JOHNS HOPKINS UNIVERSITY

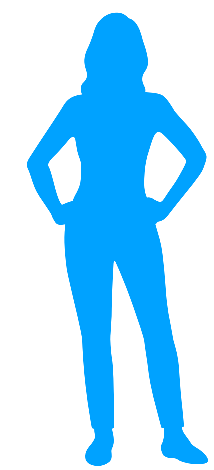
Reliable Decision Support using Counterfactual Models

Suchi Saria
Assistant Professor
Computer Science, Applied Math & Stats
and Health Policy
Institute for Computational Medicine

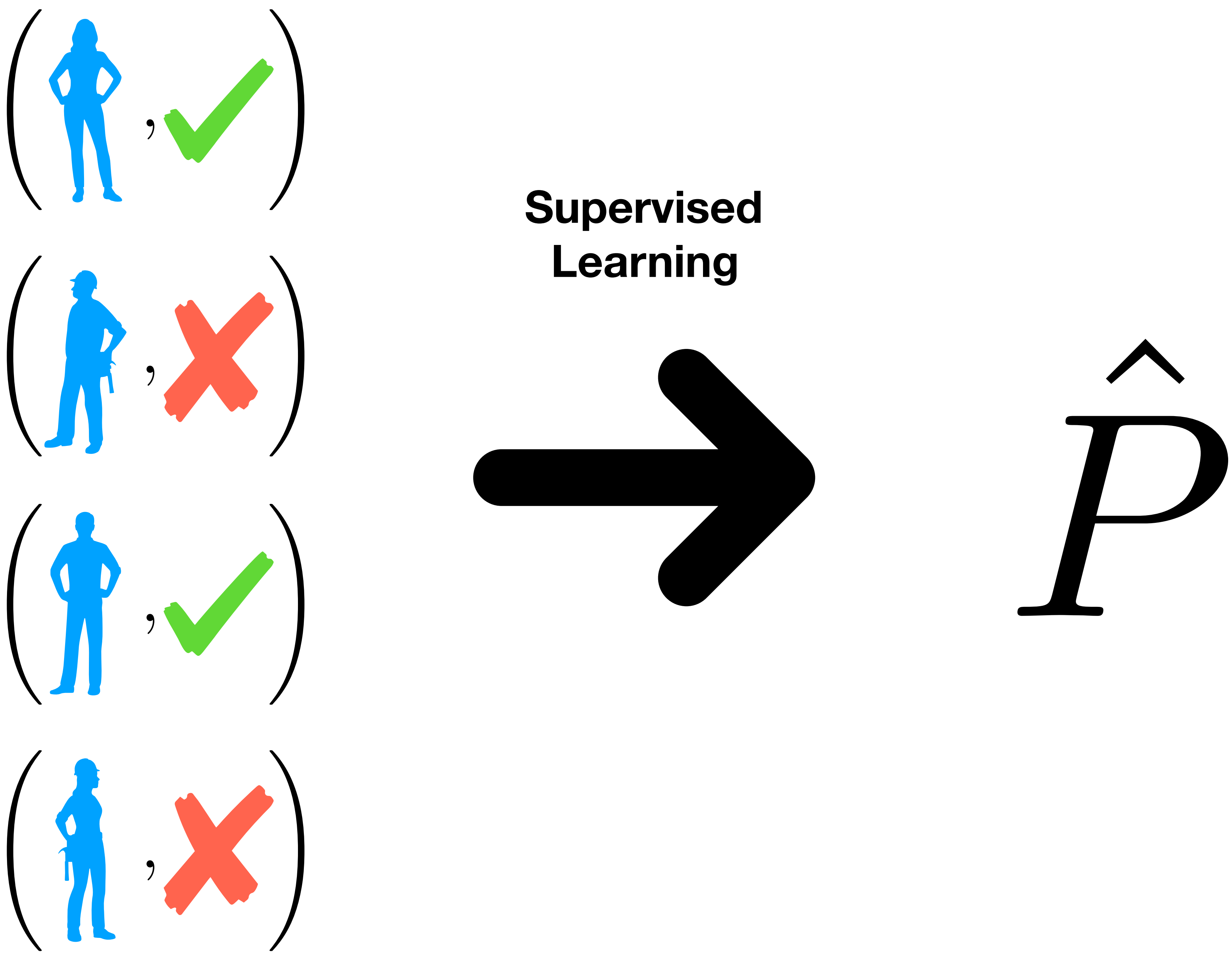
w/ Peter Schulam, PhD candidate

NSF
GORDON AND BETTY MOORE FOUNDATION
Google
THE MICHAEL J. FOX FOUNDATION FOR PARKINSON'S RESEARCH
DARPA

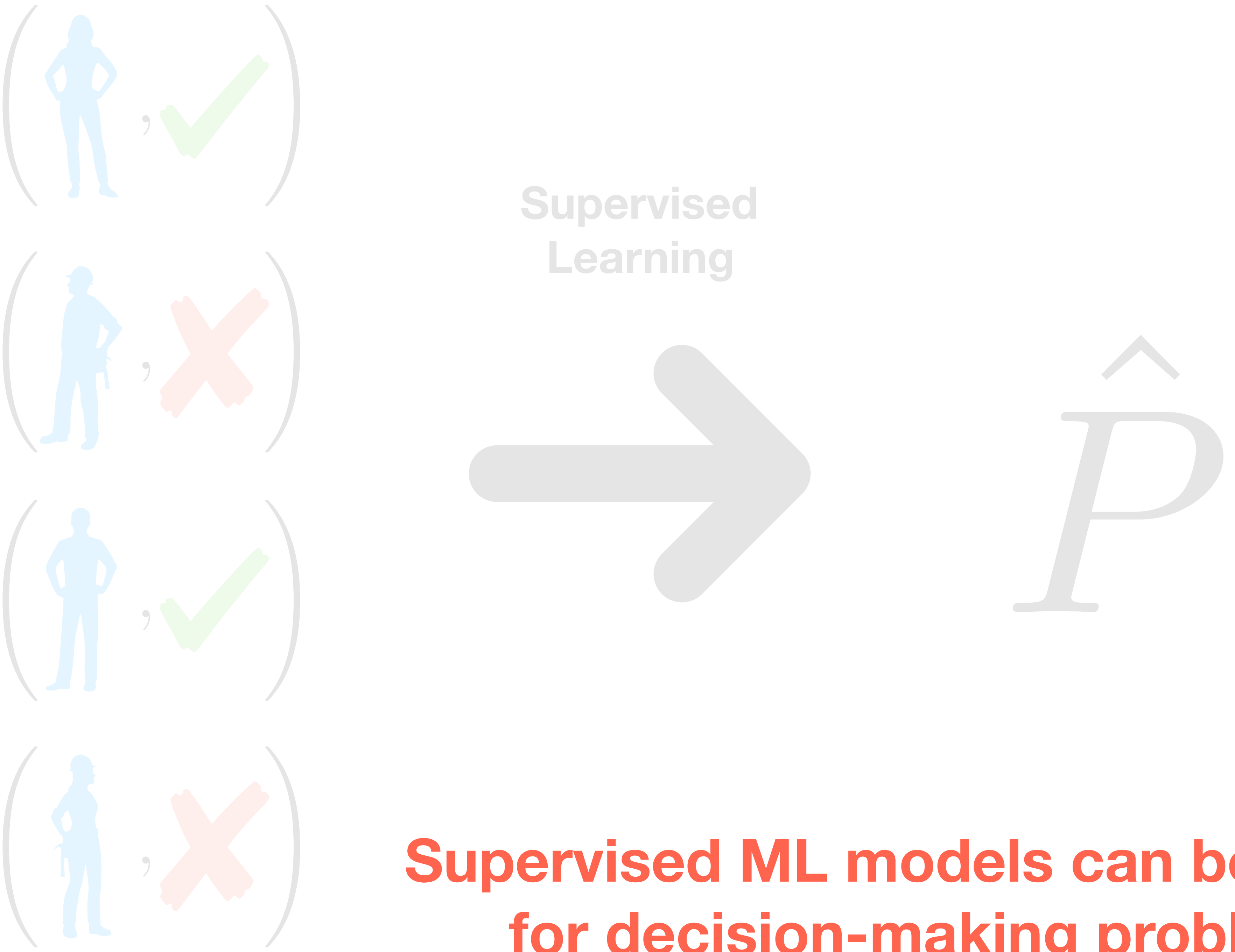
Example: Customer Churn

$P\Big(\text{Cancels Account} \mid$  $\Big)$

Example: Customer Churn

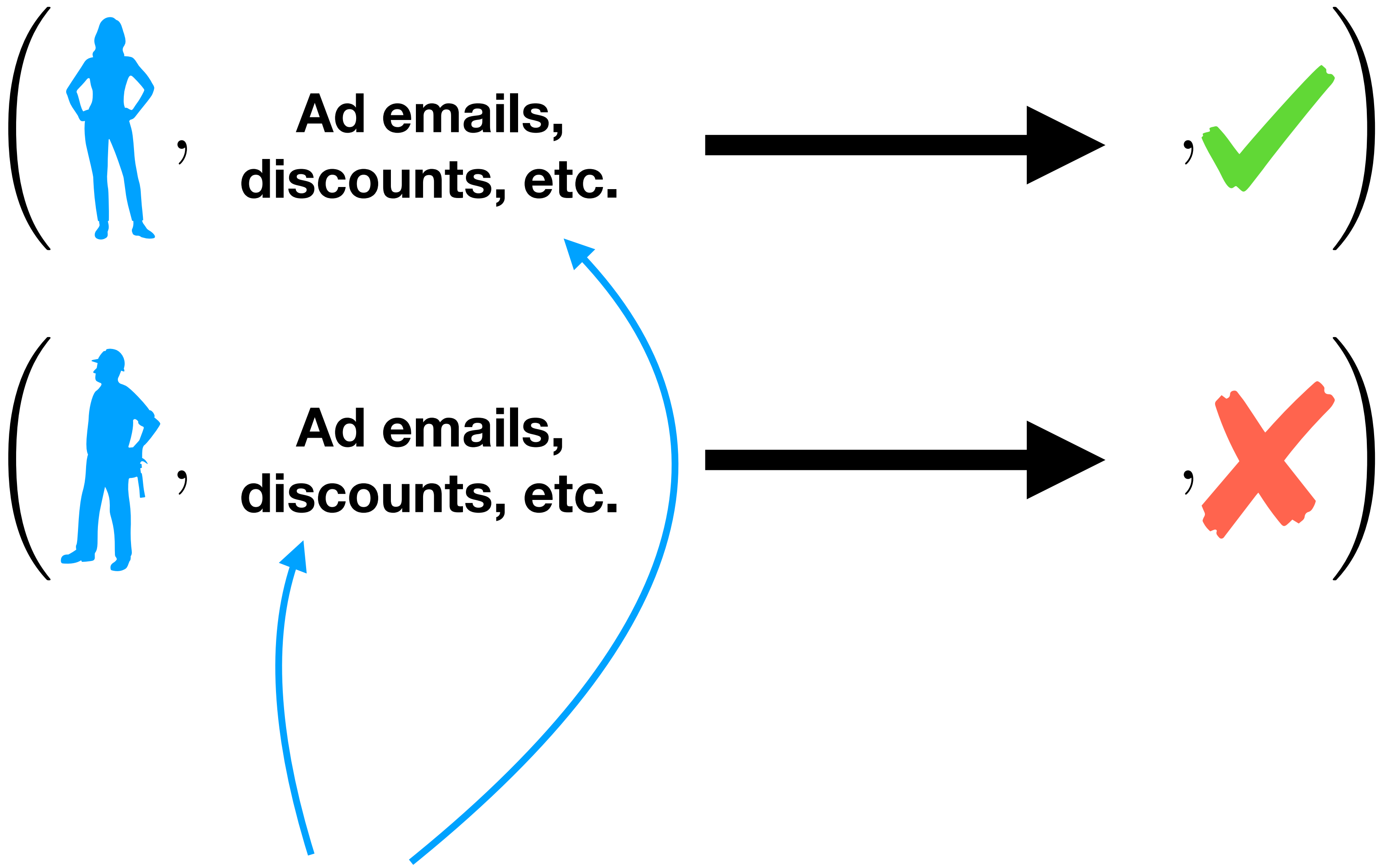


Example: Customer Churn



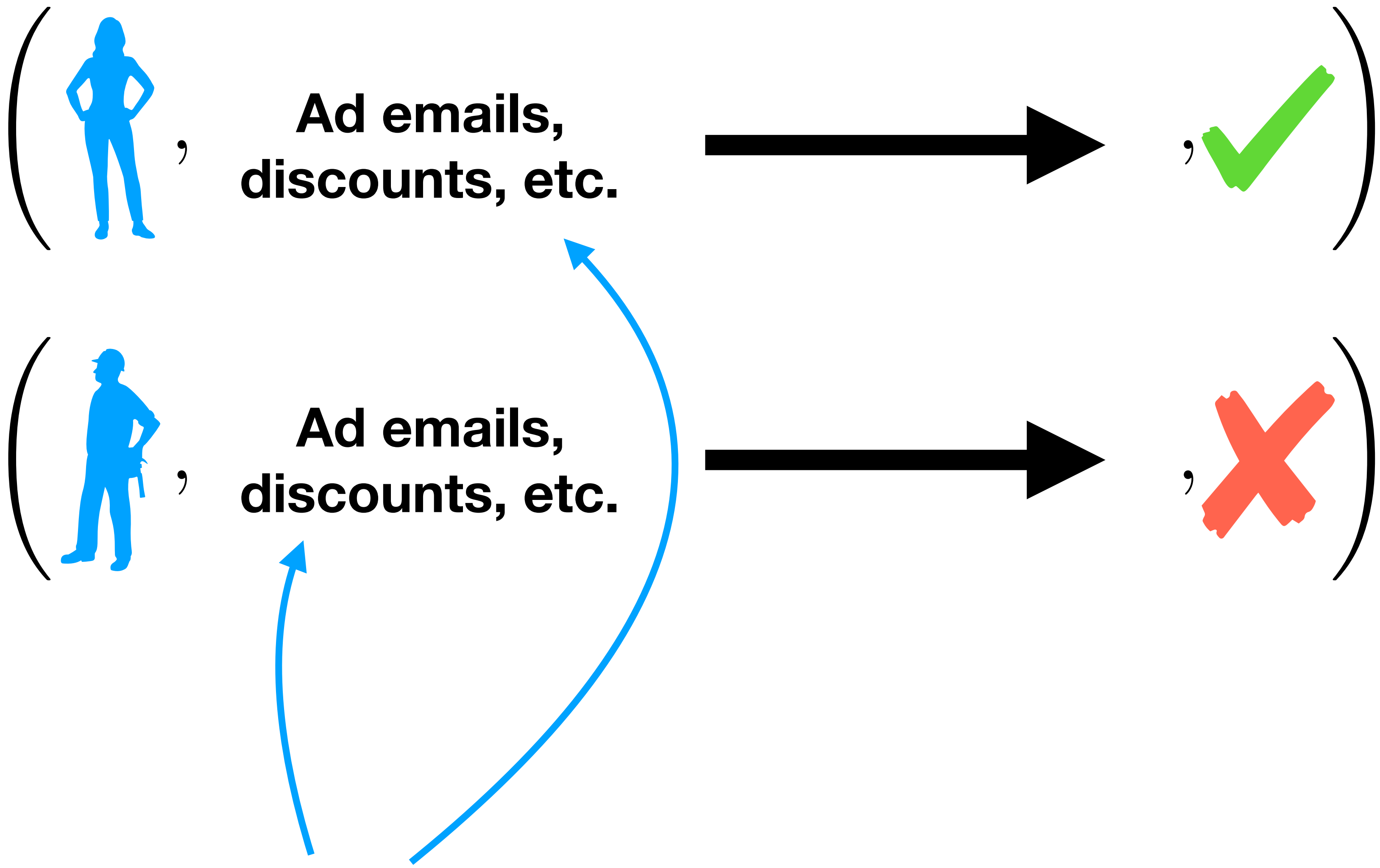
Supervised ML models can be biased for decision-making problems!

Why?



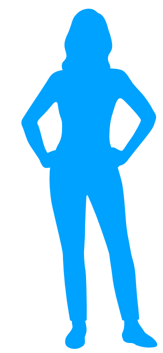
Past actions determined by some policy.

Why?

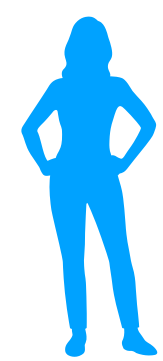


Actions determined by a policy based on your learned model $\hat{P}$

Why?

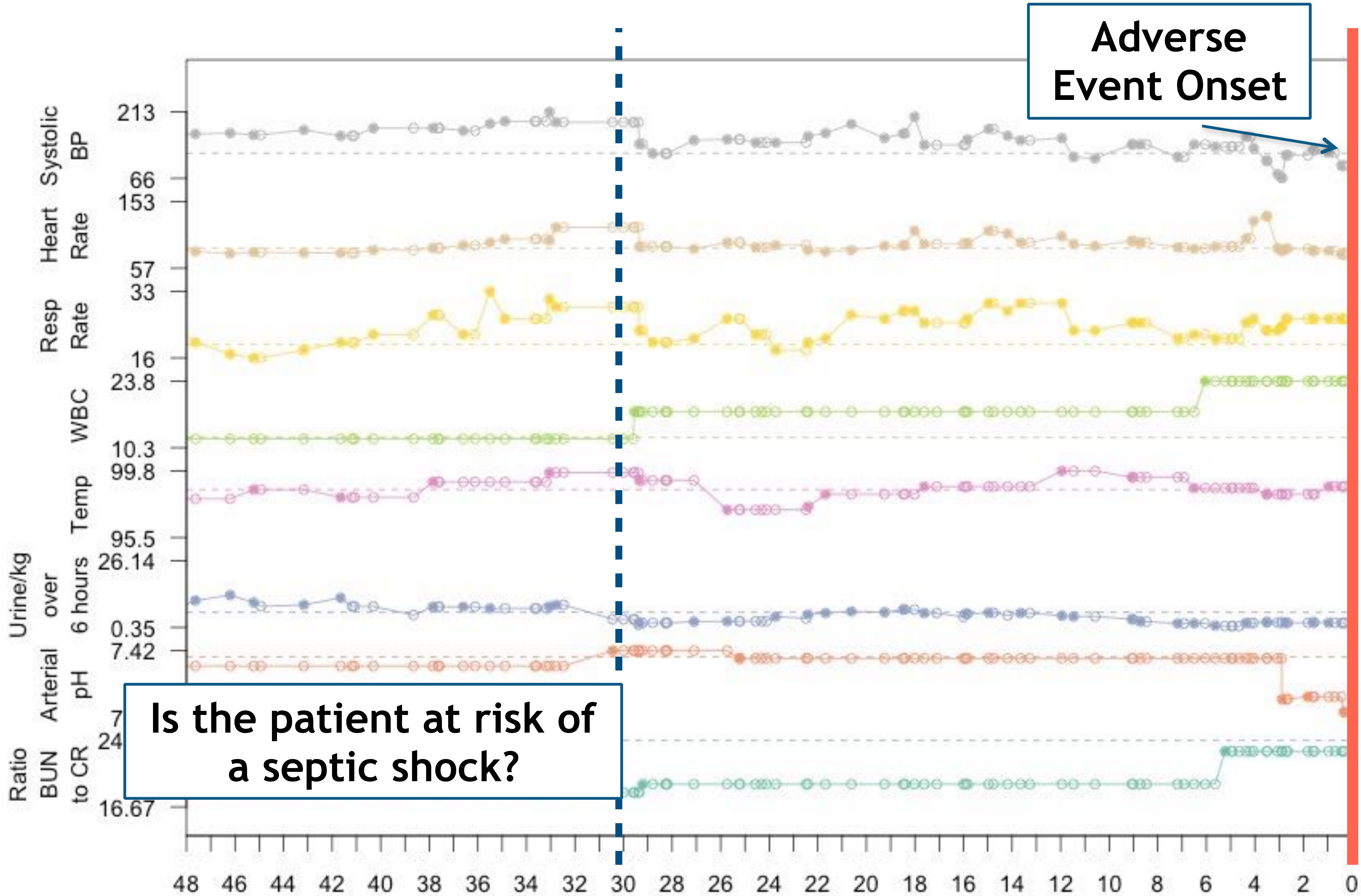
$P\Big($ Cancels Account $\mid$  $, \pi_{\text{train}}\Big)$

$$\neq$$

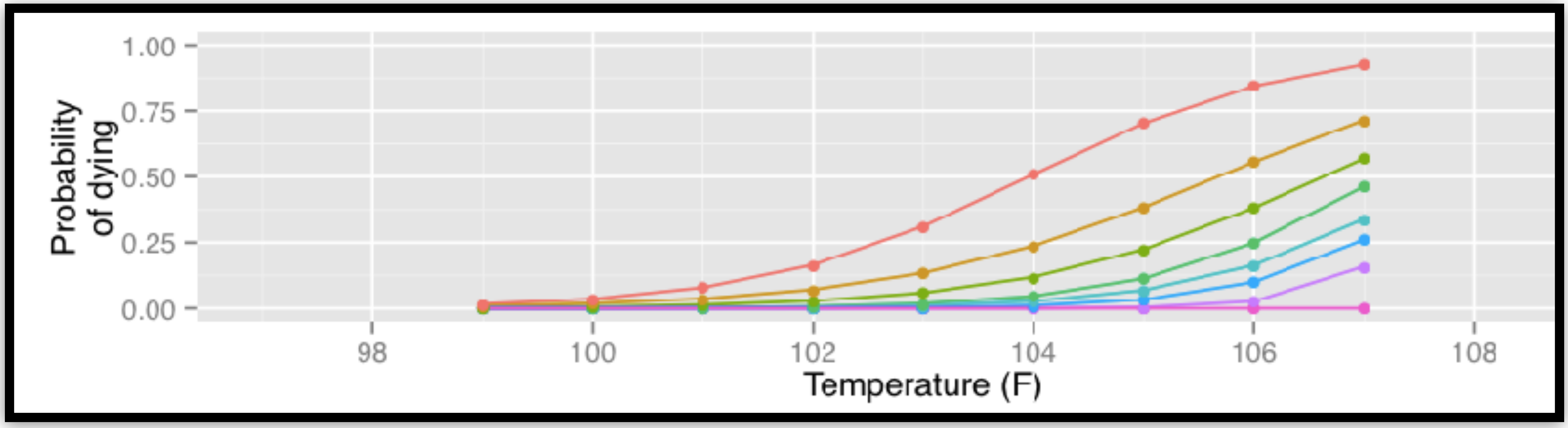
$P\Big($ Cancels Account $\mid$  $, \pi_{\text{test}}(\hat{P})\Big)$

Supervised ML leads to models that are **unstable** to shifts in the policy between the train and test

Example: Risk Monitoring



- Rise in Temperature and Rise in WBC are indicators of sepsis and death
- But, doctors in H1 aggressively treat patients with high temperature
- As doctors treat treat more aggressively, supervised learning model learns **high temperature is associated with low risk**.



Dyagilev and Saria, Machine Learning 2015

Treat based on temp

Treat based on WBC

Scenario	ρ_T^{train}	ρ_{WBC}^{train}	ρ_T^{test}	ρ_{WBC}^{test}	Logistic Regression
#1	0	0	0	0	0.974
#2	0.1	0	0.1	0	0.978
#3	0.1	0	0	0	0.963
#4	0.3	0	0	0	0.769
#5	0.3	0	0	0.3	0.510

Increasing **discrepancy** in physician prescription behavior in train vs. test **environment**

Predictive model trained using classical supervised ML creates unsafe scenarios where sick patients are overlooked.

Dyagilev and Saria, Machine Learning 2015

Run an experiment:
observe outcome under diff scenarios

- Clone the customer; give a 10% and 20% discount code to each clone

- Choose the outcome that has the better outcome

$$\left\{ \boxed{Y(d_{10})}, \quad Y(d_{20}) \right\}$$

Outcome under 10% discount.

Run an experiment: observe outcome under diff scenarios

- Clone the customer; give a 10% and 20% discount code to each clone
- Choose the outcome that has the better outcome

$$\left\{ \; Y(d_{10}) \;\; , \;\; \boxed{Y(d_{20})} \; \right\}$$

Outcome under 20% discount.

Can we learn models of these outcomes from observational data?

- Factual: outcome observed in the data

  vs.

- Counterfactual: outcome is unobserved

$$\left\{ \boxed{Y(d_{10})} , \boxed{Y(d_{20})} \right\}$$

Potential Outcomes

Random variable

Set of actions

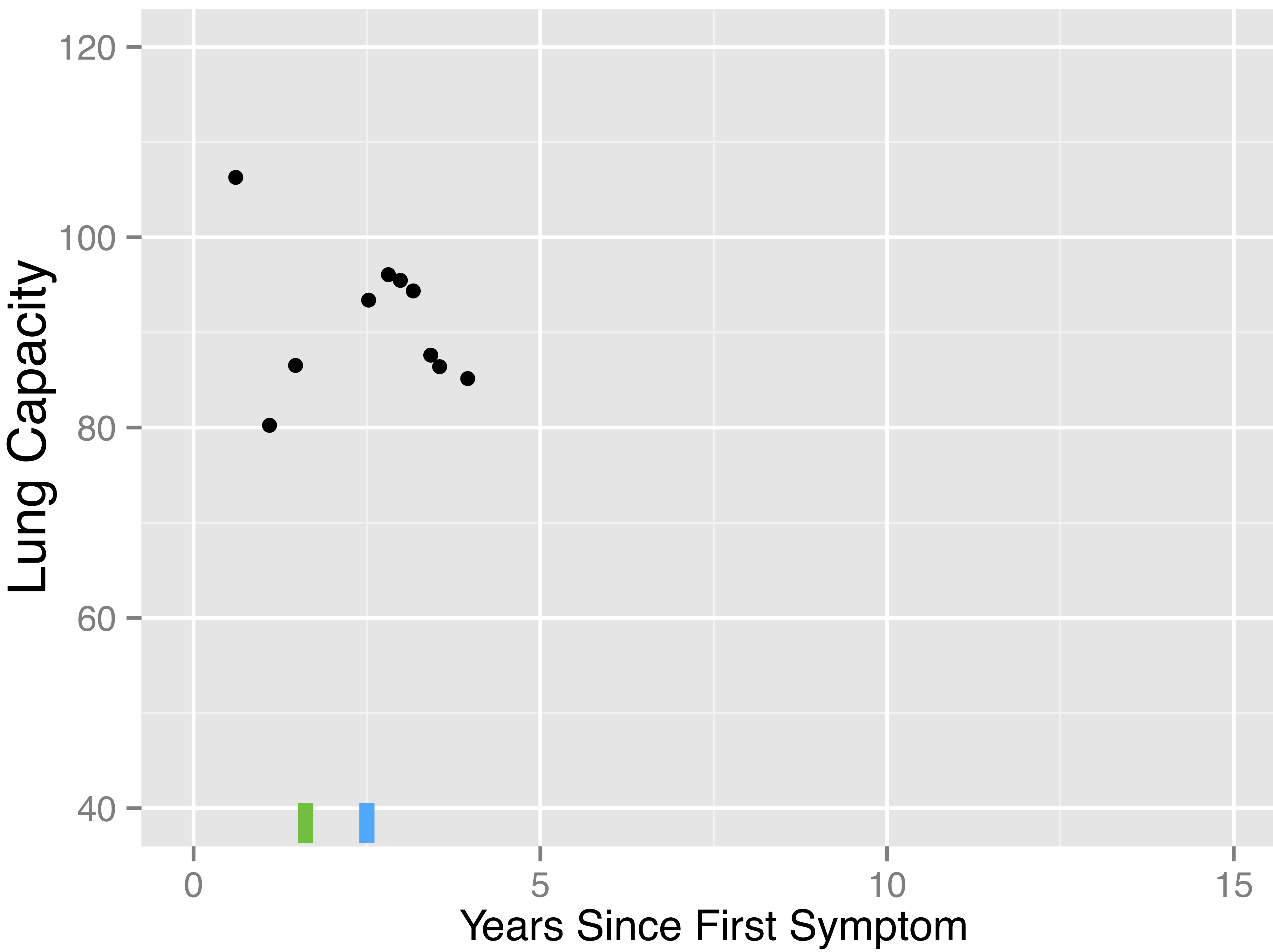
$$\{Y(a) : a \in \mathcal{A}\}$$

Action

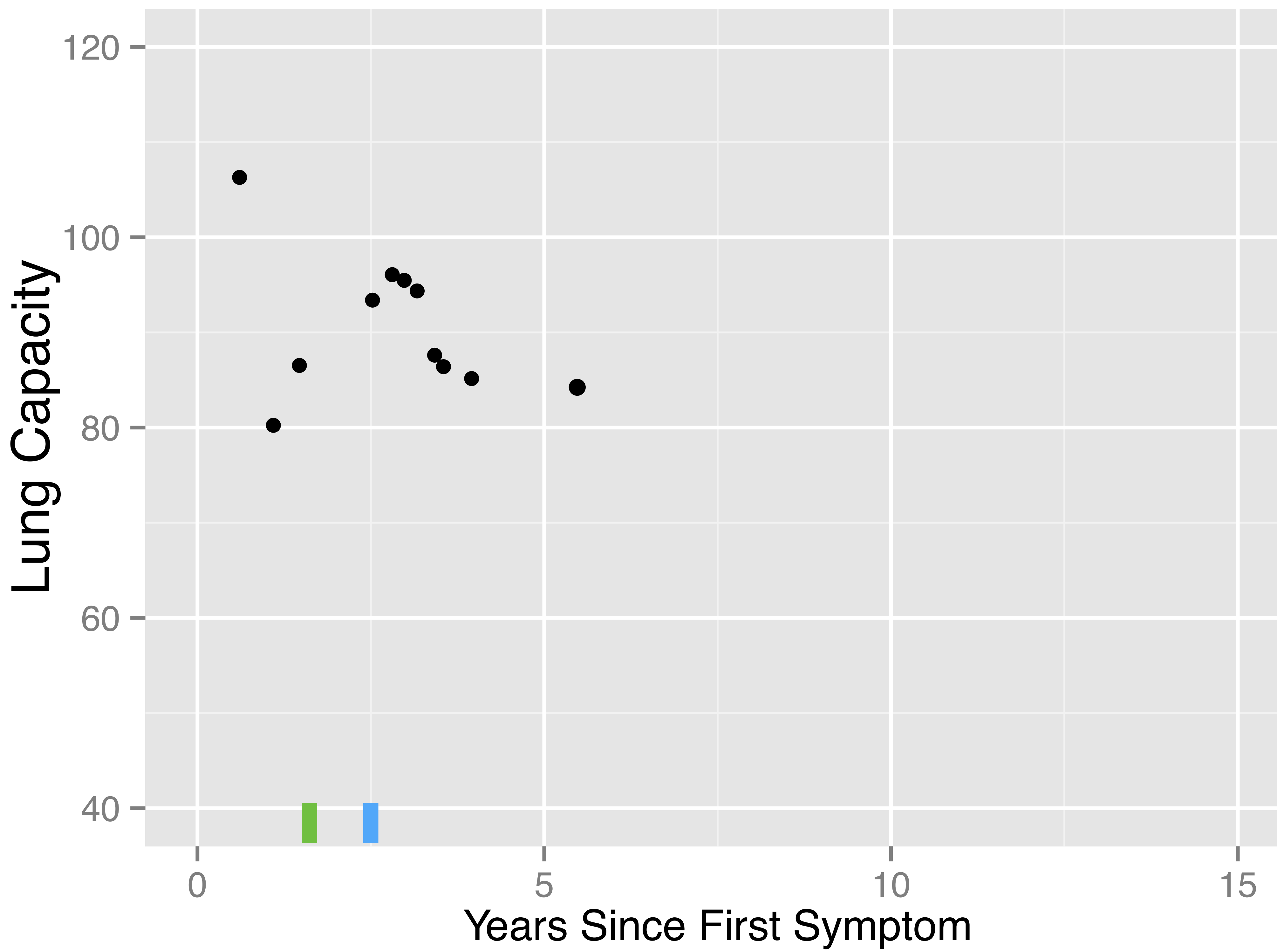
Potential outcomes model the observed outcome under each possible action (or intervention)

Rubin, 1974 **Neyman et al., 1923** **Rubin, 2005**

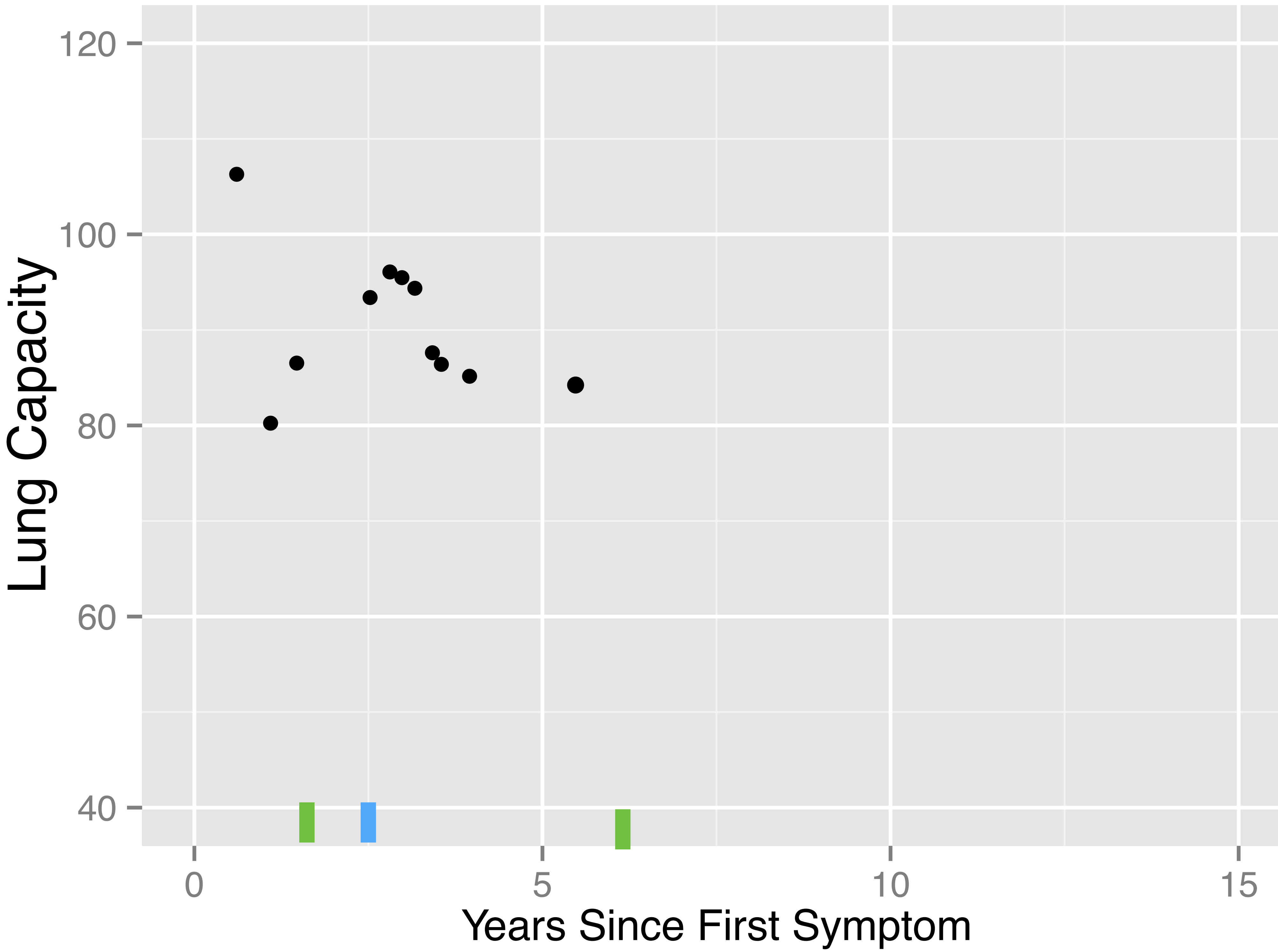
Sequential Decisions in Continuous-Time



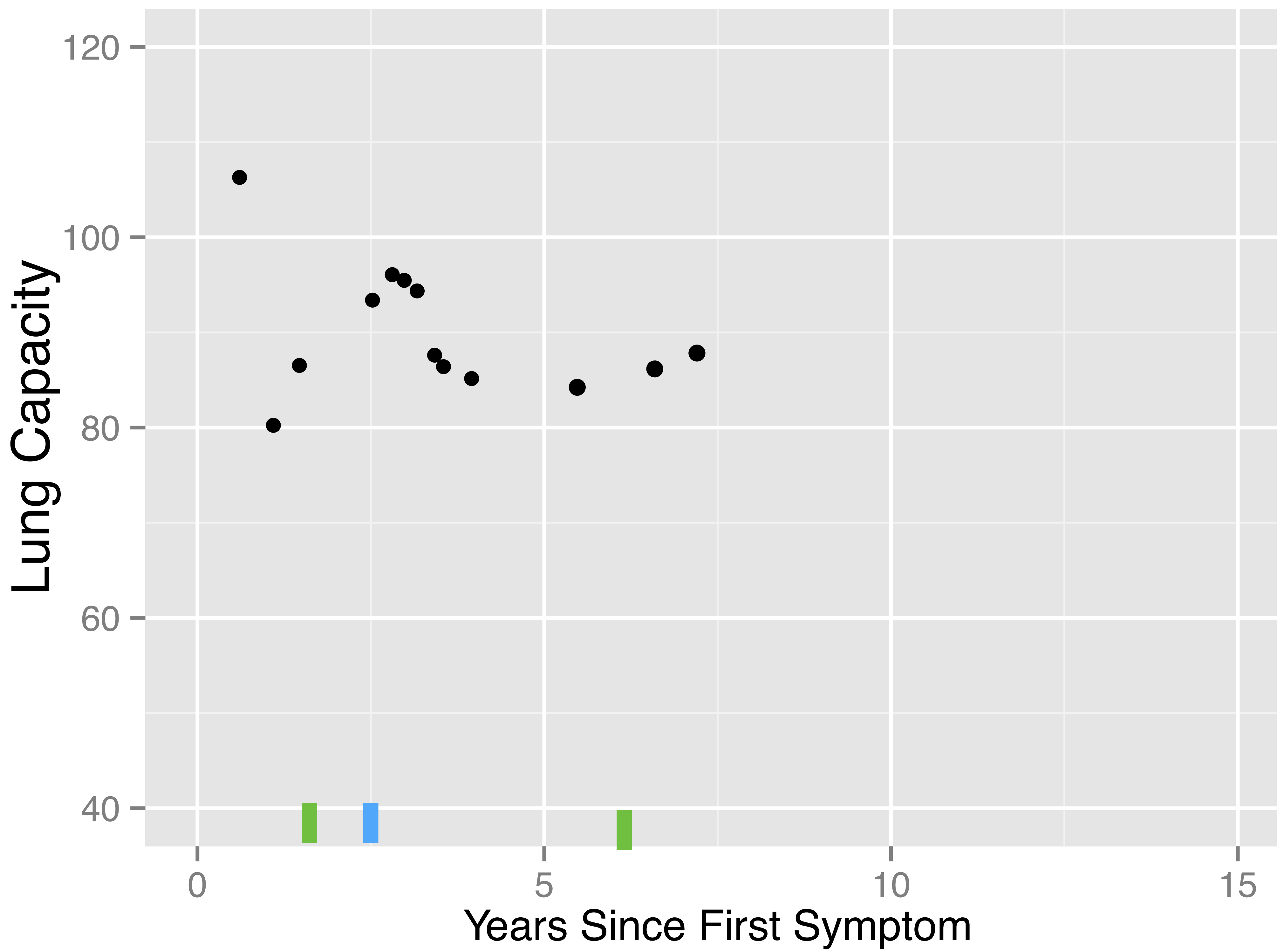
Sequential Decisions in Continuous-Time



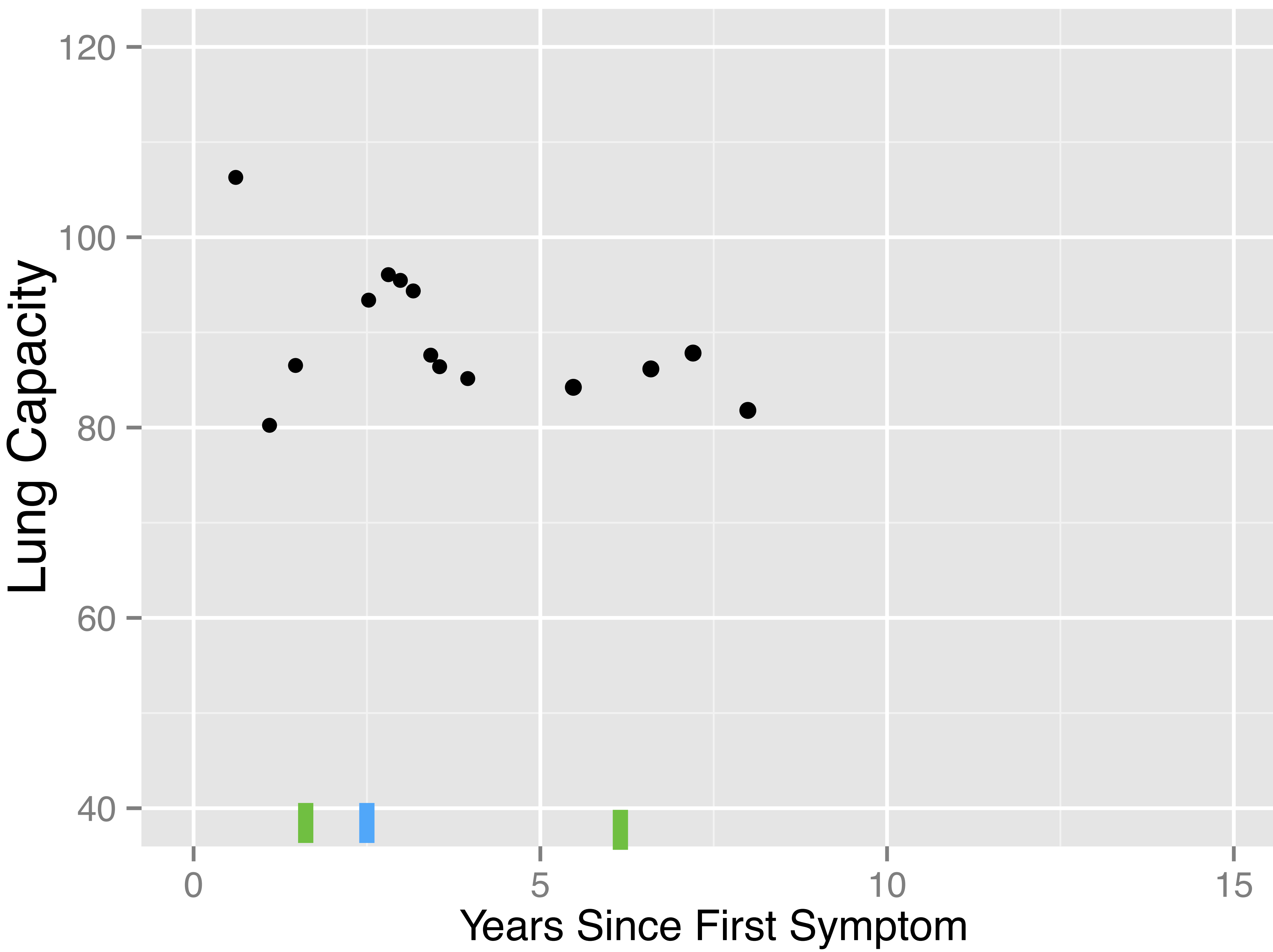
Sequential Decisions in Continuous-Time



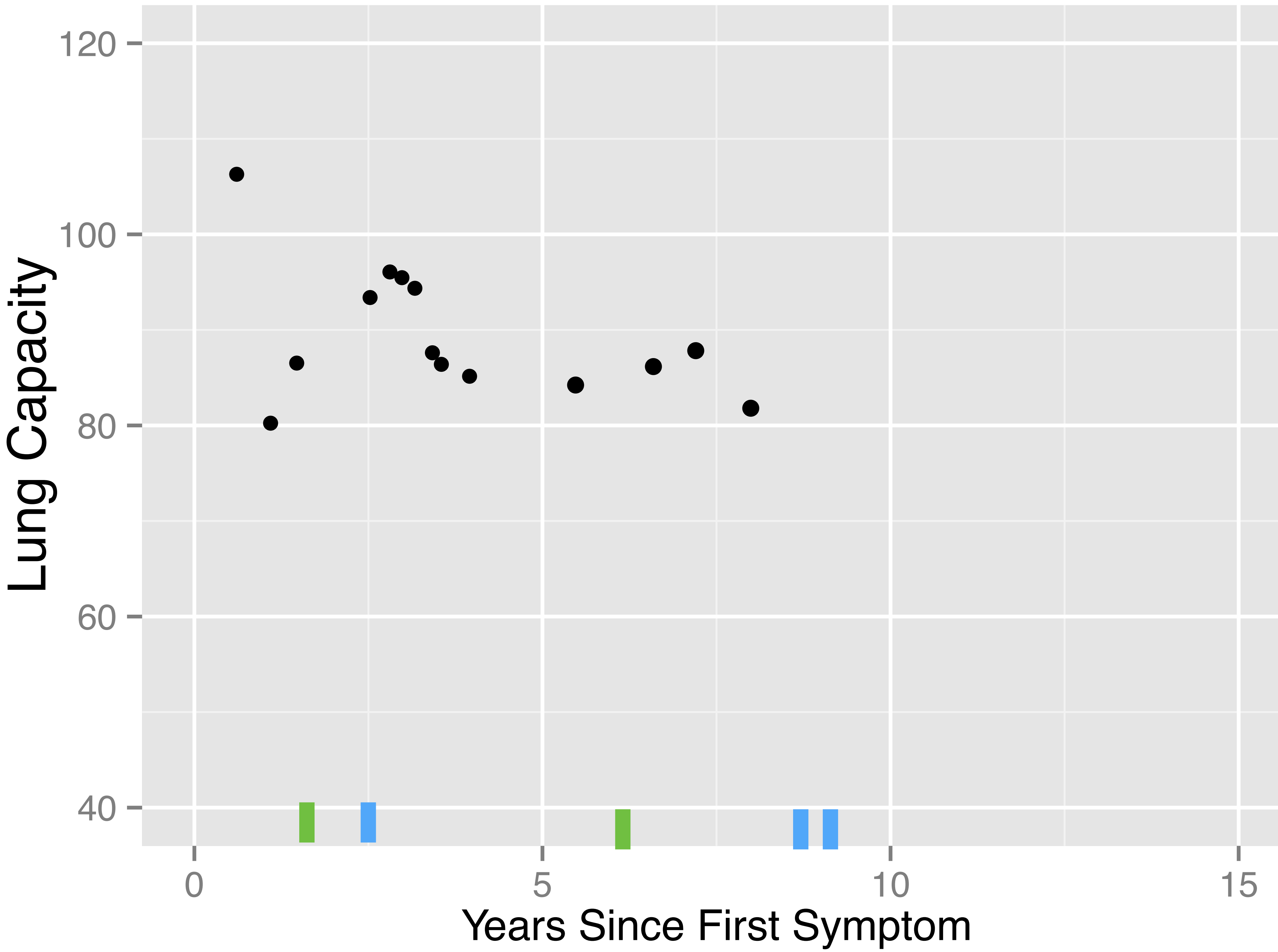
Sequential Decisions in Continuous-Time



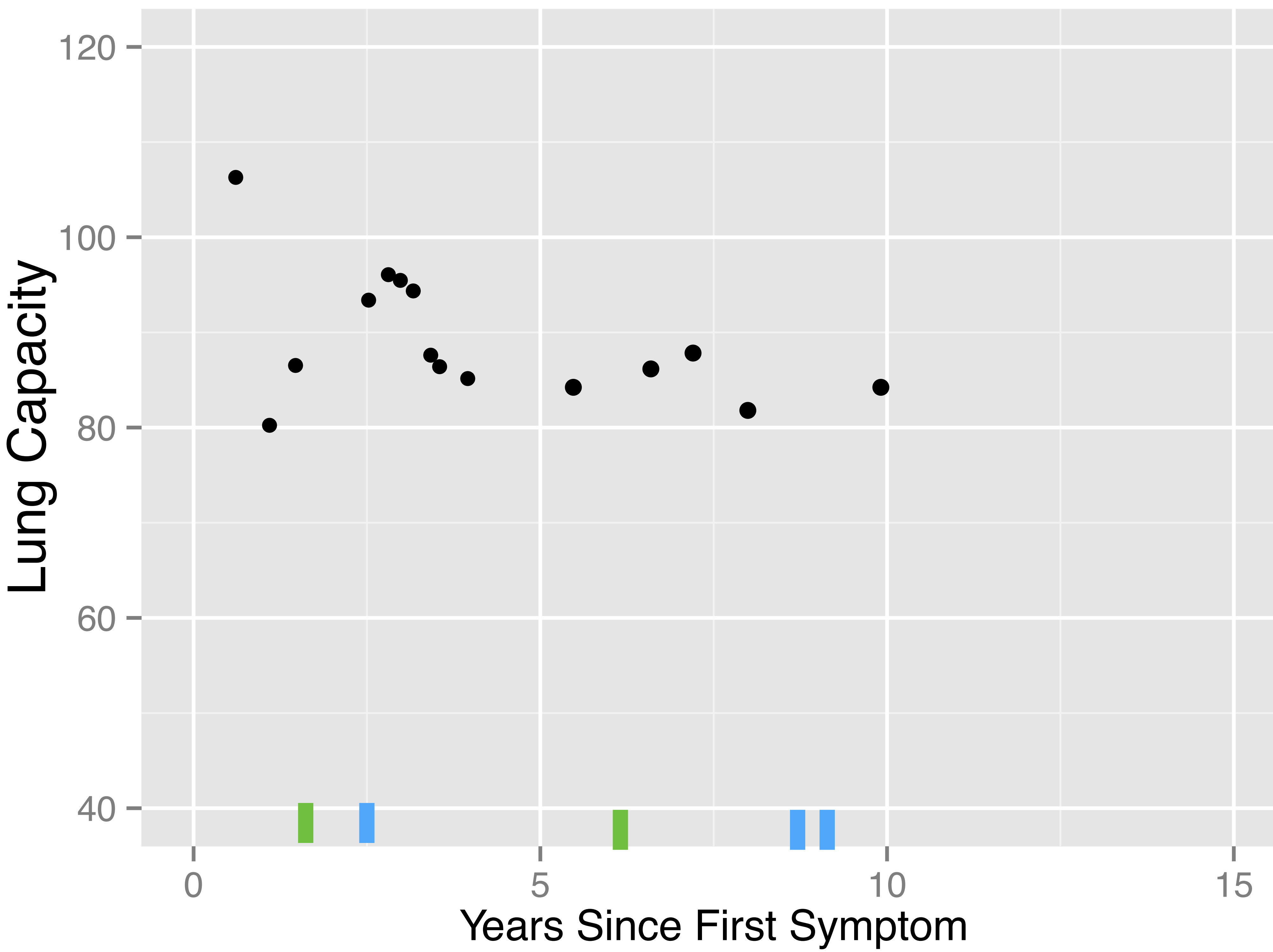
Sequential Decisions in Continuous-Time



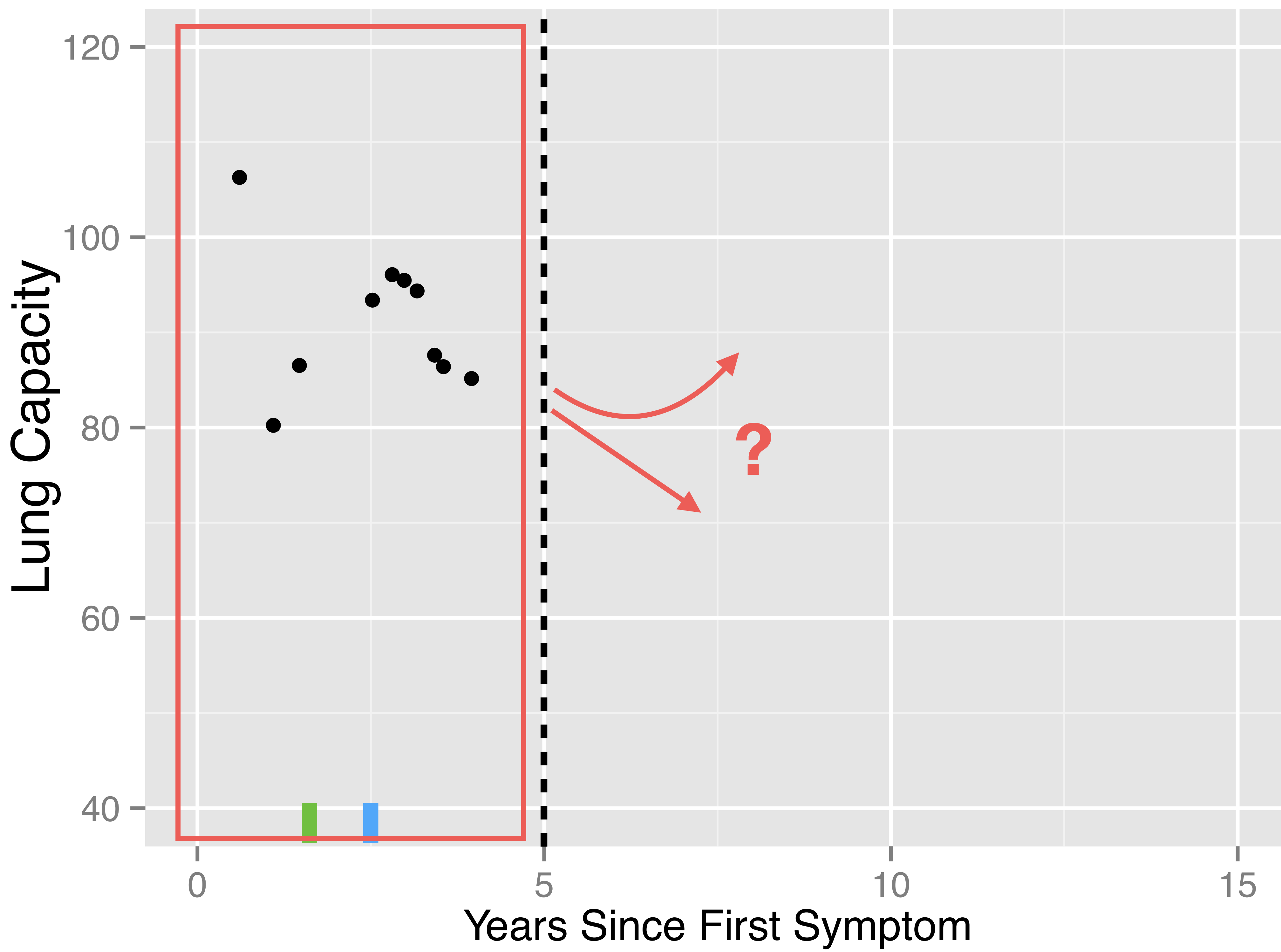
Sequential Decisions in Continuous-Time



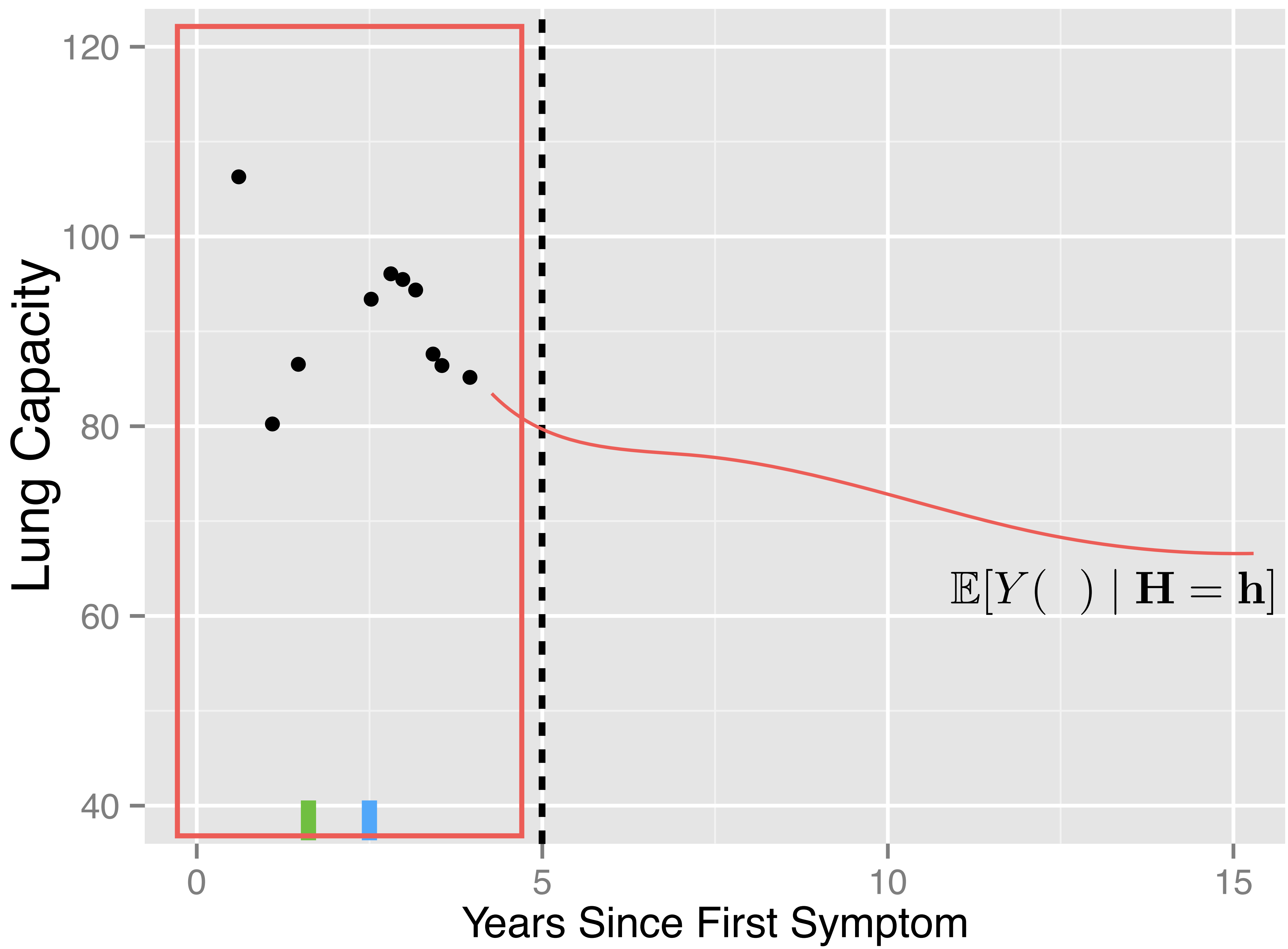
Sequential Decisions in Continuous-Time



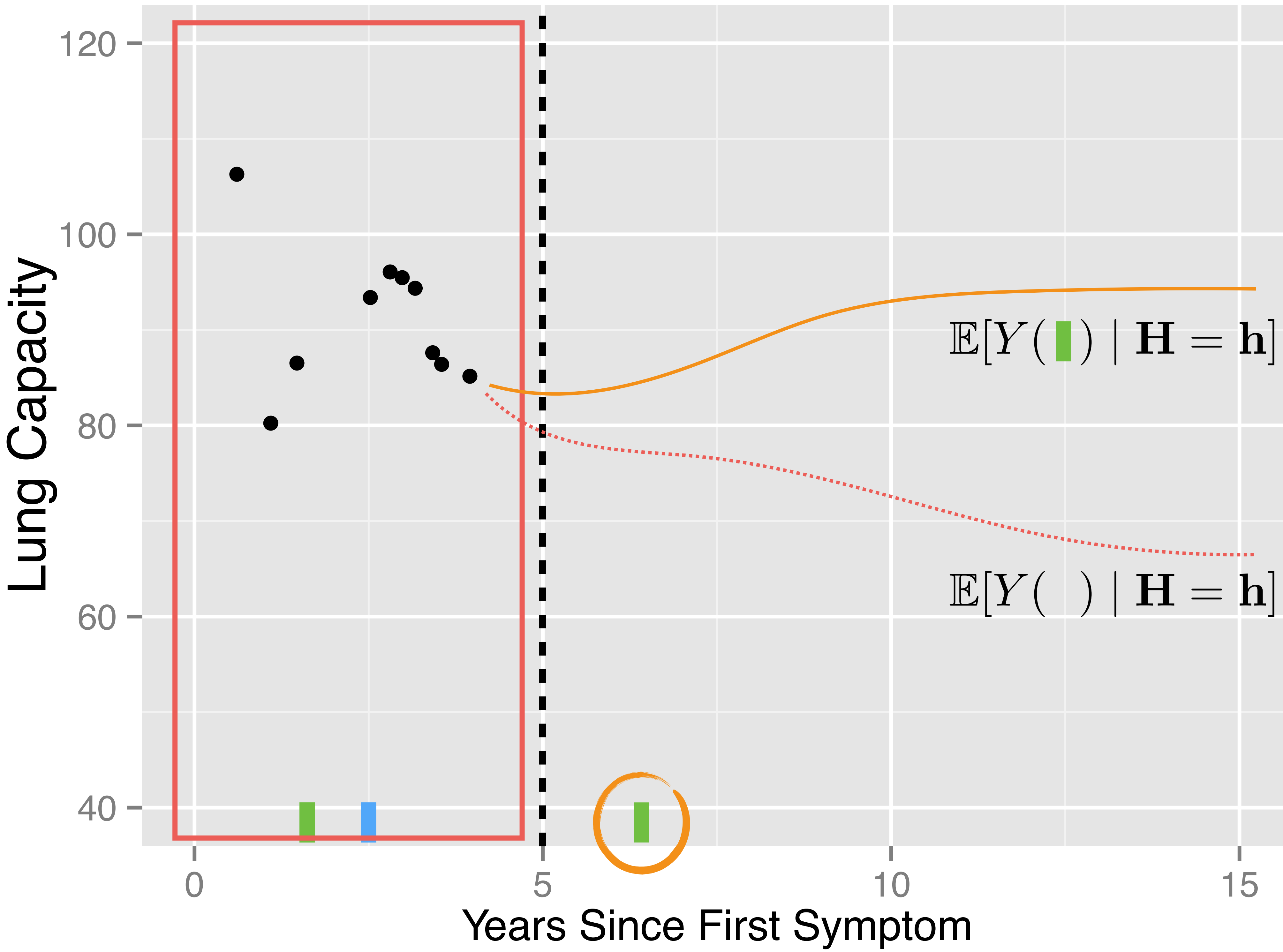
Counterfactual GP



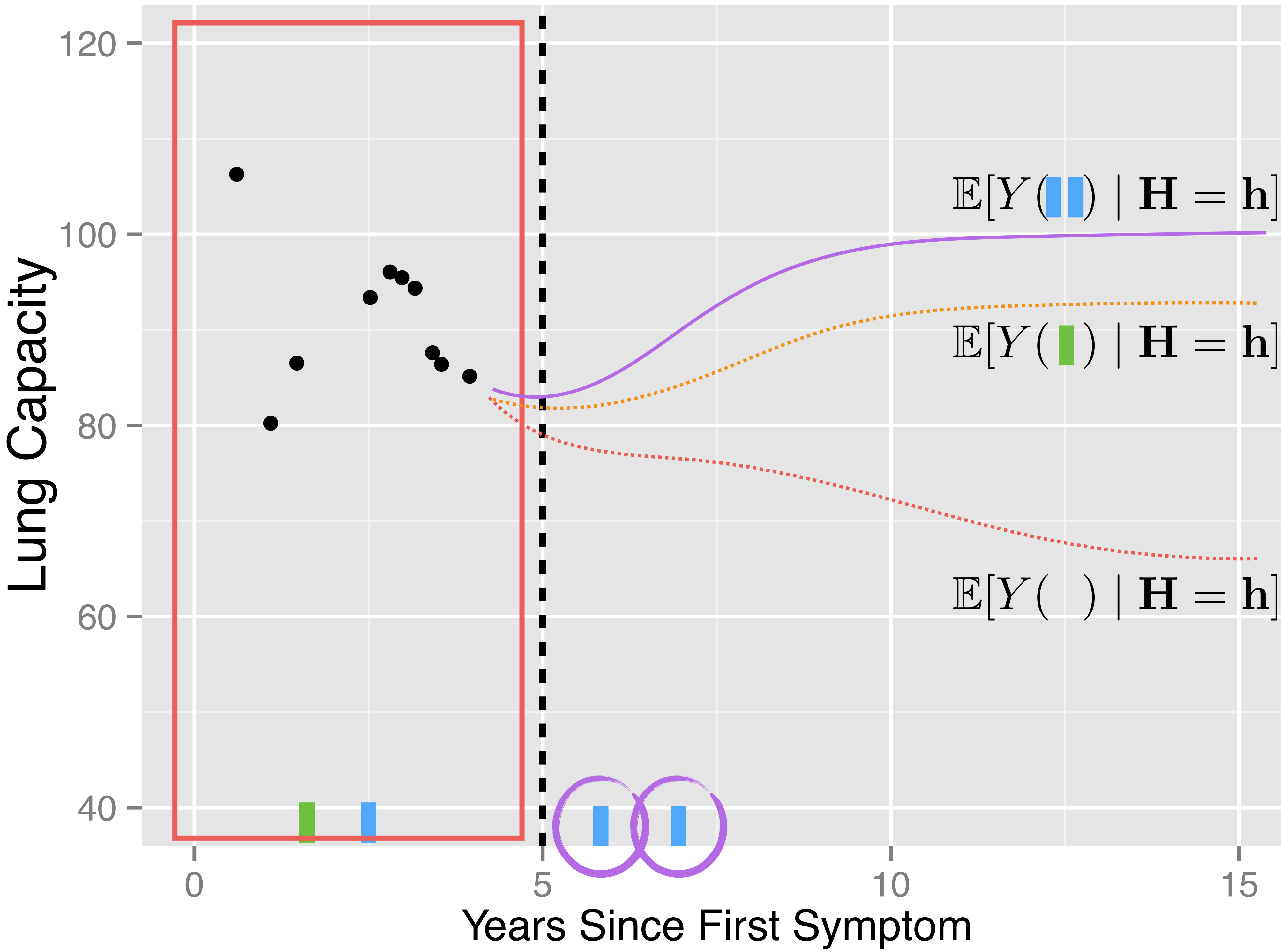
Counterfactual GP



Counterfactual GP



Counterfactual GP



Related Work

- Counterfactual models: See Schulam and Saria, NIPS 2017 for discussion of related work. **Schulam Saria, 2017**

Brodersen et al., 2015 ads; single intervention

Bottou et al., 2013

Taubman et al.,2009 epidemiology; multiple sequential interventions

---

Xu, Xu, Saria, 2016 sparse, irregularly sampled longitudinal data; functional outcomes

Lok et al., 2008

- Off-policy evaluation: Re-weighting to evaluate reward for a policy when learning from offline data.

  e.g. **Dudik et al., 2011** **Jiang and Li, 2016** **Paduraru et al. 2013**

Critical Assumptions

- To learn the potential outcome models, we will use three important assumptions:
- (1) Consistency
  - Links observed outcomes to potential outcomes
- (2) Treatment Positivity
  - Ensures that we can learn potential outcome models
- (3) No unmeasured confounders (NUC)
  - Ensures that we do not learn biased models

Rubin, 1974 **Neyman et al., 1923** **Rubin, 2005**

(1) Consistency

- Consider a dataset containing observed outcomes, observed treatments, and covariates:

$$\{y_i, a_i, \mathbf{x}_i\}_{i=1}^{n}$$

  - E.g.: blood pressure, exercise, BMI

- Consistency allows us to replace the observed response with the potential outcome of the observed treatment

$$Y \triangleq Y(a) \mid A = a$$

- Under consistency our dataset satisfies

$$\{y_i, a_i, \mathbf{x}_i\}_{i=1}^{n} \triangleq \{y_i(a_i), a_i, \mathbf{x}_i\}_{i=1}^{n}$$

(2) Positivity

- When working with observational data, for any set of covariates $\mathbf{x}$ we need to **assume a non-zero probability of seeing each treatment**
  - Otherwise, in general, cannot learn a conditional model of the potential outcomes given those covariates
- Formally, we assume that

$$\mathrm{P}_{\mathrm{Obs}}(A = a \mid \mathbf{X} = \mathbf{x}) > 0 \quad \forall a \in \mathcal{A}, \forall \mathbf{x} \in \mathcal{X}$$

(3) No Unmeasured Confounders (NUC)

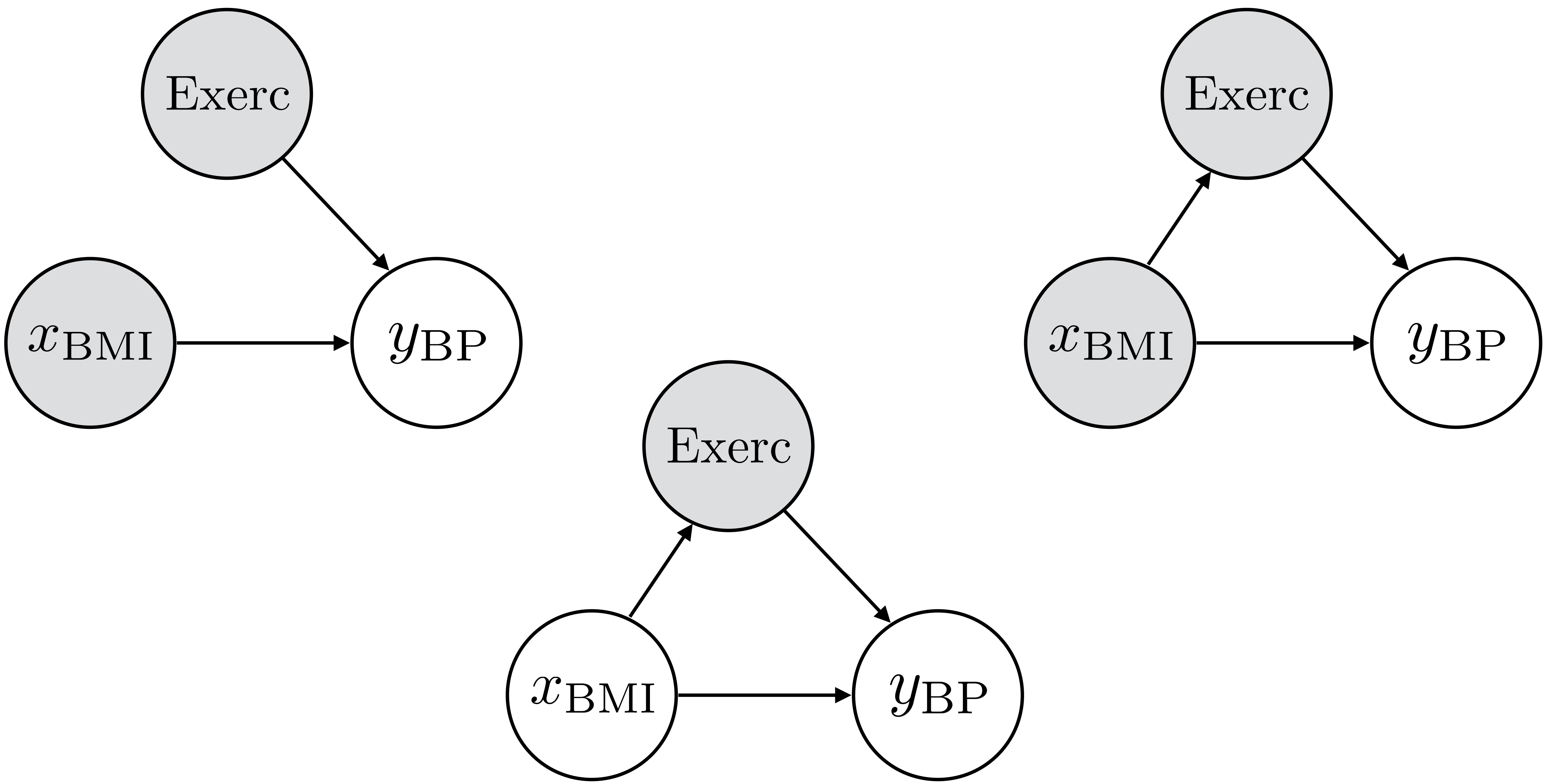
- Formally, NUC is an statistical independence assertion:

$$Y(a) \perp A \mid \mathbf{X} = \mathbf{x} \quad : \quad \forall a \in \mathcal{A}, \forall \mathbf{x} \in \mathcal{X}$$

(3) No Unmeasured Confounders (NUC)

- Formally, NUC is an statistical independence assertion:

$$Y(a) \perp A \mid \mathbf{X} = \mathbf{x} \quad : \quad \forall a \in \mathcal{A}, \forall \mathbf{x} \in \mathcal{X}$$



Learning Potential Outcome Models

- Assumptions allow estimation of potential outcomes from (observational) data:

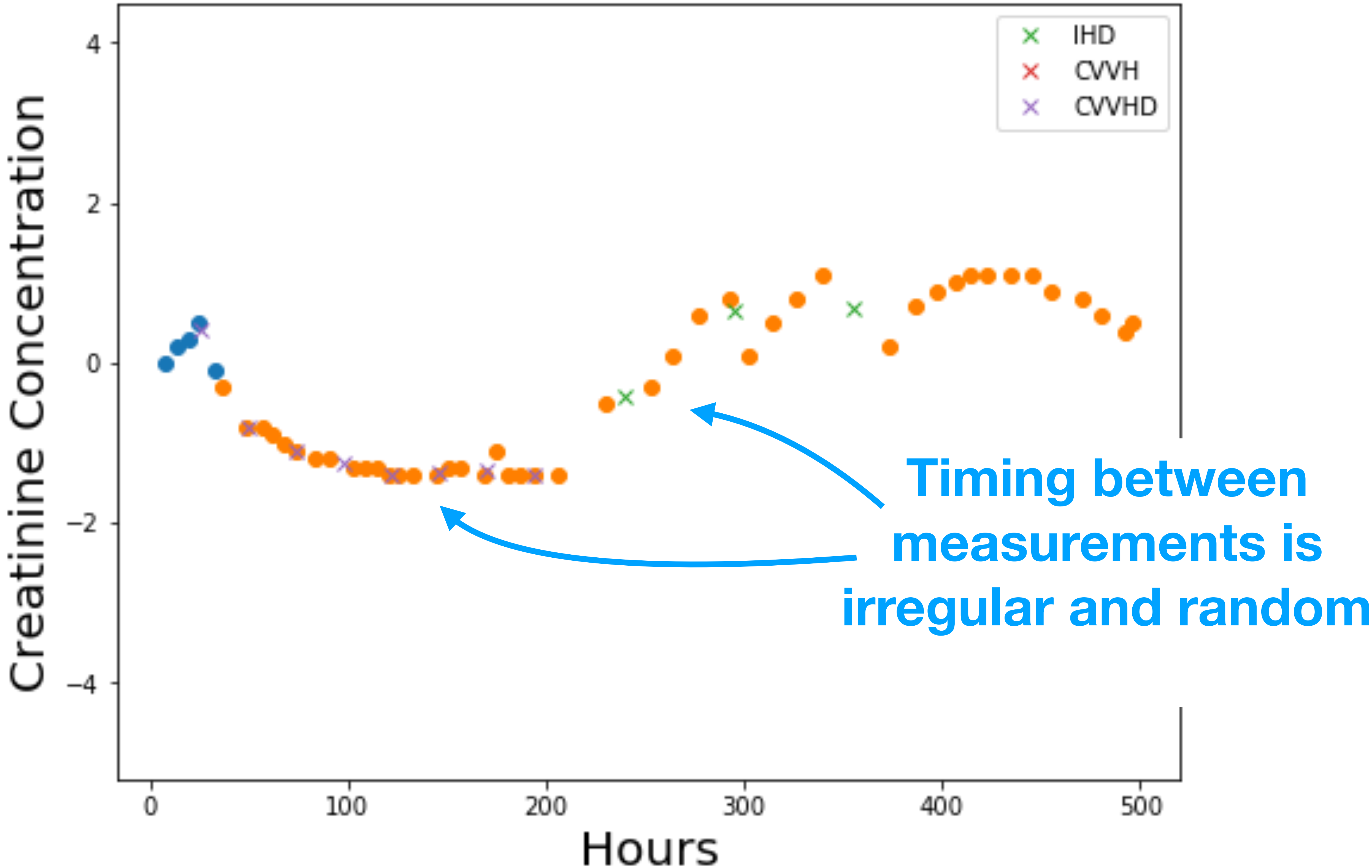
$$\mathrm{P}(Y(a) \mid \mathbf{X} = \mathbf{x}) = \mathrm{P}(Y(a) \mid \mathbf{X} = \mathbf{x}, A = a) \quad \text{(A3)}$$
$$= \boxed{\mathrm{P}(Y \mid \mathbf{X} = \mathbf{x}, A = a)} \quad \text{(A1)}$$

Estimation requires a statistical model for estimating conditionals

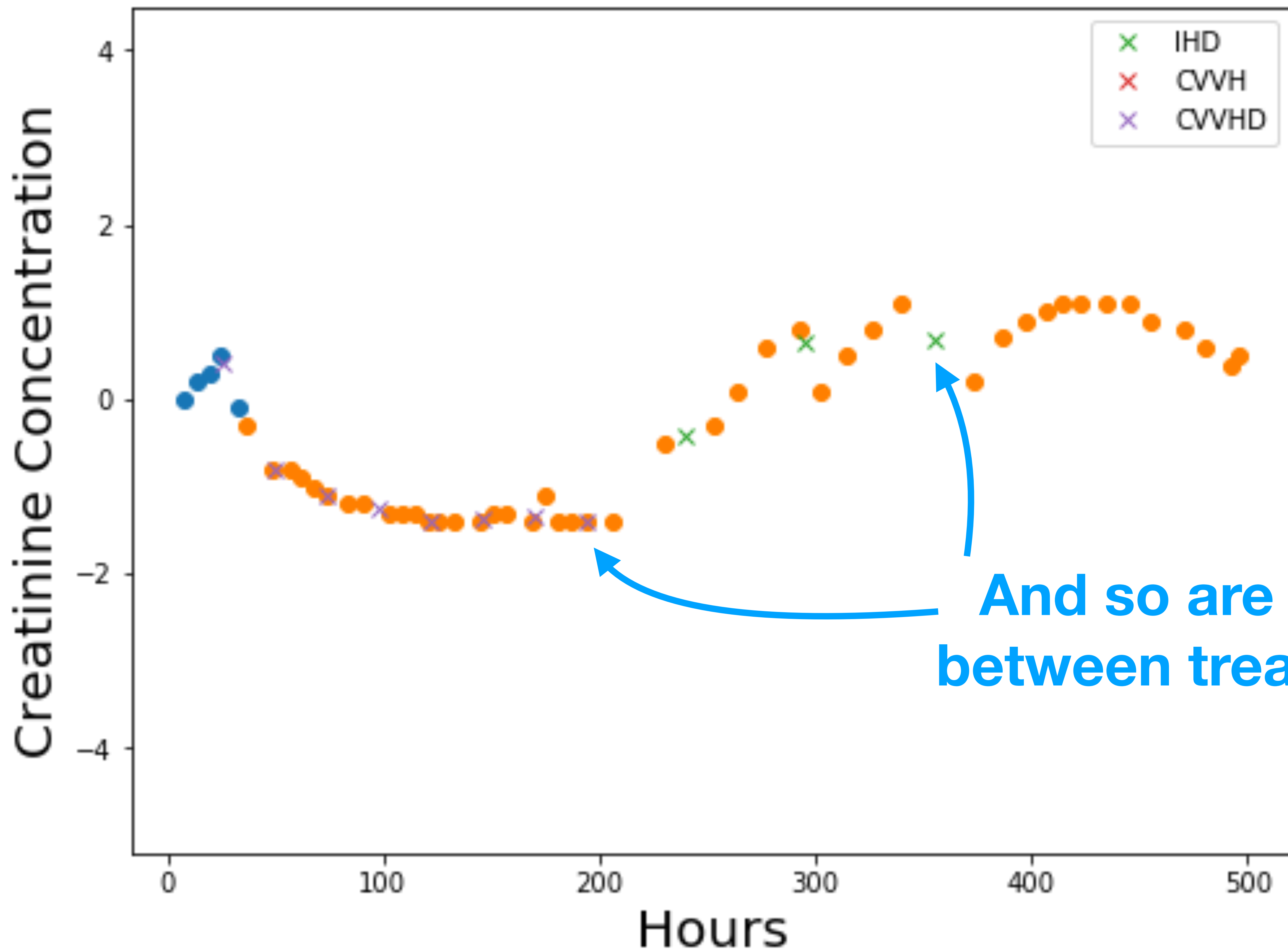
- To simulate data from a new policy, we need to learn the potential outcome models
  - If we have an observational dataset where assumptions 1-3 hold, then this is possible!

Observational Traces

- Creatinine is a test used to measure kidney function.

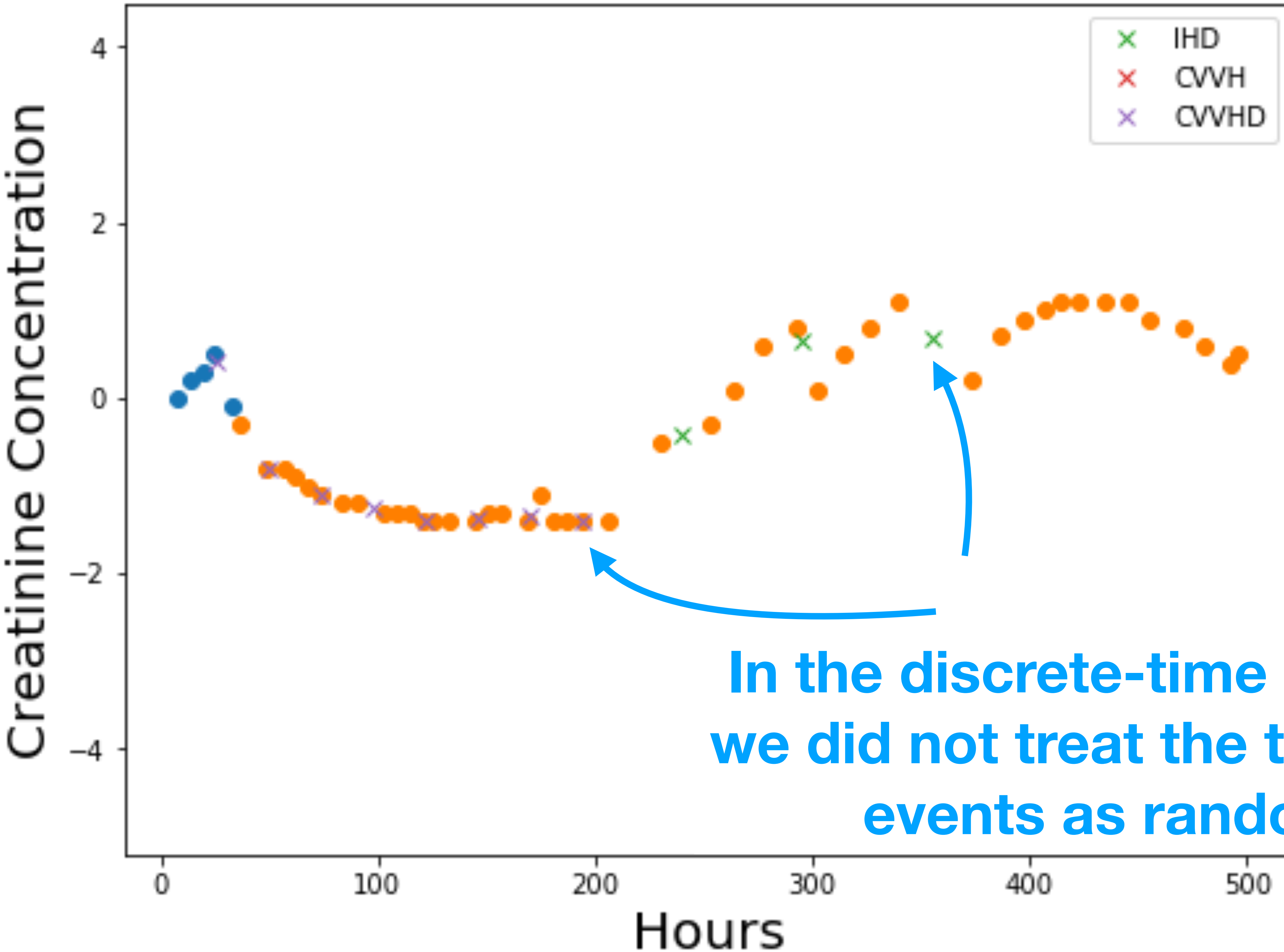


Observational Traces



And so are times between treatments

Challenges w/ Observational Traces



In the discrete-time setting, we did not treat the timing of events as random

Counterfactual GP

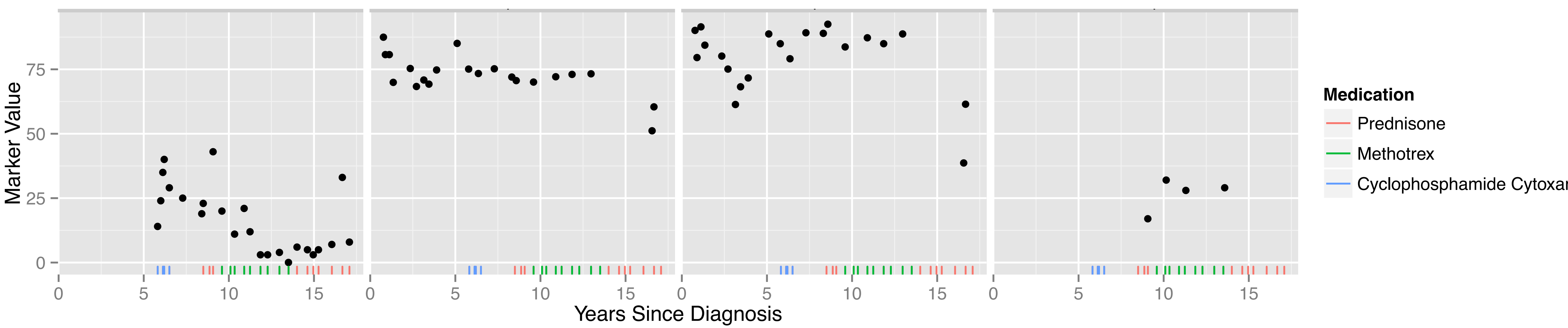
- Collection of Gaussian processes

$$\left\{ \{Y_t(\boldsymbol{a}) : t \in [0, \tau]\} : \boldsymbol{a} \in \mathcal{C} \right\}$$

Fixed time period

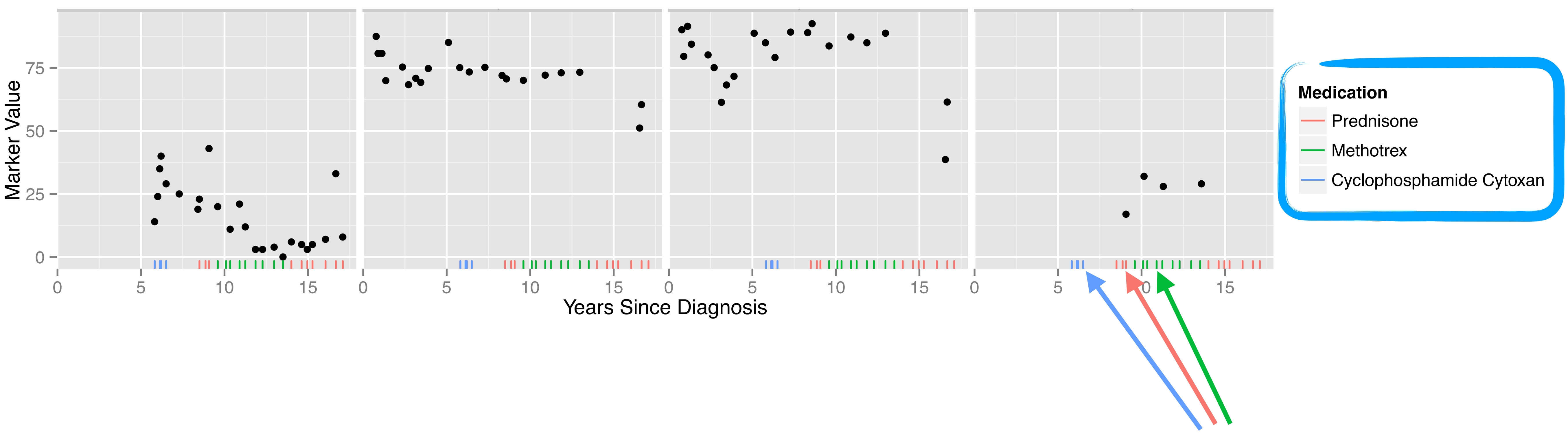
Set of finite sequences of actions

Learning from Observational Traces



$$\mathcal{D} \triangleq \left\{ \mathbf{h}_i = \{(t_{ij}, y_{ij}, a_{ij})\}_{j=1}^{n_i} \right\}_{i=1}^{m}$$

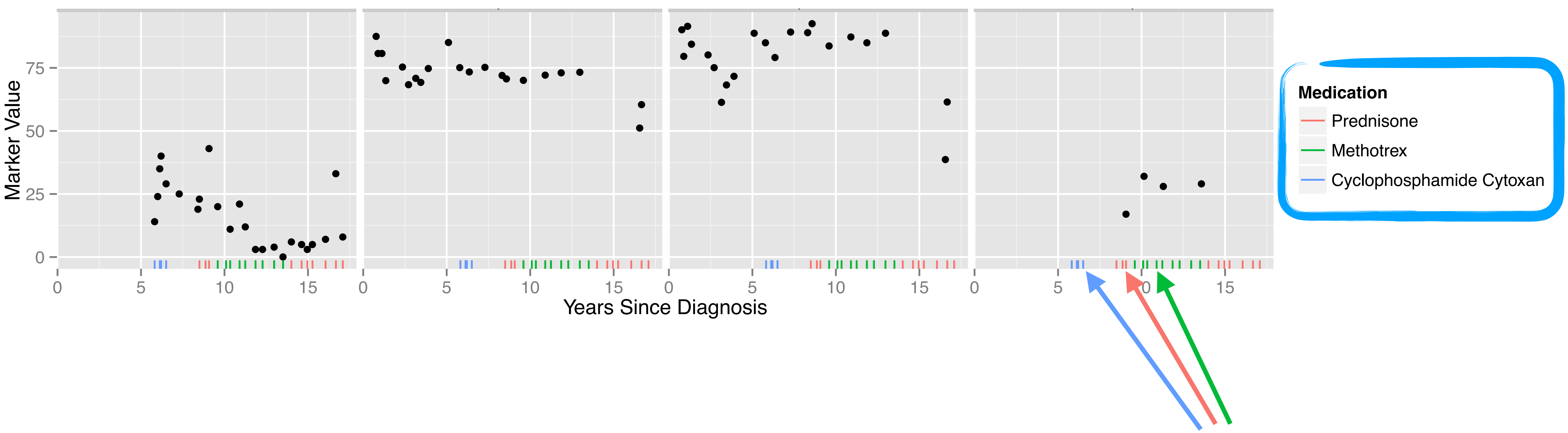
Learning from Observational Traces



$$\mathcal{D} \triangleq \left\{ \mathbf{h}_i = \{(t_{ij}, y_{ij}, a_{ij})\}_{j=1}^{n_i} \right\}_{i=1}^{m}$$

Treatments administered according to unknown policy (i.e. not an RCT)

Learning from Observational Traces



$$\mathcal{D} \triangleq \left\{ \mathbf{h}_i = \{(t_{ij}, y_{ij}, a_{ij})\}_{j=1}^{n_i} \right\}_{i=1}^{m}$$

Learning is especially difficult because there is time-dependent *feedback* between actions and outcomes

Robins 1986

Learning Models from Observational Traces

- Road map:
  - (1) Establish assumptions that connect probabilistic of observational traces to *target counterfactual model*
  - (2) Posit probabilistic model of observational traces
  - (3) Derive maximum likelihood estimator

$$P(\{Y_s[\mathbf{a}] : s > t\} \mid \mathcal{H}_t)$$

Schulam and Saria, NIPS 2017

Modeling Observational Traces

- We use a marked point process (MPP):

$$\{(T_i, X_i)\}_{i=1}^{\infty}$$

- Points model the *event times*: measurements or actions

- Mark models the type of event

$$\mathcal{X} = (\mathbb{R} \cup \{\varnothing\}) \times (\mathcal{C} \cup \{\varnothing\}) \times \{0, 1\} \times \{0, 1\}$$

Schulam and Saria, NIPS 2017

Modeling Observational Traces

- We use a marked point process (MPP):

$$\{(T_i, X_i)\}_{i=1}^{\infty}$$

- Points model the *event times*: measurements or actions

- Mark models the type of event

$$\mathcal{X} = (\mathbb{R} \cup \{\varnothing\}) \times (\mathcal{C} \cup \{\varnothing\}) \times \{0, 1\} \times \{0, 1\}$$

z_y

Did we measure an outcome?

Modeling Observational Traces

- We use a marked point process (MPP):

$$\{(T_i, X_i)\}_{i=1}^{\infty}$$

- Points model the *event times*: measurements or actions

- Mark models the type of event

$$\mathcal{X} = (\mathbb{R} \cup \{\varnothing\}) \times (\mathcal{C} \cup \{\varnothing\}) \times \underset{z_y}{\{0, 1\}} \times \underset{z_a}{\{0, 1\}}$$

Did we take an action?

Modeling Observational Traces

- We use a marked point process (MPP):

$$\{(T_i, X_i)\}_{i=1}^{\infty}$$

- Points model the *event times*: measurements or actions

- Mark models the type of event

$$\underbrace{\mathcal{X} = (\mathbb{R} \cup \{\varnothing\})}_{y} \times (\mathcal{C} \cup \{\varnothing\}) \times \underbrace{\{0, 1\}}_{z_y} \times \underbrace{\{0, 1\}}_{z_a}$$

What is the value of the outcome?

Modeling Observational Traces

- We use a marked point process (MPP):

$$\{(T_i, X_i)\}_{i=1}^{\infty}$$

- Points model the *event times*: measurements or actions

- Mark models the type of event

$$\mathcal{X} = \underbrace{(\mathbb{R} \cup \{\varnothing\})}_{y} \times \underbrace{(\mathcal{C} \cup \{\varnothing\})}_{a} \times \underbrace{\{0,1\}}_{z_y} \times \underbrace{\{0,1\}}_{z_a}$$

What action did we take?

Modeling Observational Traces

- Parameterize MPP using hazard and mark density:

$$\lambda^*(t, x) = \lambda^*(t) p^*(x \mid t)$$

Schulam and Saria, NIPS 2017

Modeling Observational Traces

- Parameterize MPP using hazard and mark density:

$$\lambda^*(t, x) = \lambda^*(t) p^*(x \mid t)$$

Probability of event happening at this time

Probability of mark given event time

Schulam and Saria, NIPS 2017

Modeling Observational Traces

- Parameterize MPP using hazard and mark density:

$$\lambda^*(t, x) = \lambda^*(t)p^*(x \mid t)$$

Probability of event happening at this time

Probability of mark given event time

Star denotes dependence on history

Schulam and Saria, NIPS 2017

Modeling Observational Traces

- Parameterize MPP using hazard and mark density:

$$\lambda^*(t, x) = \lambda^*(t) p^*(x \mid t)$$

- Estimate MPP by maximizing probability of traces

$$\ell(\theta) = \sum_{j=1}^{n} \log p_\theta^*(y_j \mid t_j, z_{yj}) + \sum_{j=1}^{n} \log \lambda_\theta^*(t) p_\theta^*(a_j, z_{yj}, z_{aj} \mid t_j, y_j) - \int_0^\tau \lambda_\theta^*(s) ds$$

Model the conditional probability of the outcome using a GP

Schulam and Saria, NIPS 2017

Recovering the CGP

- When does the MPP GP recover the CGP?
- In addition to Consistency, we define two assumptions

Recovering the CGP

- When does the MPP GP recover the CGP?
- In addition to Consistency, we define two assumptions
- Continuous-time NUC
    - Analogue of NUC for MPP

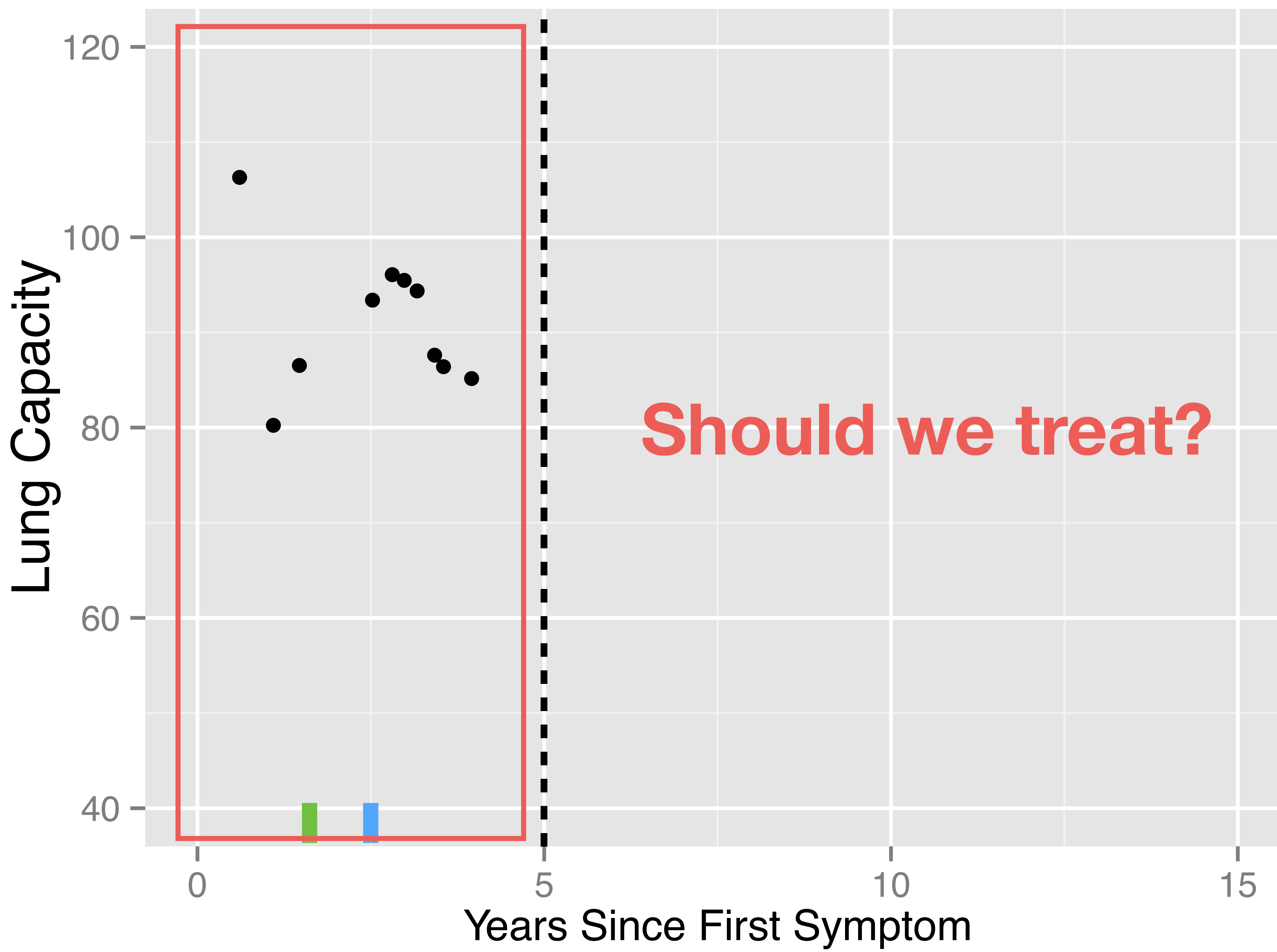
Schulam and Saria, NIPS 2017

Recovering the CGP

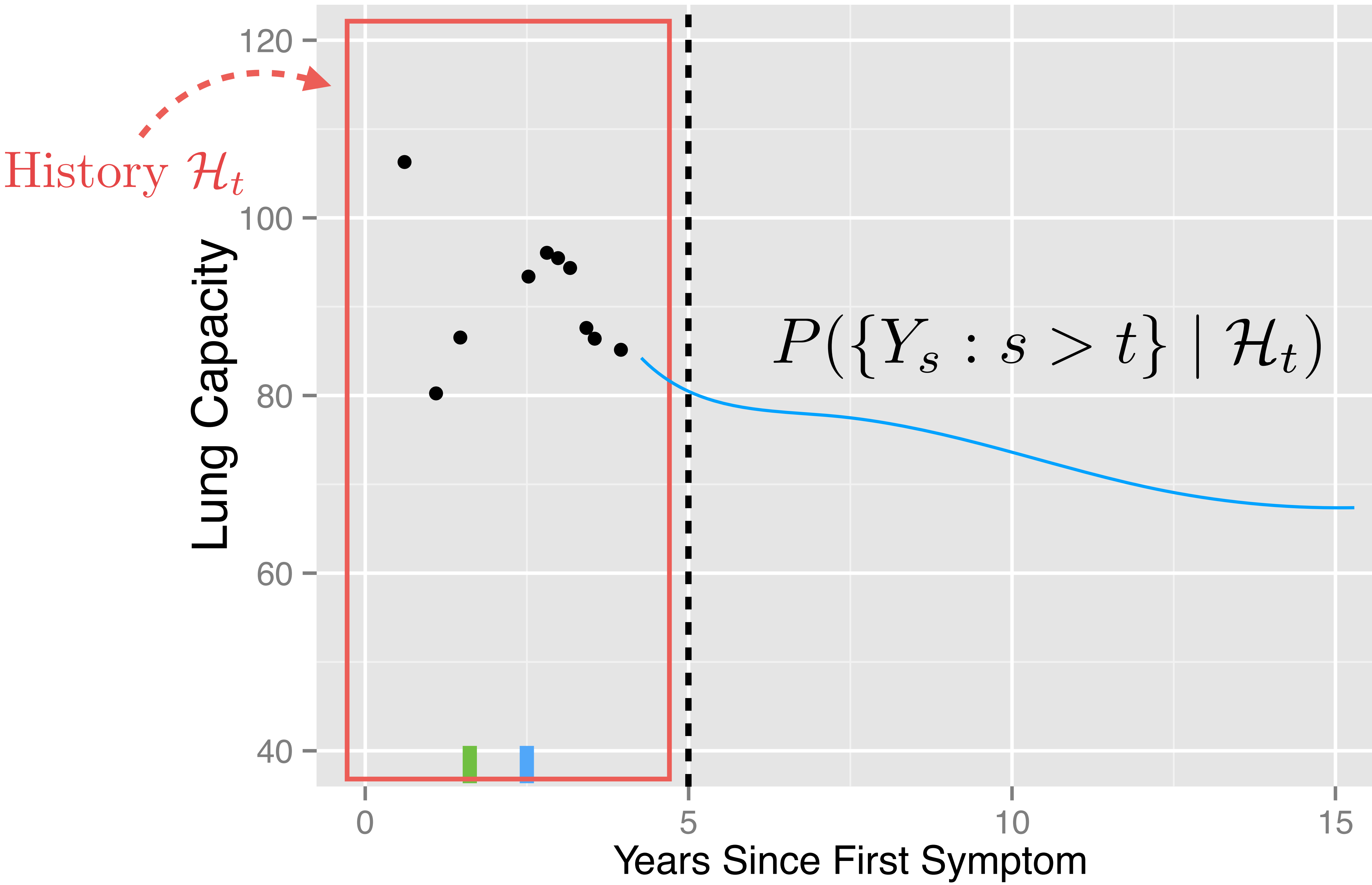
- When does the MPP GP recover the CGP?
- In addition to Consistency, we define two assumptions
- Continuous-time NUC
  - Analogue of NUC for MPP
- Non-informative measurement times
  - Measurement and action times are conditionally independent of potential outcomes

Schulam and Saria, NIPS 2017

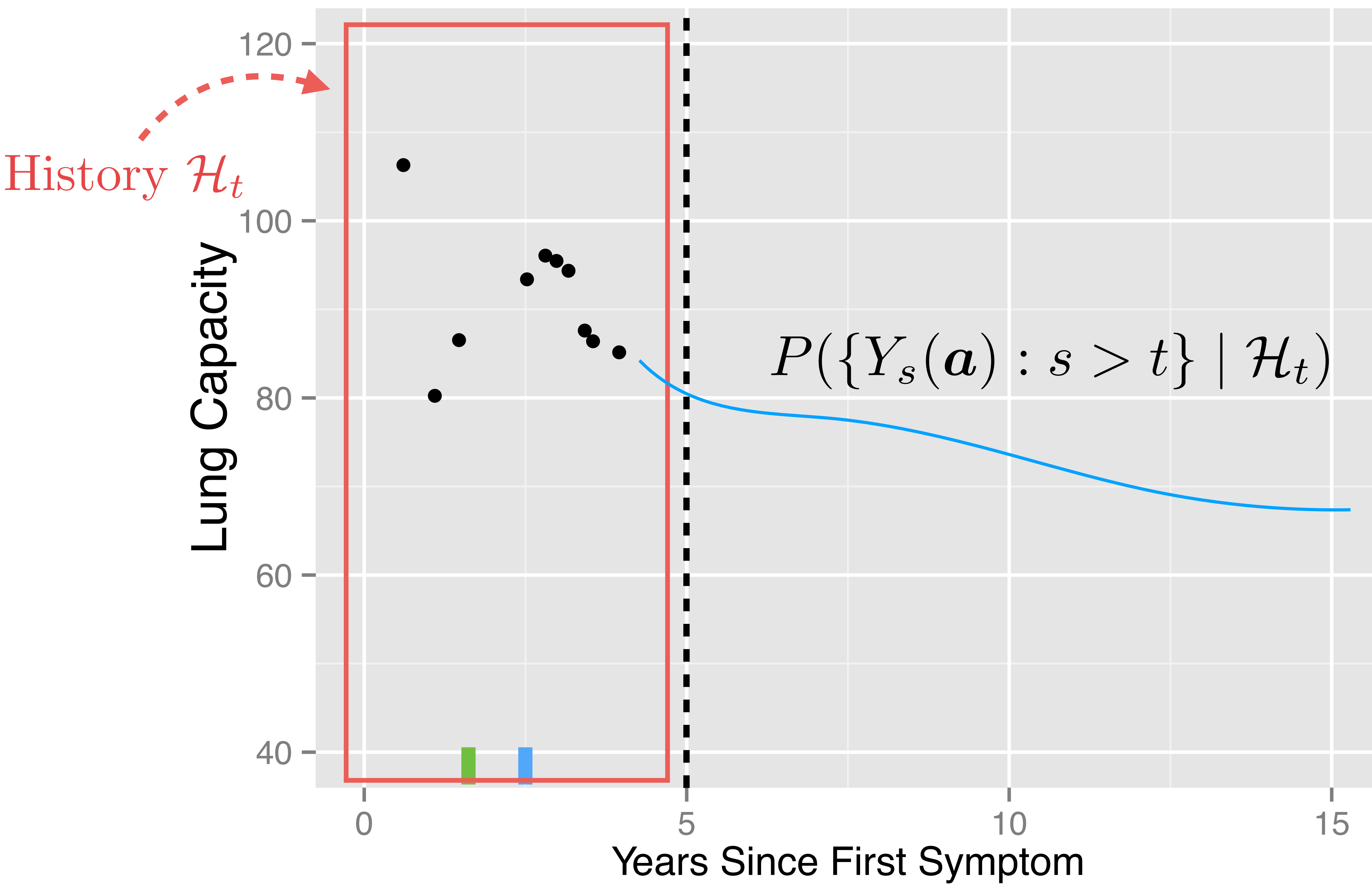
Reliable Decisions with CGPs



Classical Supervised Model



Counterfactual GP



Simulated Data

- Simulate observational traces from multiple regimes
- Traces are treated by policies unknown to learners
- In regimes A and B, policies satisfy our assumptions
- In regime C, policy violates our assumptions
- Simulate three training sets (regimes A, B, and C)
- Simulate one common test set (regime A)

Results

- Risk scores:
  - Use Baseline and CGP to predict final severity marker
  - Normalize predictions to [0, 1]

Results

- Risk scores:
  - Use Baseline and CGP to predict final severity marker
  - Normalize predictions to [0, 1]

CGP risk scores are stable across regime A and B training data

	Regime A		Regime B		Regime C	
	Baseline GP	CGP	Baseline GP	CGP	Baseline GP	CGP
Risk Score Δ from A	0.000	0.000	0.083	0.001	0.162	0.128
Kendall's τ from A	1.000	1.000	0.857	0.998	0.640	0.562
AUC	0.853	0.872	0.832	0.872	0.806	0.829

Results

- Risk scores:
  - Use Baseline and CGP to predict final severity marker
  - Normalize predictions to [0, 1]

Baseline GP scores change

	Regime A		Regime B		Regime C	
	Baseline GP	CGP	Baseline GP	CGP	Baseline GP	CGP
Risk Score Δ from A	0.000	0.000	0.083	0.001	0.162	0.128
Kendall's τ from A	1.000	1.000	0.857	0.998	0.640	0.562
AUC	0.853	0.872	0.832	0.872	0.806	0.829

Results

- Risk scores:
  - Use Baseline and CGP to predict final severity marker
  - Normalize predictions to [0, 1]

CGP relative risk across patients is also stable across training data A and B

	Regime A		Regime B		Regime C	
	Baseline GP	CGP	Baseline GP	CGP	Baseline GP	CGP
Risk Score Δ from A	0.000	0.000	0.083	0.001	0.162	0.128
Kendall's τ from A	1.000	1.000	0.857	0.998	0.640	0.562
AUC	0.853	0.872	0.832	0.872	0.806	0.829

Results

- Risk scores:
  - Use Baseline and CGP to predict final severity marker
  - Normalize predictions to [0, 1]

Baseline GP's relative risk changes

	Regime A		Regime B		Regime C	
	Baseline GP	CGP	Baseline GP	CGP	Baseline GP	CGP
Risk Score Δ from A	0.000	0.000	0.083	0.001	0.162	0.128
Kendall's τ from A	1.000	1.000	0.857	0.998	0.640	0.562
AUC	0.853	0.872	0.832	0.872	0.806	0.829

Results

- Risk scores:
  - Use Baseline and CGP to predict final severity marker
  - Normalize predictions to [0, 1]

CGP AUC is constant across regimes A and B

	Regime A		Regime B		Regime C	
	Baseline GP	CGP	Baseline GP	CGP	Baseline GP	CGP
Risk Score Δ from A	0.000	0.000	0.083	0.001	0.162	0.128
Kendall's τ from A	1.000	1.000	0.857	0.998	0.640	0.562
AUC	0.853	0.872	0.832	0.872	0.806	0.829

Results

- Risk scores:
  - Use Baseline and CGP to predict final severity marker
  - Normalize predictions to [0, 1]

Baseline GP's AUC is unstable

	Regime A		Regime B		Regime C	
	Baseline GP	CGP	Baseline GP	CGP	Baseline GP	CGP
Risk Score Δ from A	0.000	0.000	0.083	0.001	0.162	0.128
Kendall's τ from A	1.000	1.000	0.857	0.998	0.640	0.562
AUC	0.853	0.872	0.832	0.872	0.806	0.829

Simulated Data

- Simulate observational traces from three regimes
- Traces are treated by policies unknown to learners
- In regimes A and B, policies satisfy our assumptions
- In regime C, policy violates our assumptions
- Simulate three training sets (regimes A, B, and C)
- Simulate one common test set (regime A)

Results

- Risk scores:
  - Use Baseline and CGP to predict final severity marker
  - Negate predictions and normalize to [0, 1]

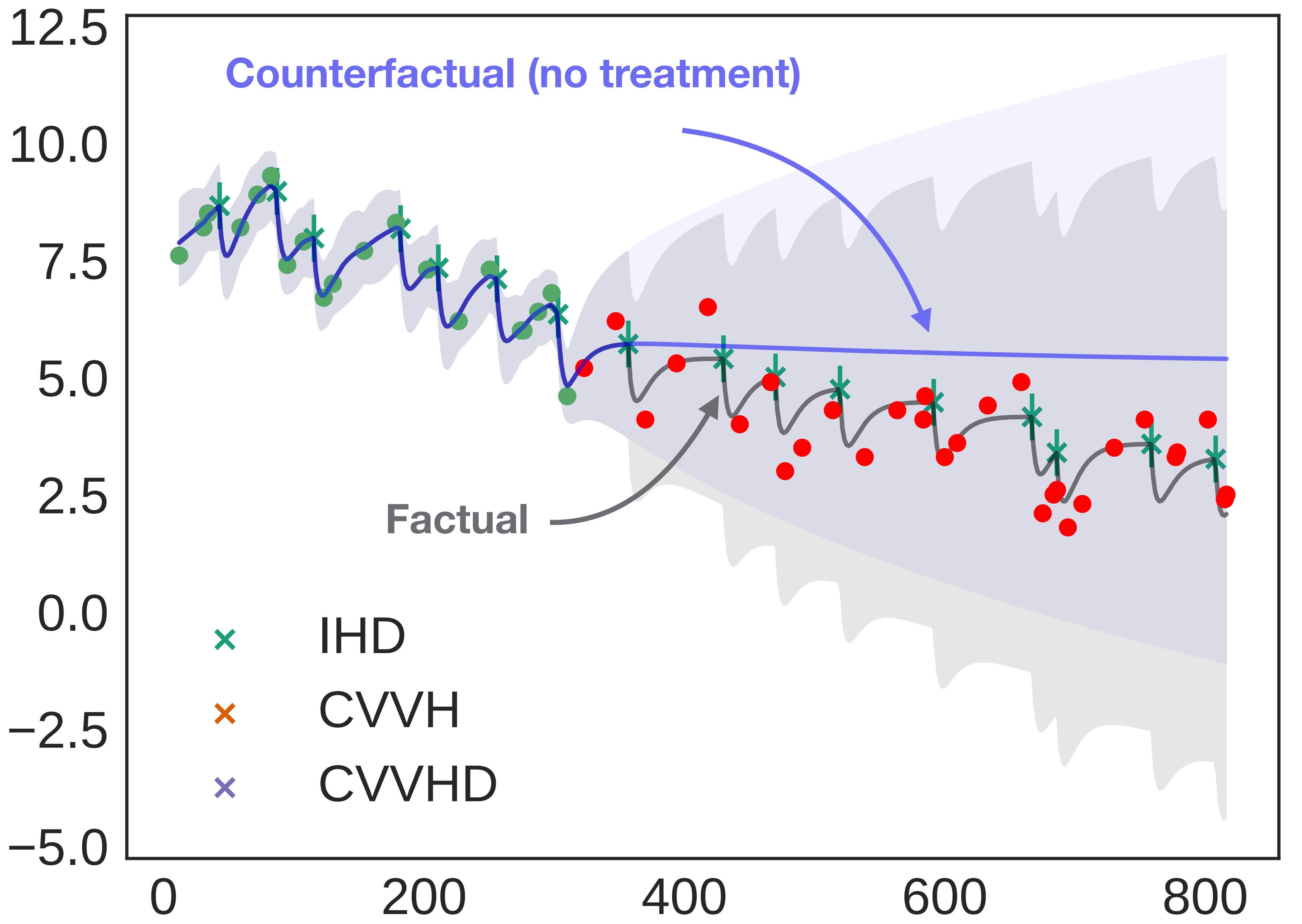
CGP risk scores are unstable if the policy in the training data violates our assumptions

	Regime A		Regime B		Regime C	
	Baseline GP	CGP	Baseline GP	CGP	Baseline GP	CGP
Risk Score Δ from A	0.000	0.000	0.083	0.001	0.162	0.128
Kendall's τ from A	1.000	1.000	0.857	0.998	0.640	0.562
AUC	0.853	0.872	0.832	0.872	0.806	0.829

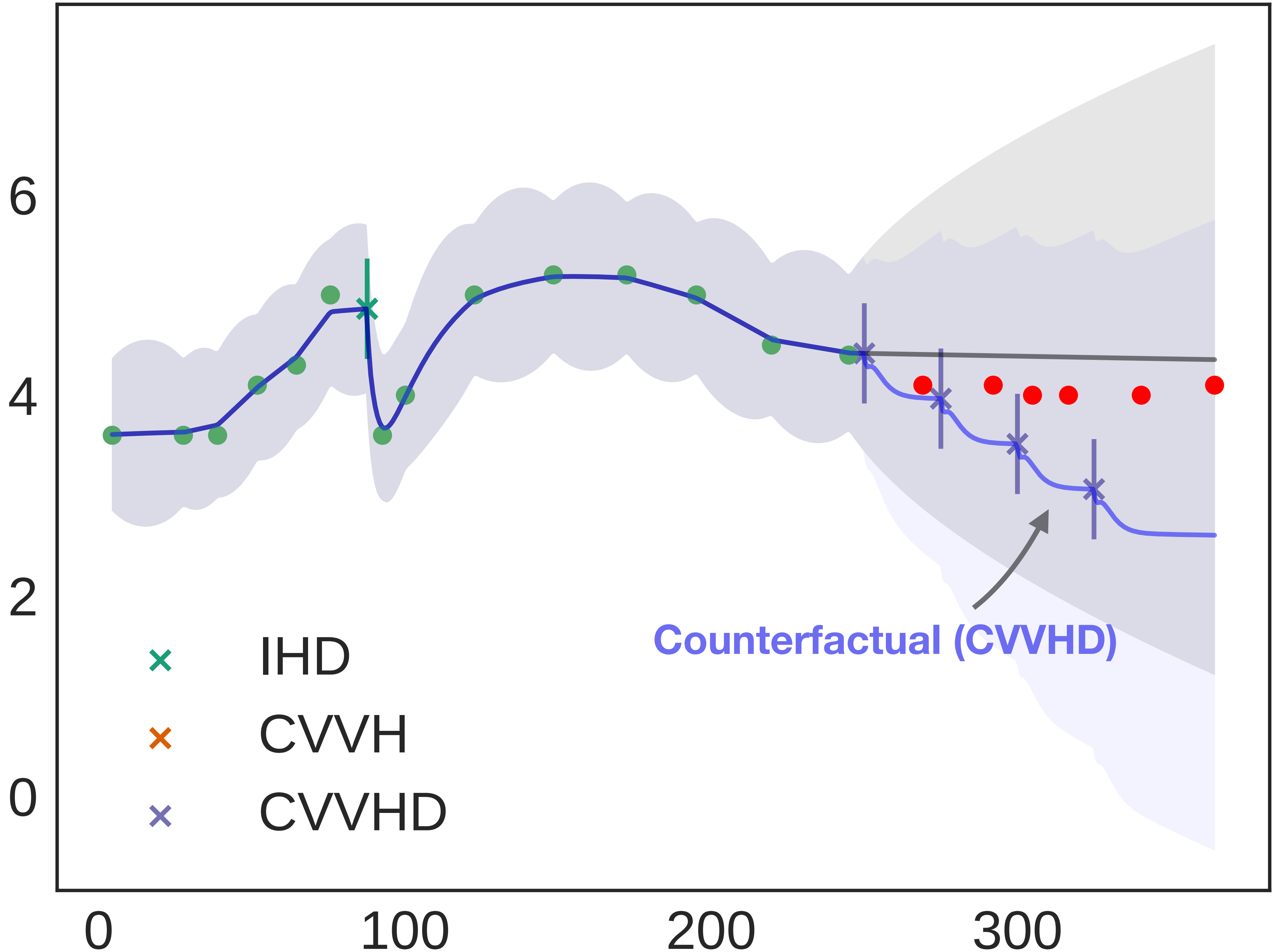
Medical Decision-Support using CGPs

- Dialysis is expensive, but necessary when kidneys fail
- Important questions for decision-making:
  - (1) Will this individual be okay if I remove dialysis?
  - (2) Will this individual benefit from dialysis?
- CGP can help to answer these questions

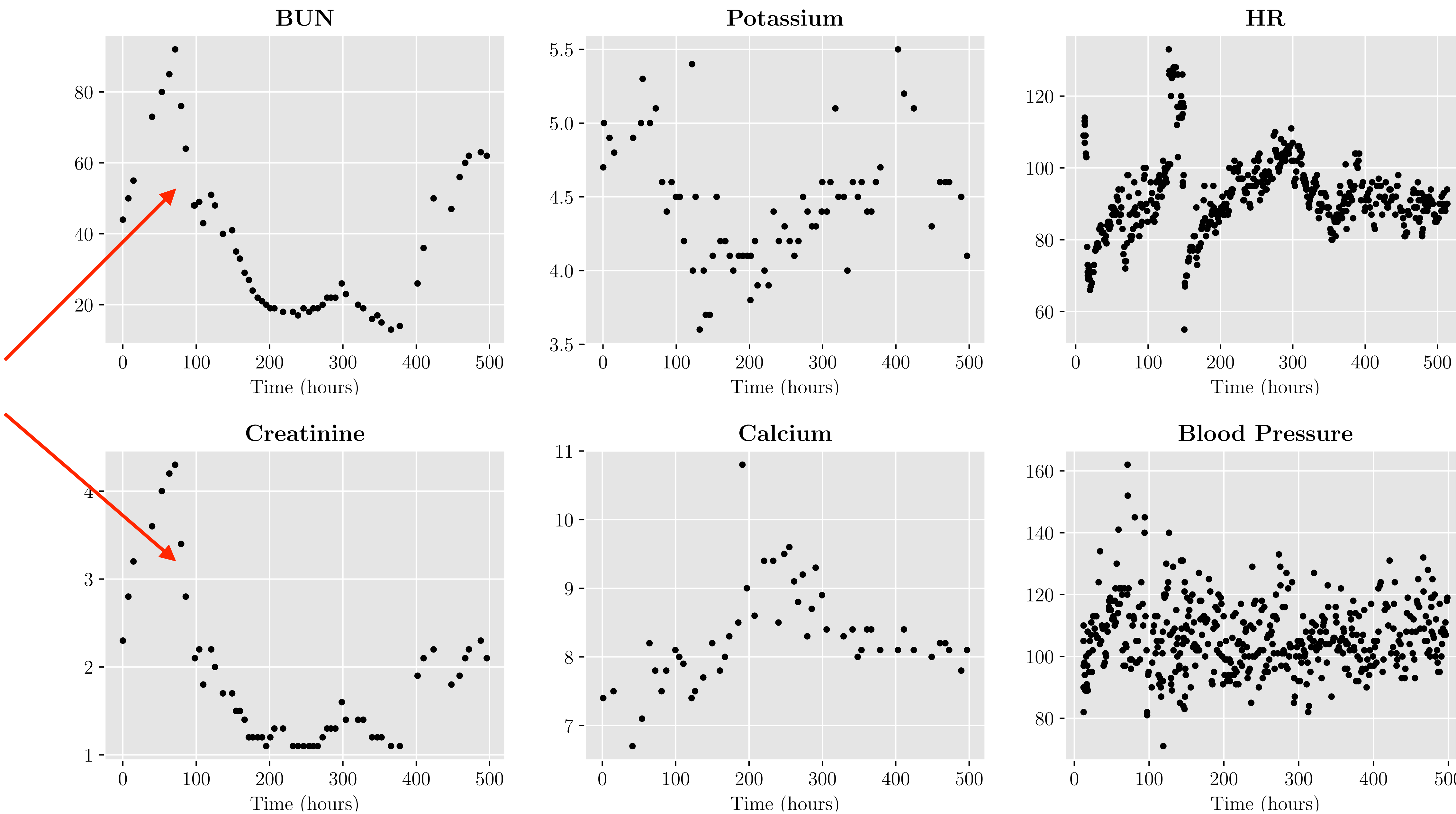
Medical Decision-Support



Medical Decision-Support



A Real ICU Patient with AKI

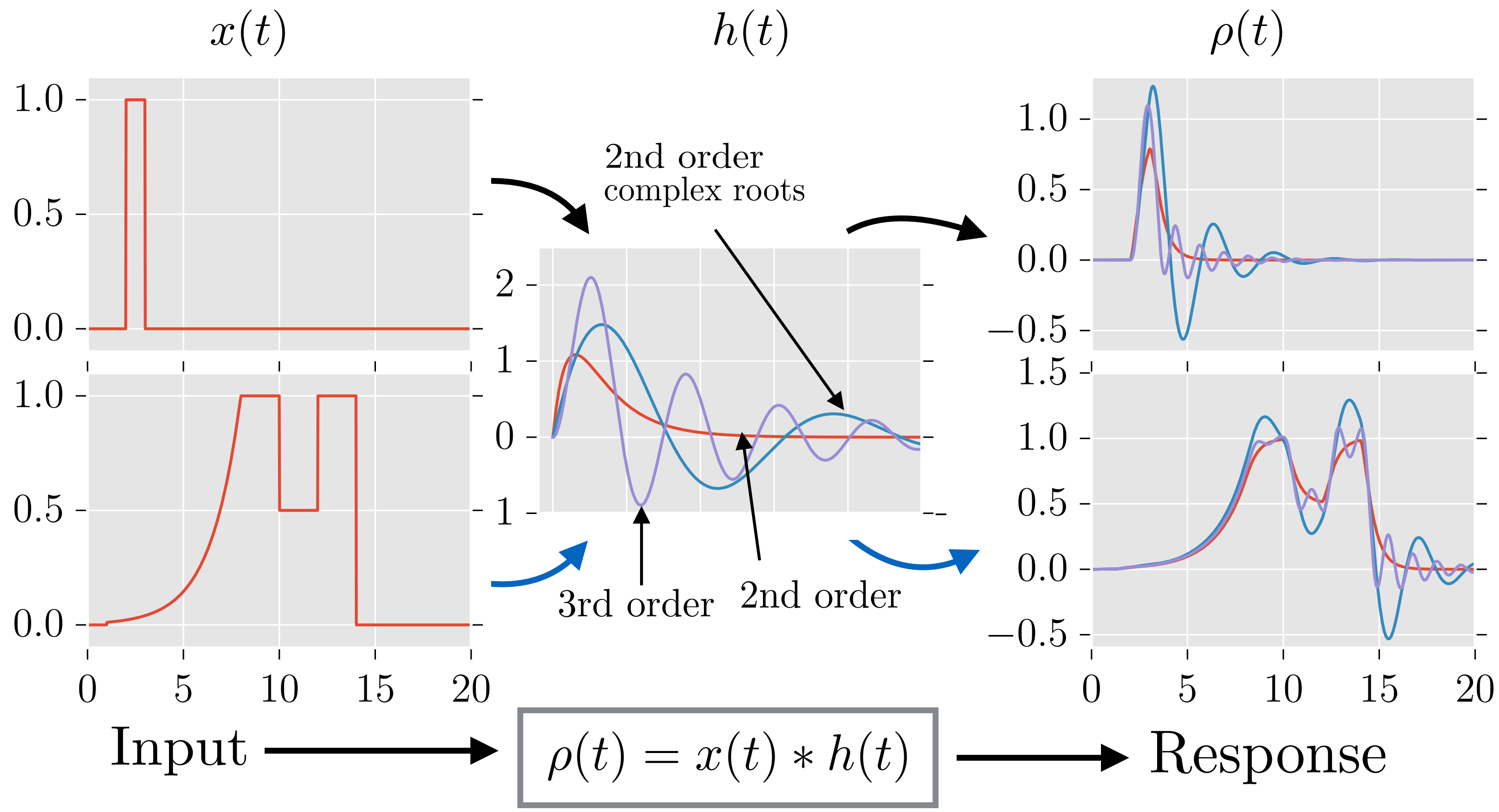


1. Irregularly sampled
2. Unaligned signals
3. Cross correlations

Continuous-time actions, continuous-time multi-variate trajectories

Input $x(t)$ convolved with *impulse-response* $h(t)$ to generate response $\rho(t)$

$$\rho(t) = x(t) * h(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)\mathrm{d}\tau$$



Similar ideas in pharmacokinetics:

Cutler, 1978

Rich et al., 2016

Shargel et al. 2005

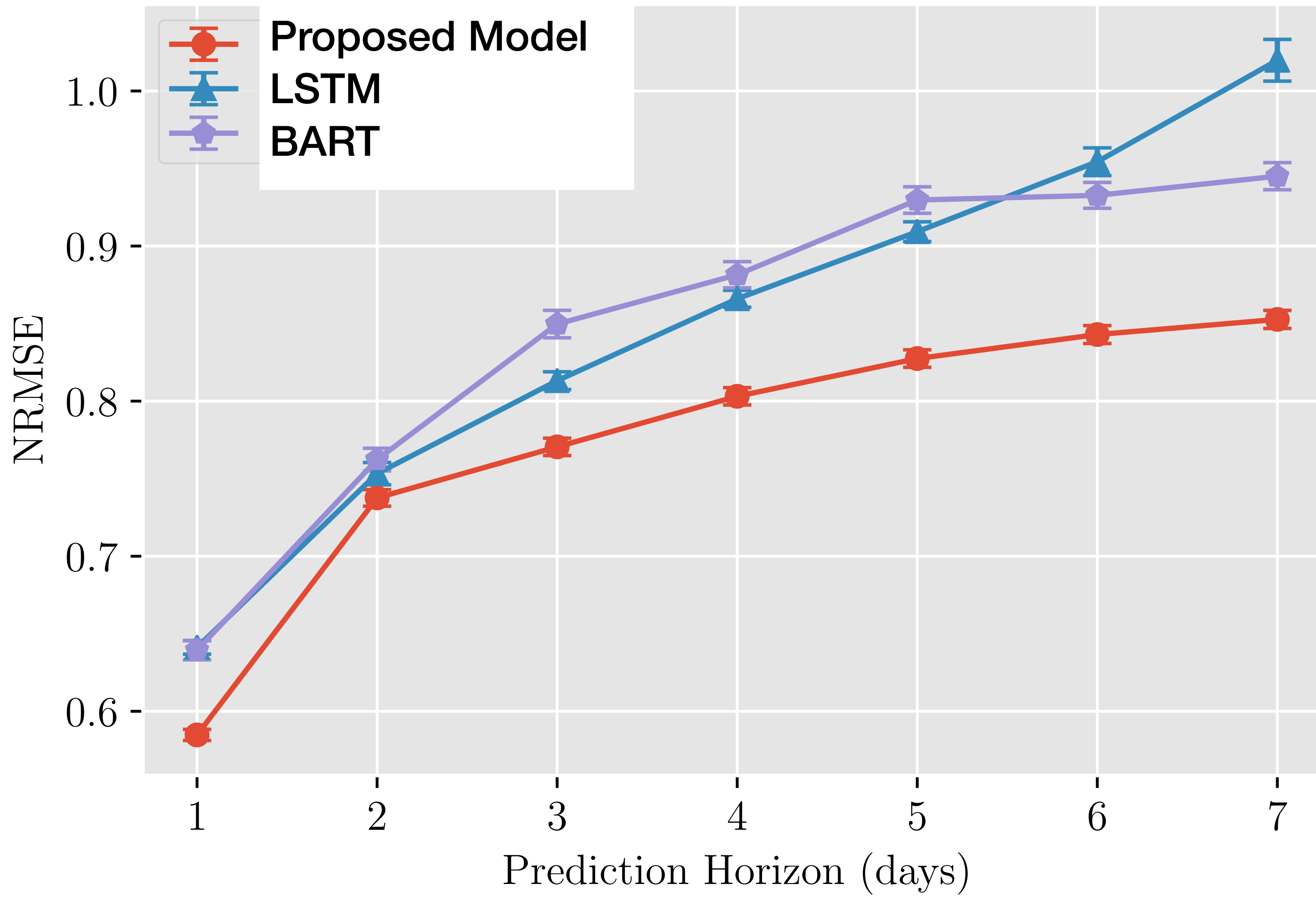
Example: $h(t) = \dfrac{\alpha\beta}{\beta - \alpha}(\mathrm{e}^{-\alpha t} - \mathrm{e}^{-\beta t})1(t \geq 0)$

To allow sharing across signals: $g_d(t) = \psi \underbrace{\rho_0(t)} + (1-\psi) \underbrace{\rho_d(t)}$

$\psi \in [0, 1]$

Soleimani, Subbaswamy, Saria, UAI 2017

Quantitative Results



- Better relative performance at longer prediction horizons

- For horizon 7: on test regions with treatment, 15% than BART and 8% better than LSTM

Soleimani, Subbaswamy, Saria, UAI 2017

Conclusions

- Use counterfactual objectives for training predictive models
- Assumptions are critical for counterfactual models
  - But they are <u>not</u> statistically testable
  - Can we develop formal sensitivity analyses?
- Are the other structural assumptions where CGP's can be learned?
- Counterfactual reasoning is orthogonal to other efforts in interpretability and accountability
  - Counterfactual objective tells us what to fit
  - Interpretable models: how to parameterize for transparency

Key References

- **Potential Outcomes**
    - **Neyman et al., 1923**l. 1990 (English)
    - **Rubin, 1974** **Rubin, 2005**
- **Treatment-Confounder Feedback and G-computation**
    - **Robins 1986**
    - **Robins and Hernan 2009**
- **Counterfactual Reasoning and Reliable Decision Support**
    - **Schulam and Saria, NIPS 2017**
    - **Soleimani, Subbaswamy, Saria, UAI 2017**
    - **Xu, Xu, Saria, MLHC 2016 (JMLR-to appear)**
    - **Dyagilev and Saria, Machine Learning 2015**
    - **Soleimani and Saria, UAI 2017**
    - Saria and Schulam, NIPS Tutorial 2016

Thank you!
ssaria@cs.jhu.edu
www.suchisaria.com
@suchisaria

All references throughout the slides are active links and clickable.
For errors and edits, please contact: ssaria@cs.jhu.edu Thanks!