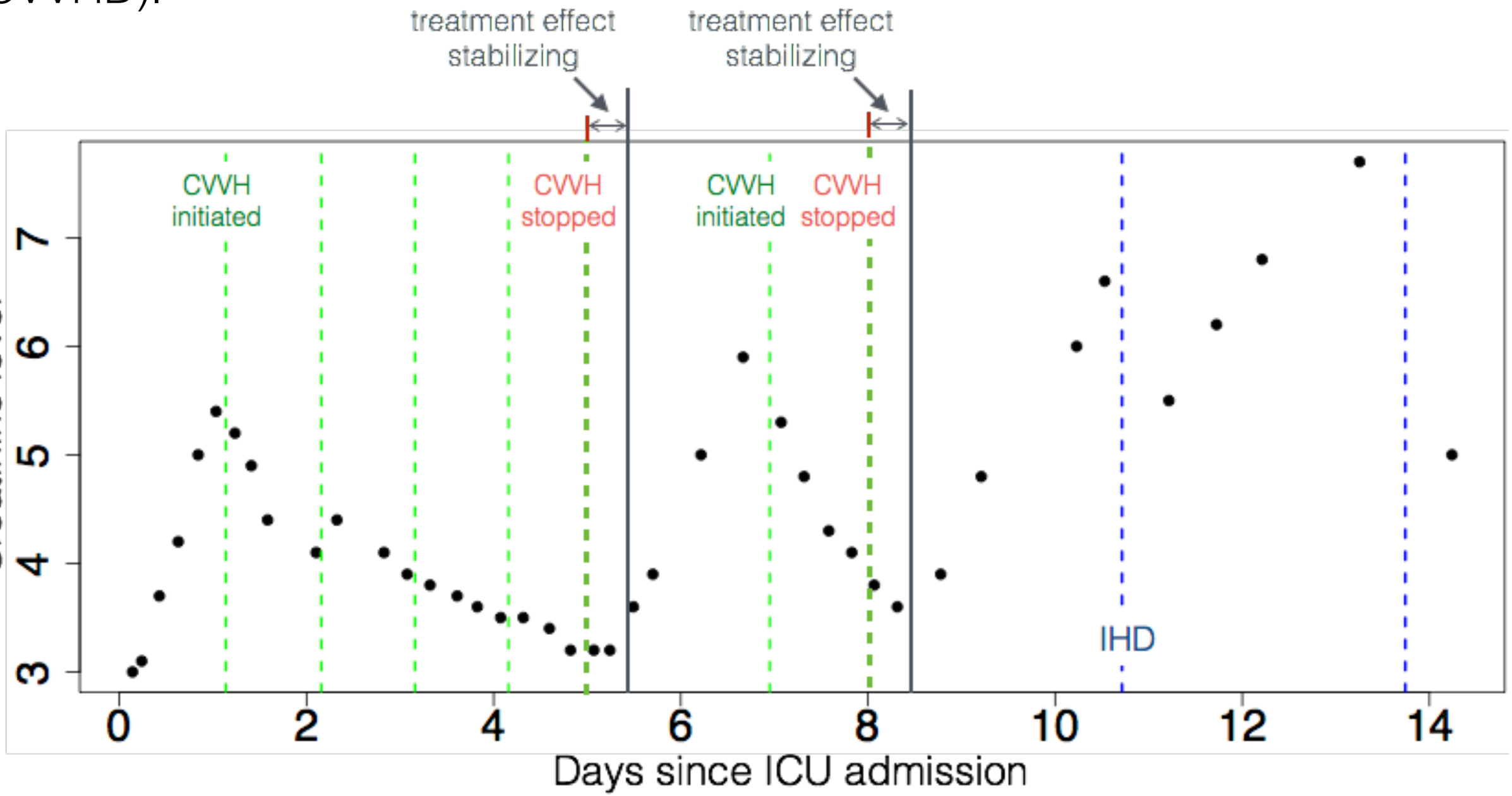
Rcpp implementation: https://github.com/YanxunXu/BayesianITR

To **Cluster** model parameters such that **individuals** with similar responses can **share statistical strength**, we generalize **Dirichlet Process Mixture** to the two components..



Numerical analysis

Goal: estimate heterogeneous response curves to renal replacement therapy.
Marker: creatinine, a measure for kidney function.
Treatments: renal replacement therapy (RRT): intermittent hemodialysis (IHD), continuous Veno-Venous Hemofiltration (CVVH), and CVV Hemodialysis (CVVHD).



Numerical analysis

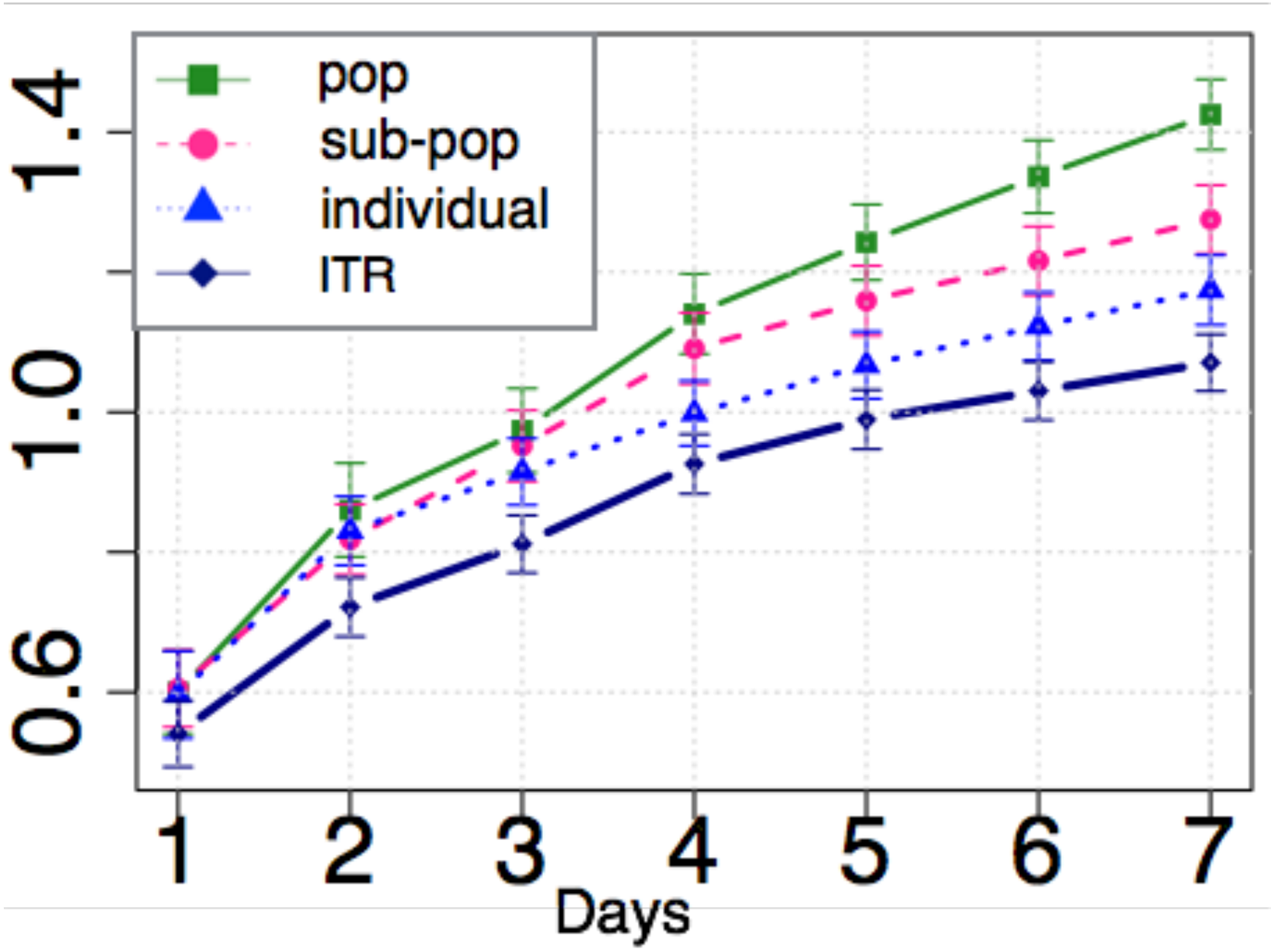
Data: publicly available in the MIMIC-II Clinical Database (Saeed et al., 2002).

- We have 428 trajectories with 16,593 creatinine observations.
- 525 instances of IHD, 186 of CVVH, and 981 of CVVHD.

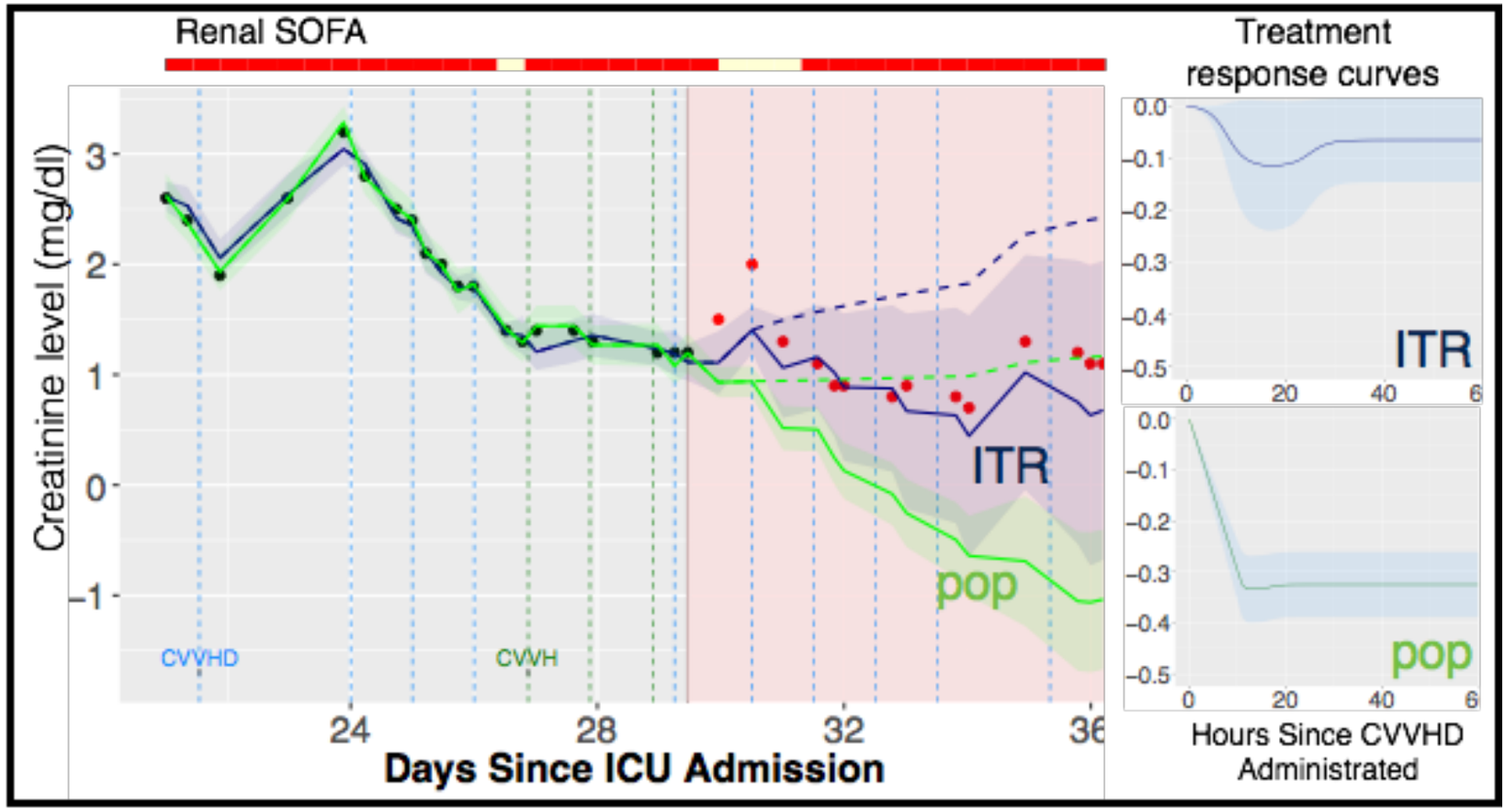
Numerical analysis

Comparison:

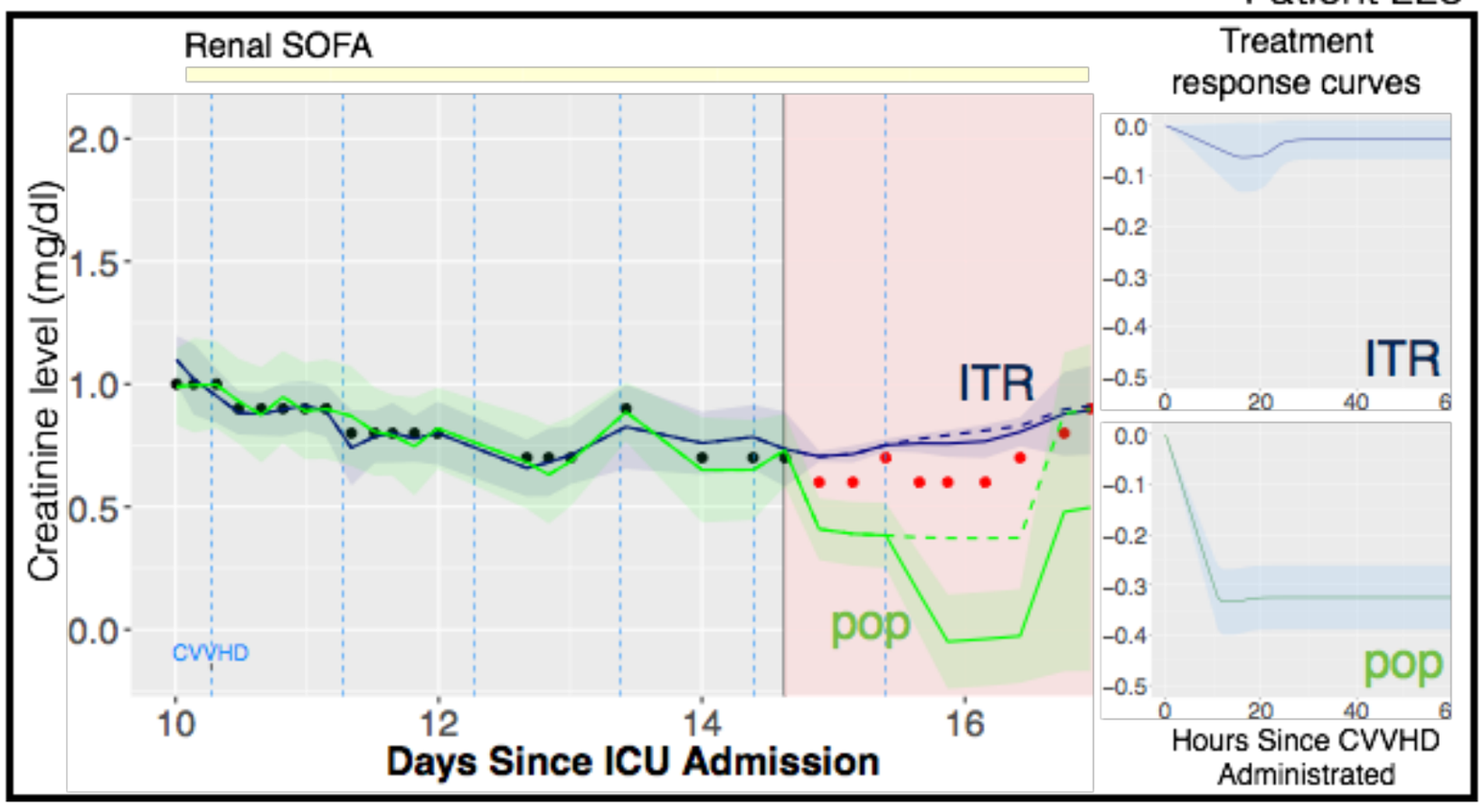
- **pop model:** estimate treatment effect at the population level.
- **individual model:** estimate treatment effect at the individual level.
- **sub-pop model:** treatment effect vary by subgroups.



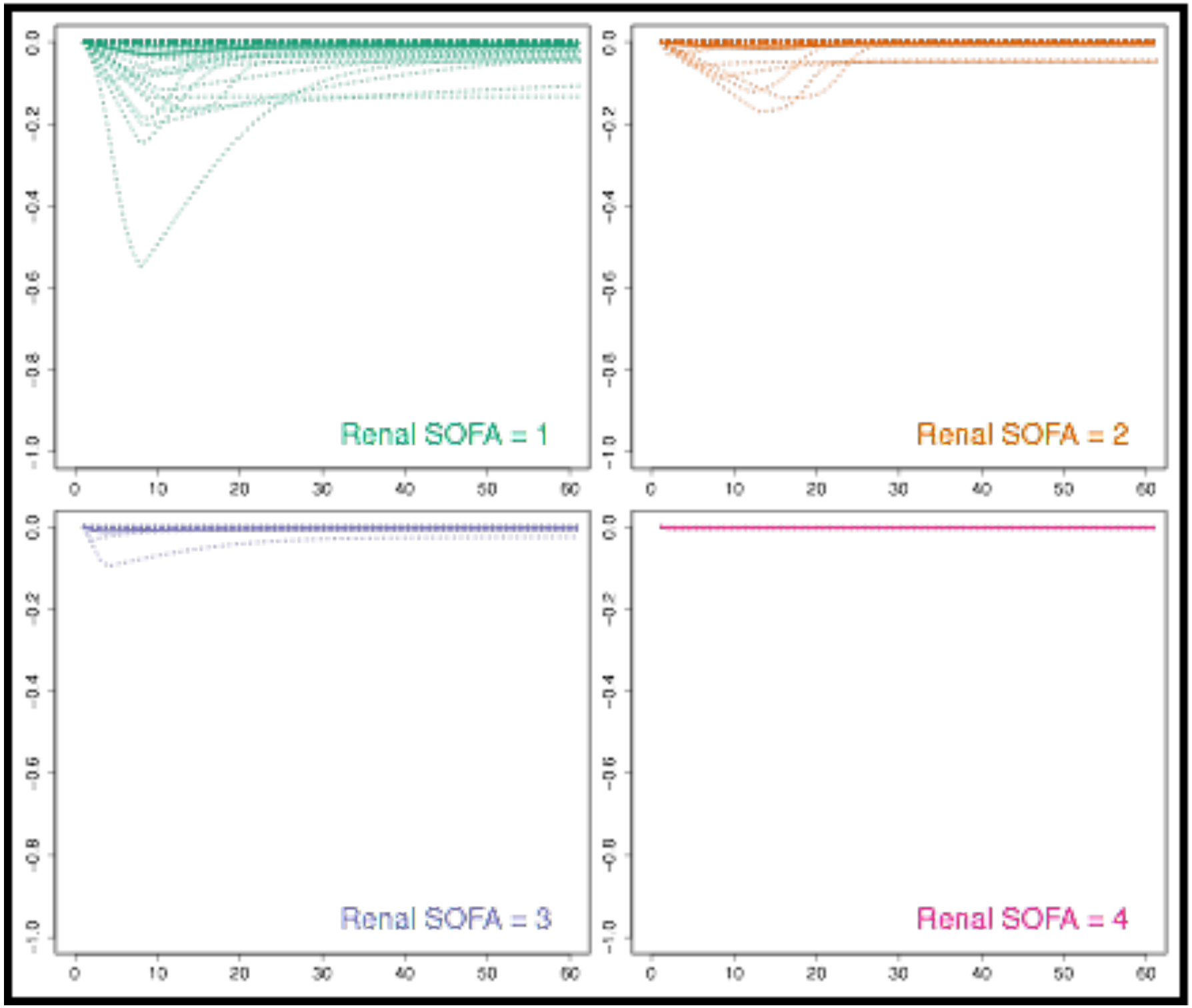
Patient 44



Patient 228

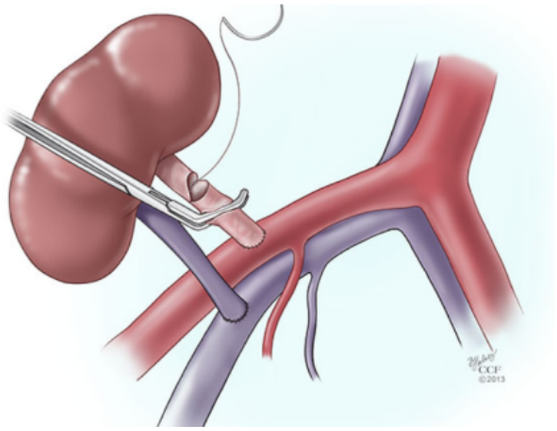


(a) Patients with 4-level kidney failures



Motivation: Kidney Transplant

Patient undergoes kidney transplantation surgery



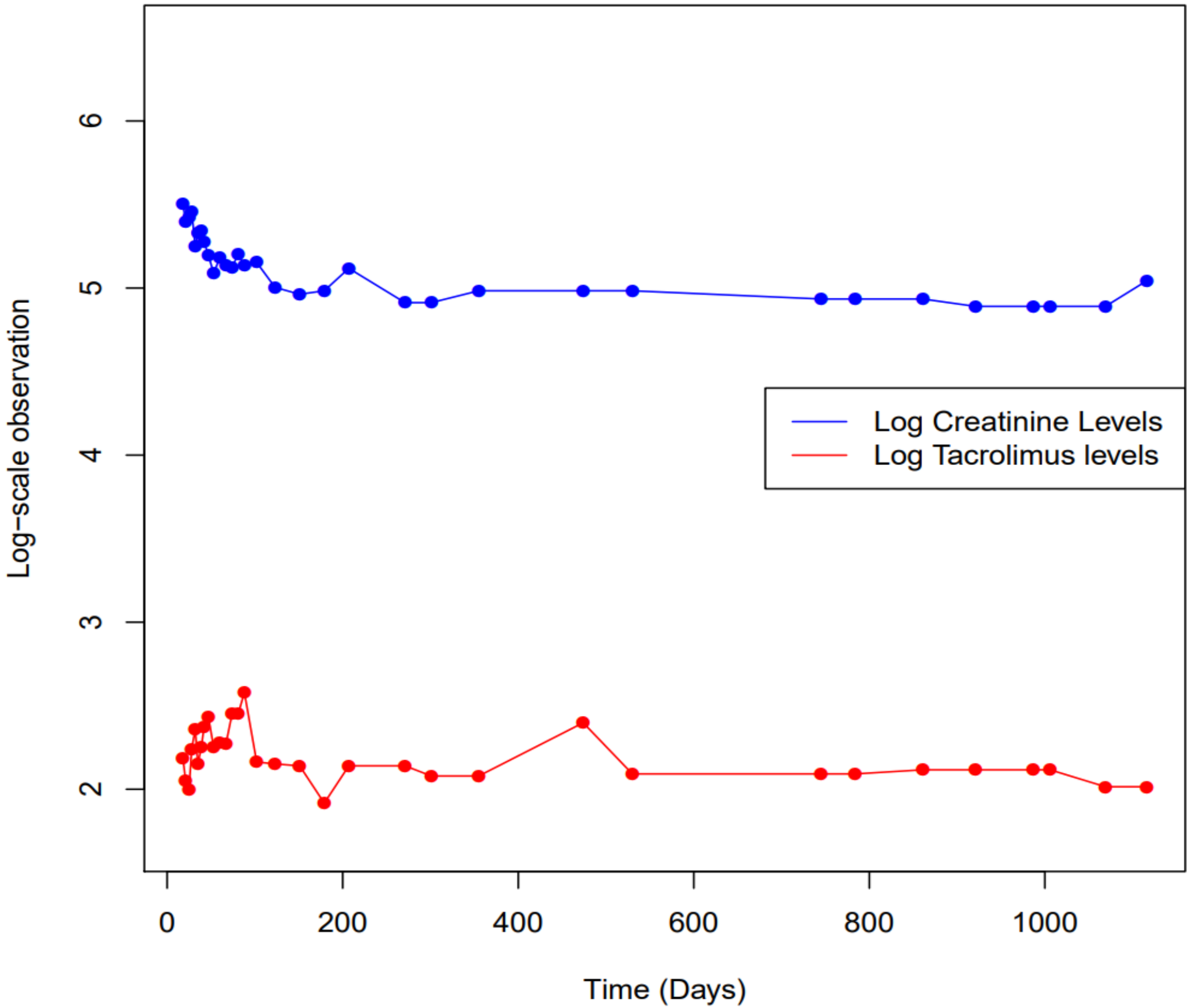
Tacrolimus

Patient follows up with doctor and records longitudinal observations

Doctor gives patient new dosage levels and determines next follow up visitation time

At some time, either the patient's death is observed or the survival time is censored.

Example Observed Data



Goals

- Understand how creatinine changes along the time

Longitudinal modeling

- Study how creatinine affects survival

Joint modeling of longitudinal data and survival

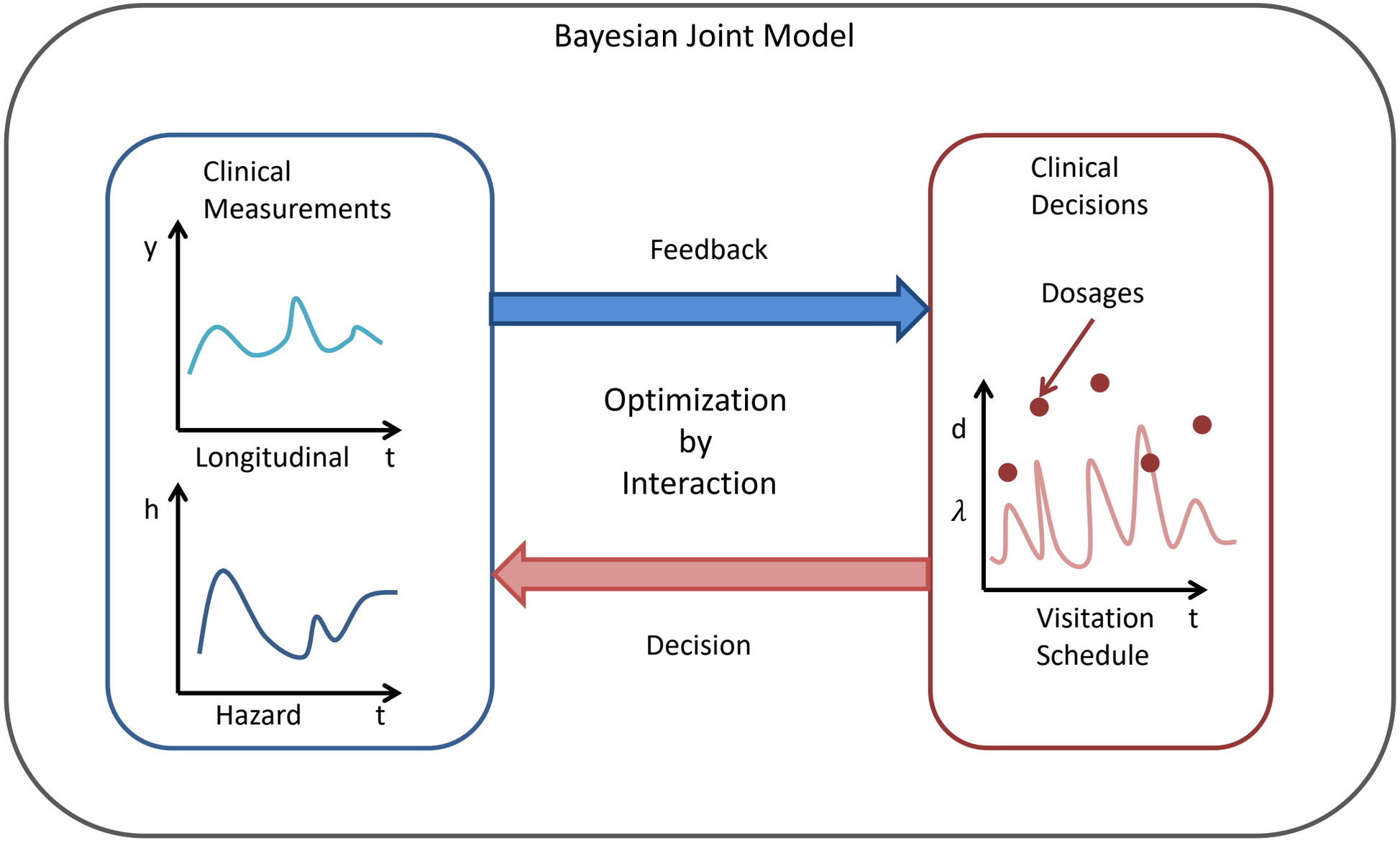
- Learn how doctors treat patients

Visitation schedule and dosage

- Find an optimal visitation and dosing strategy to maximize survival outcomes.

Optimization

Approach Overview



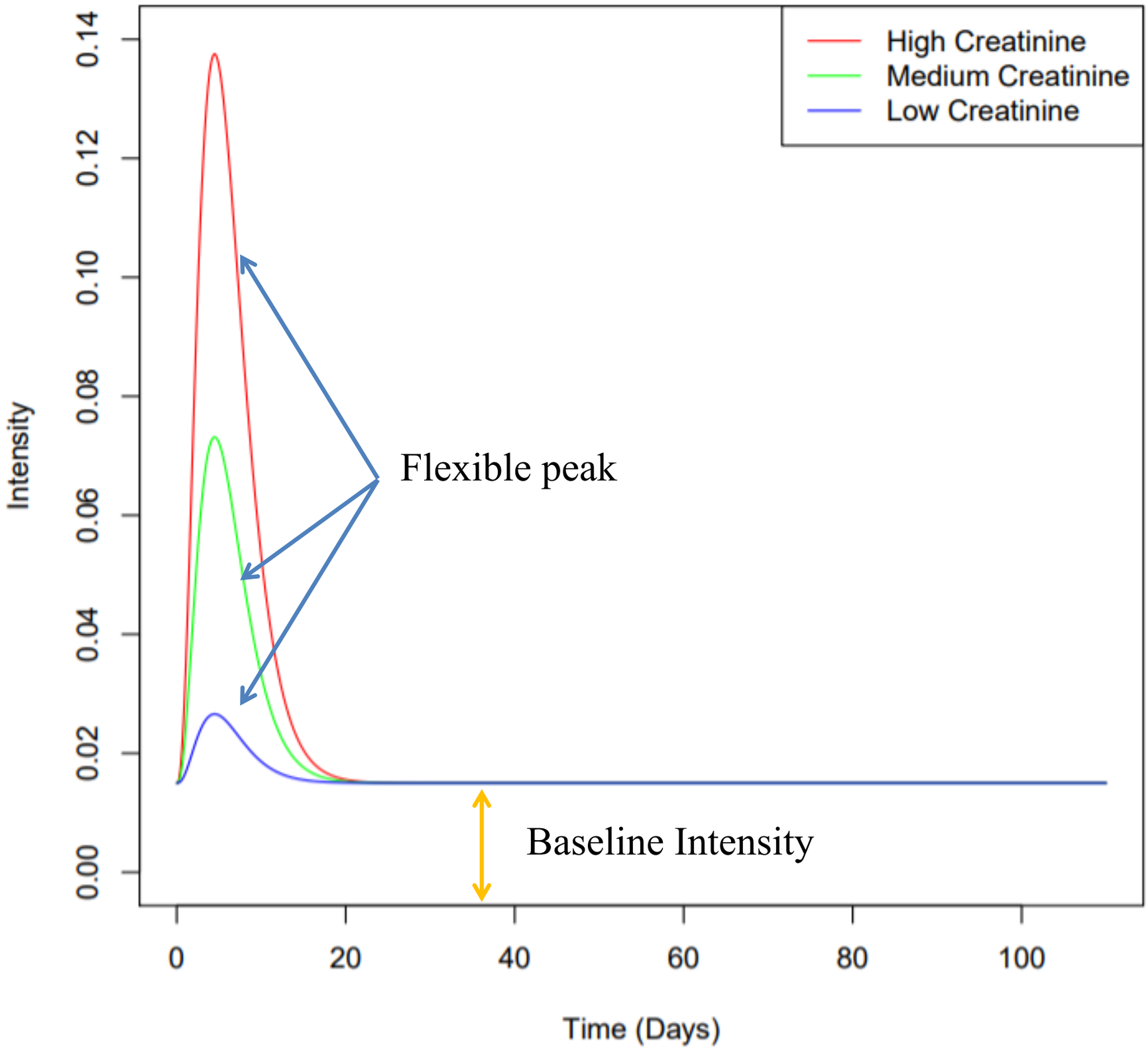
p(Longitudinal, Survival, Visitation, Dosage)

argmax **Reward (Visitation, Dosage)**

Visitation, Dosage

Hua et al., 2021

Proposed Visitation Modeling



Marked Temporal Point Process

$$\mathcal{A}_{i,T} = \{(t_{i,1}, D_{i,1}), \ldots, (t_{i,n_i}, D_{i,n_i})\}$$

$$p(\mathcal{A}_{i,T}) = \underbrace{\exp\Big(-\int_0^{t_{i,n_i}} \lambda_i(x)dx\Big)}_{\text{Prob. of no visits at } t\in[0,T]\setminus\{t_{i,j}\}_{j=1}^{n_i}} \prod_{j=1}^{n_i} \Big(\overbrace{\lambda_i(t_{i,j})}^{(2.1)} \underbrace{p(D_{i,j} \mid A_{i,j}, \boldsymbol{\beta}_d, \sigma_d^2)}_{\text{Prob. of dosage } D_{i,j}} \Big)$$

Prob. of an action at $t_{i,j}$

Bayesian Joint Modeling

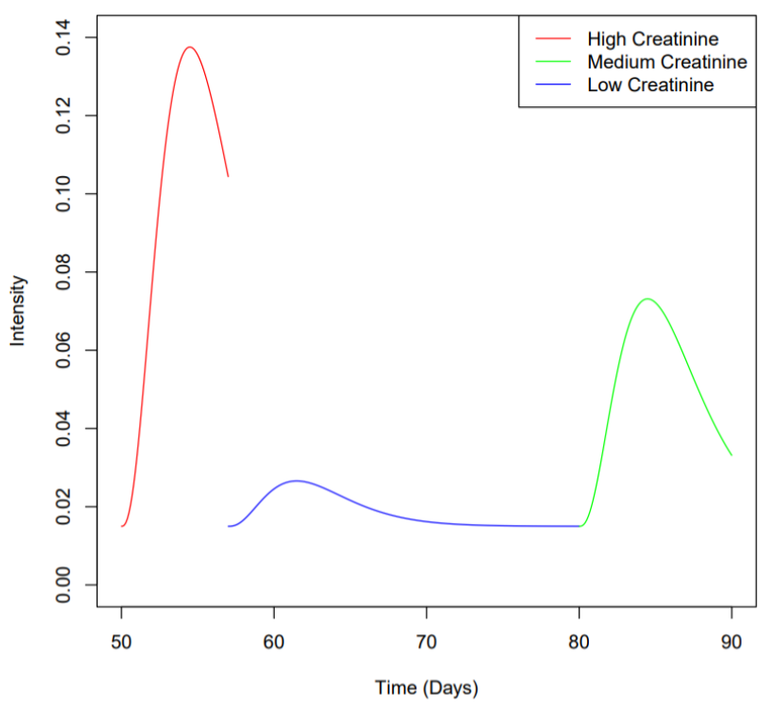
$$\prod_{i=1}^{N} p(\boldsymbol{Y_i}, \mathcal{A}_{i,T_i}, T_i, \delta_i \mid \boldsymbol{X}_i, \boldsymbol{\theta}_\alpha, \boldsymbol{\beta}_l, \boldsymbol{\beta}_d, \boldsymbol{\theta}_v, \boldsymbol{\theta}_s, \boldsymbol{b}_i, \sigma_l^2, \sigma_d^2)$$

$$\propto \prod_{i=1}^{N} \Big(p(\mathcal{A}_{i,T_i} \mid \boldsymbol{Y_i}, \boldsymbol{X}_i, \boldsymbol{\theta}_v, \boldsymbol{\theta}_\alpha, \boldsymbol{\beta}_d, \sigma_d^2) p(\boldsymbol{Y_i} \mid \mathcal{A}_{i,T_i}, \boldsymbol{X}_i, \boldsymbol{\beta}_l, \sigma_l^2, \boldsymbol{b}_i) p(T_i, \delta_i \mid \boldsymbol{X}_i, \boldsymbol{Y_i}, \mathcal{A}_{i,T_i}, \boldsymbol{\beta}_l, \boldsymbol{\theta}_s, \boldsymbol{\theta}_\alpha) \Big)$$

Marked temporal point process for visitations and dosages

Longitudinal creatinine

Survival



Optimal Treatment

$$\prod_{i=1}^{N} p(\boldsymbol{Y_i}, \mathcal{A}_{i,T_i}, T_i, \delta_i \mid \boldsymbol{X}_i, \boldsymbol{\theta}_\alpha, \boldsymbol{\beta}_l, \boldsymbol{\beta}_d, \boldsymbol{\theta}_v, \boldsymbol{\theta}_s, \boldsymbol{b}_i, \sigma_l^2, \sigma_d^2)$$

Action parameters: $\Theta = (\beta_{\nu 1}, \beta_{\nu 2}, \mu, \beta_d)$

Other parameters: ϕ

Goal:

Posterior distribution of ϕ

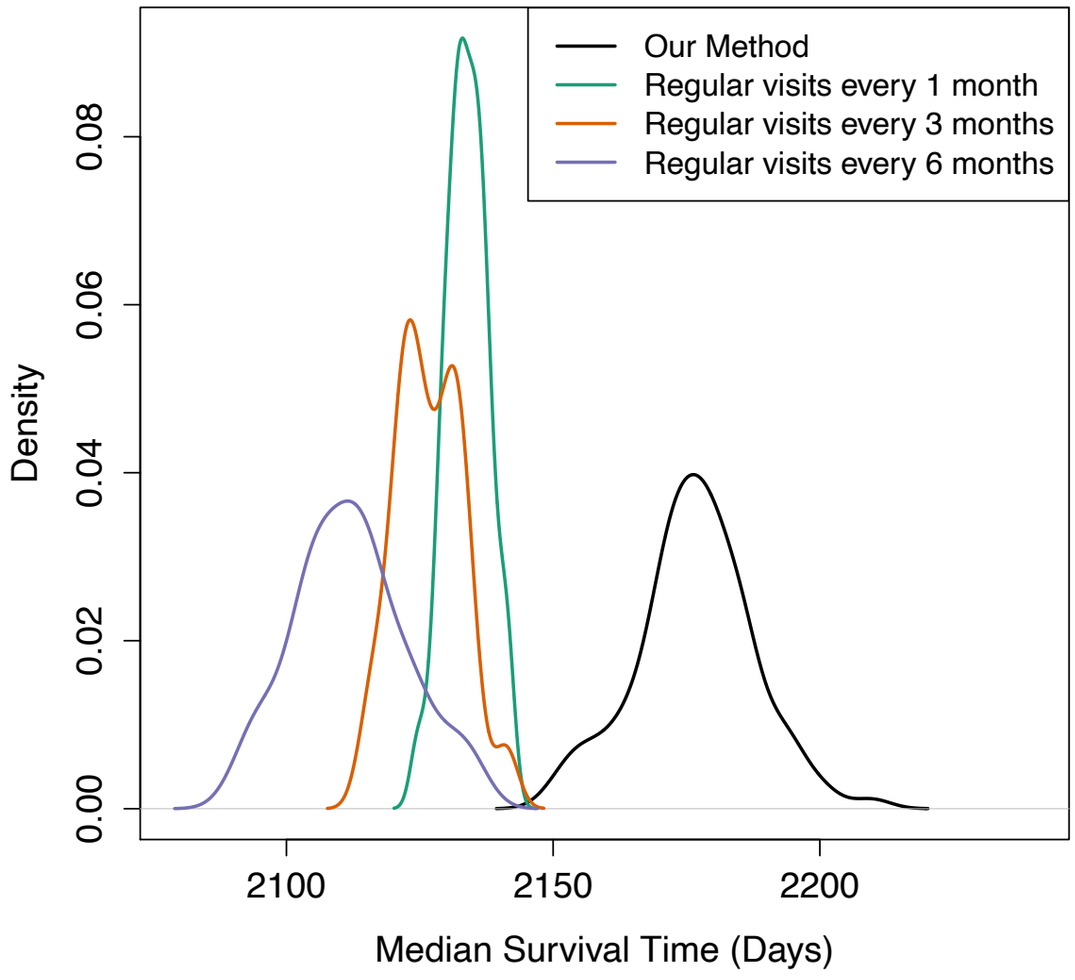
$$\text{maximize}_\Theta E_{\Theta,\phi}[R_i(T)]$$

$$\text{maximize}_\Theta \int E_{\Theta,\phi}[R_i(T)] p(\phi \mid \mathcal{D}) d\phi$$

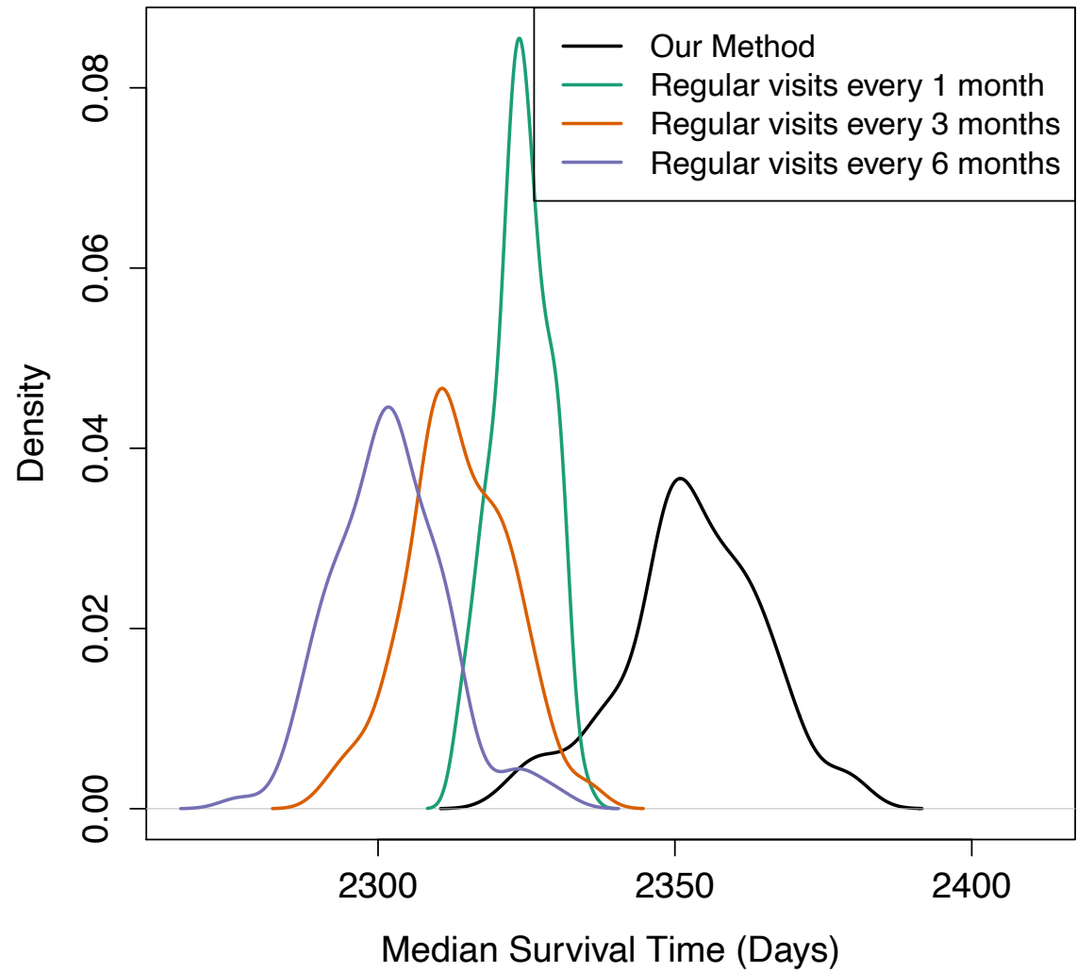
Median survival time

Reinforcement learning: policy gradient

Optimal Treatment



(c) Patient S1



(d) Patient S2

Connection to Offline Reinforcement Learning

Offline reinforcement learning: contextual bandits (Dudik et al., 2011), sequential decision-making problems (Jiang and Li, 2015).

Markov Decision Process

A Markov Decision Process is a tuple $M = (\mathcal{S}, \mathcal{A}, T, r, \mu_0, \gamma)$

- $\mathcal{S}$ is the state space
- $\mathcal{A}$ is the action space
- $T(s' \mid s, a)$ is the transition dynamics
- $r(s, a)$ is the reward function
- μ_0: the initial state distribution
- $\gamma \in (0,1)$: discount factor

Denote $\pi(a \mid s)$ the policy function,

$$\pi^* = \text{argmax}\mathbb{E}_{\pi,T,\mu_0}[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t)]$$

MOPO: Model-based Offline Policy Optimization

Uncertainty-penalized MDP:

$$\tilde{M} = (\mathcal{S}, \mathcal{A}, \hat{T}, \tilde{r}, \mu_0, \gamma)$$

Estimated dynamic model

$$\tilde{r}(s,a) = r(s,a) - \lambda u(s,a).$$

Yu et al., 2020

Take-home Messages

What **BNP** brings to treatment estimation from observational data:

- Easy and flexible modeling for individual treatment effects
- Uncertainty quantification in decision process
- Complex settings