Clinical Trials and Personalized Medicine

Michael R. Kosorok

University of North Carolina at Chapel Hill

What is Personalized Medicine?

The idea is that a person's genetics (biomarkers) can tell us how to medically treat that individual.

Clearly, the idea extends to epigenetics and/or prognostic data.

Most often, one biomarker is used to identify which patients will benefit from a certain treatment.

An Interesting Event at Duke

In the *New York Times*, July 20, 2010:

Duke Suspends Researcher and Halts Cancer Studies

What happened?

The researcher was Dr. Anil Potti, and the three clinical studies were clinical trials evaluating personalized medicine (individualized treatments) developed by Dr. Potti and others.

Going Back in Time

The individualized treatments studied in the trials were developed in a 2006 *NEJM* publication by Dr. Potti and a 2006 *Nature Medicine* publication of Drs. Potti and Nevins.

Biostatisticians, Drs. Baggerly and Coombes of MD Anderson, used "forensic biostatistics" to identify 5 serious errors (data and statistical) that could potentially lead to patient harm.

The Baggerly and Coombes *Annals of Applied Statistics* paper was not published until December 2009, but preprints were available which led to shutting down two clinical trials at Duke on October 6, 2009 (October 9, 2009, *Cancer Letter*).

In February 2010, Duke resumed the clinical trials (now three) anyway.

In the July 16, 2010, *Cancer Letter*, evidence that Dr. Potti had falsely claimed in his CV that he was a Rhodes Scholar came out.

On Monday, July 19, 2010, a letter from 31 biostatisticians was sent to NCI Director Dr. Harold Varmus and simultaneously to Duke articulating the dangers in the trials from the bad science.

On Tuesday, July 20, 2010, Duke halted the three trials and the American Cancer Center halted funding to Dr. Potti: now a real independent investigation is under way.

The main points in the Varmus Letter were that the Baggerly and Coombes (2009) criticisms needed to be addressed, including establishing reproducibility of results, data validity, and validity of statistical methodology.

The personalized medicine industry is high pressure, cut throat, and often insufficiently guided by statistical principles.

Statistical Methods for Cancer Clinical Trials

Duke, NC State and UNC have obtained a P01 grant from NCI: “Statistical Methods for Cancer Clinical Trials.”

The grant is lead by three PIs: Michael Kosorok (UNC), Marie Davidian (NC State) and Stephen George (Duke).

The P01 consists of 5 projects and 3 cores, 2.5 of the projects are in personalized medicine discovery and evaluation.

The basic idea is for the three universities to work together to make significant advances in clinical trial methodology leading to changes in implementation.

There are over 40 investigators and 10 students working on this project with expertise in:

- Statistics and Biostatistics
- Medicine
- Computer Science
- Biology
- Pharmacology
- Health Services

There are five projects:

1. Innovative Clinical Trial Design and Analysis. Leader: Jianwen Cai.
2. Methods for Missing and Auxiliary Data in Clinical Trials. Leader: Marie Davidian.
3. Methods for Post Marketing Surveillance and Comparative Effectiveness Research. Leader: Joseph Ibrahim.
4. Methods for Pharmacogenomics and Individualized Therapy Trials. Leader: Danyu Lin.
5. Methods for Discovery and Analysis of Dynamic Treatment Regimes. Leader: Anastasios (Butch) Tsiatis.

Supported by three cores:

A Administrative. Director: Michael Kosorok.

B Data Compilation. Director: Stephen George.

C Computational Resource and Dissemination. Director: Marie Davidian.

Overall leadership provided by:

- Executive Committee. Consisting of the Principal Investigators: Michael Kosorok (contact PI), Marie Davidian and Stephen George.
- Steering Committee. Consisting of the Executive Committee plus the co-principal investigators (Joseph Ibrahim, Anastasios Tsiatis and Sin-Ho Jung) and remaining project leaders (Jianwen Cai and Danyu Lin).

As mentioned before, 2.5 of the projects focus on personalized medicine:

- Part of Project 3 focuses on comparative effectiveness research which, through causal inference methodology, can be used to generate hypothesized personalized treatment regimes.
- Project 4 focuses on several pre-clinical personalized medicine areas in addition to the single-decision personalized medicine trial set-up.
- Project 5 focuses on dynamic treatment regimens, a complicated but very promising area of personalized medicine.

I will focus hereafter on personalized medicine methodology, including statistical challenges.

Personalized Medicine Challenges

- Establish a principled approach for discovery and evaluation.
- Develop statistically efficient clinical trial designs and analysis methods for discovering personalized medicine regimes, of both the single-decision and multi-decision (dynamic) type.
- Develop efficient confirmatory clinical trial methods for evaluation and validation of personalized regimens.
- Find ways of leveraging existing clinical trial and other kinds of data for personalized medicine discovery and evaluation.
- Solve the many associated statistical theory, methodology and computational challenges.
- Get medical scientists to use the new methods.

The Single-Decision Setup

The basic challenge is to estimate the relationship between a clinical outcome R and prognostic variables X and treatment choice A.

Machine learning techniques are very useful for estimating $Q(x, a) = E[R|X = x, A = a]$.

For a patient presenting with prognostic values (e.g., genomics, epigenetics, etc.) $X = x$, the optimal treatment policy is estimated as $\hat{\pi}(x) = \arg\max_a \hat{Q}(x, a)$.

How do we do design, estimation, inference and validation for $\hat{Q}$ and $\hat{\pi}$?

The Multi-Decision Setup

In the multi-decision setup, the clinician has several decision times to decide treatment (e.g., cancer with multiple lines of therapy).

Now one must estimate $\pi_t(x_t)$, for time points $t = 1, \ldots, T$, i.e., "what is the best treatment at time t based on prognostic (and historic) data x_t?"

Now a dynamic version of artificial intelligence, reinforcement learning or Q-learning, must be used to estimate these "dynamic treatment regimes;" part of this involves combining several machine learning steps.

Design, estimation, inference and validation are much more complex.

Non-Small Cell Lung Cancer

In typical regimens for advanced cancer (in breast, lung, and ovarian) patients utilize

- a single agent
- in combination with a platinum-based compound
- in multiple stages (lines) of treatment.

In non-small cell lung cancer (NSCLC), 2–3 lines of treatment increases survival.

Can we improve survival by personalizing the treatment at each decision point (at the beginning of a treatment line) based on prognostic data?

Colorectal Cancer

Another important example of an advanced cancer is metastatic colorectal cancer:

- Over half the patients are older than 65.
- The older age group is typically precluded from participating in clinical trials.
- There are many approved therapies but very little knowledge about which treatment is best for which patients.

Can we improve clinical outcome by personalizing treatment selection:

- to either FOLFOX or FOLFIRI?
- in combination with either bevacizumab (BEV) or cetuximab (CET)?

Cystic Fibrosis

A major challenge in patients with cystic fibrosis (CF) is lung infections caused by *Pseudomonas aeruginosa* (Pa):

- Young CF patients acquire Pa off and on, but eventually the infection does not clear up.
- After several years of chronic infection, Pa infections can transform to a severe mucoid variant, leading to death or lung transplant.
- Delaying the onset of the mucoid variant is a primary goal in CF care.

Is it possible to improve on existing treatment methods by using a personalized treatment rule—to be applied at each time a Pa infection is detected—based on prognostic data?

Weighted SVM for the Single-Decision Setup

For each subject, participants are randomized to the treatment options:

1. An individualized treatment rule (ITR) π is a deterministic decision rule from $\mathcal{X}$ into the treatment space $\mathcal{A}$.

2. Let P^{π} denote the distribution of (X, A, R) where π is used to assign treatments (the domain of π):

$$V(\pi) = E^{\pi}(R) = EQ(X, \pi(X)).$$

3. Optimal ITR

$$\pi_0 \in \arg\max_{\pi} V(\pi)$$

.

Value function:

$$\begin{aligned} V(\pi) &= E^{\pi}(R) \\ &= \int R\pi P^{\pi} \\ &= E\left[\frac{1_{A=\pi(X)}}{p(A|X)}R\right]. \end{aligned} \tag{1}$$

Goal: Estimate π_0 that maximizes (1).

Challenges for finding optimal individualized treatment rules:

- Difficult to compute via the usual approach (Qian and Murphy, 2010) where

$$\pi_0(X) \in \arg\max_{a \in \mathcal{A}} Q_0(X, a).$$

- High dimensional pretreatment variables with only a small subset needed by the individualized treatment rule.

Equivalence to a Learning Problem

Conditional on the reward R,

$$\begin{aligned} V(\pi) &= E\left[\frac{1_{A=\pi(X)}}{p(A|X)}R\right] \\ &= E\left[\frac{R}{p(A|X)}E(1_{A=\pi(X)}|R)\right] \\ &= E\left[\frac{R}{p(A|X)}P(A=\pi(X)|R)\right] \\ &\propto E\left[RP(A=\pi(X)|R)\right]. \end{aligned}$$

Thus maximizing $V(\pi) \Leftrightarrow$ maximizing $E\left[RP(A=\pi(X)|R)\right]$

- Recall that the optimal treatment is $\pi_0 = \arg\max_\pi V(\pi)$, and

$$V(\pi_0) = E\left[RP(A = \pi_0|R)\right]$$

.

- Consider the loss function

$$\begin{aligned} L(\pi) &= E\left[RP(A \neq \pi(X)|R)\right] \\ &= E\left[RE(1_{A\neq\pi(X)}|R)\right], \end{aligned} \tag{2}$$

which will be minimized at π_0, and

$$L^*(\pi) = \inf\{L(\pi) : \pi : X \to \{-1, 1\}\text{measurable}\}$$

We need to approximate the above loss function:

- Goal:

  Choose $f : \mathcal{X} \to \mathbb{R}$ from a class of functions $\mathcal{F}$, and the decision $\pi(X)$ based on the sign of $f(X)$ minimize (2).

- An estimation procedure based on minimization of a convex surrogate for the loss,

$$L_\phi(f) = E(R\phi(Af(X))) = E(RE(\phi(Af(X))|R)).$$

- Hinge loss: $\phi(\alpha) = \max\{1 - \alpha, 0\}$ (much easier computationally).

- Theorem: This hinge loss approach is equivalent to the targeted loss and, moreover, is Fisher consistent and gives the Bayes optimal decision rule.

Simulation Study

- $n = 100$ for training data, $n = 50$ for testing data;
- $X = (X_1, \ldots, X_{20}) \sim U[-1, 1]^{20}$;
- A is generated from $\{-1, 1\}$ independently of X;
- The response $R \sim N(Q_0, 1)$, where
  $Q_0 = 1 + 2X_1 + X_2 + 0.5 * X_3 + T_0(X, A)$.

Optimal Value Comparison

Table 1: Binary Response

Treatment Effect T_0	Method		
	Proposed	OLS	Probit
$X_1 A$	-0.2160(0.0018)	-0.1878(0.0032)	-0.2192(0.0017)
$0.442(1 - X_1 - X_2)A$	-0.0349(0.0000)	-0.2805(0.0120)	-0.0367(0.0000)
$X_2 sin(X_1)A$	-0.2295(0.0027)	-0.2248(0.0027)	-0.2291(0.0025)
$(1 - X_1 + sin(X_3)sin(X_4) + X_5 + 0.6X_6)A$	-0.1316(0.0030)	-0.2809 (0.0207)	-0.1333(0.0033)
$(sin(X_1 + 0.5X_2 - 0.2X_3) - X_4^2)A$	-0.2680(0.0139)	-0.3699(0.0152)	-0.9215(0.0187)
$(1 - X_1 + sin(X_3).*cos(X_4) + exp(X_5) + 0.6 * X_6 - (X_7 + X_8).^2)A$	-0.1908(0.0099)	-0.3994(0.0351)	-0.1892 (0.0092)

Dynamic Treatment Regimes

In contrast with classic adaptive designs, where the clinical trial adapts, dynamic treatment regime designs (or “adaptive treatment strategies”) (Murphy, 2005) involve adaptation of the treatment to the patient.

Dynamic treatment regime designs are able to provide information

- not only on the best treatment choices from the beginning
- but also on treatment choices that maximize outcomes at each new decision time, taking into account long-term affects.

Dynamic treatment regimes

- are a new paradigm for treatment and long term management of chronic disease and drug and alcohol dependency, and
- have been incorporated into new trial designs such as sequential multiple assignment randomized trials (SMART) (See Murphy, 2005, *Statistics in Medicine*, and 2007, *Drug and Alcohol Dependence*).
- have more recently been the basis for trial designs for treating cancer, an example of a disease with a short therapeutic time window (Zhao, Zeng and Kosorok, 2009; Zhao, Zeng, Socinski and Kosorok, 2011).

Clinical Reinforcement Trials

In the Zhao et al. (2009 and 2011) papers, we propose a variation/adaptation of the SMART concept, “clinical reinforcement trials,” wherein:

- Each patient is randomized at each decision time to a possibly continuous range of treatment possibilities (drug, dose, timing, etc.).
- At the end of the first stage, reinforcement learning is used to estimate optimal treatment per prognostic values at each decision time.
- It is not necessary that any single patient receive the optimal therapy.
- The first stage is followed by a second, confirmatory stage.

These new designs have two special features:

- First, without relying on pre-specified mathematical models, reinforcement learning based on Q-learning and Q-functions (from artificial intelligence)
  - carries out treatment selection sequentially
  - with time-dependent outcomes
  - to determine which next treatment is best for which patients at each decision time and considering all future possibilities.
- Second, the proposed approach improves longer-term outcomes by considering delayed effects.

Clinical reinforcement trials can extract the optimal treatment regimen while taking into account a drug's efficacy and toxicity simultaneously.

Challenges may arise due to the complexity of the true Q-function, including:

- the non-smooth operation of maximizing over treatment,
- the high-dimension of the state and action variables X and A, or
- the continuity of the action variable.

We utilize two recent nonparametric techniques from machine learning:

- support vector regression (SVR) (Vapnik, et al., 1997) and
- extremely randomized trees (ERT) (Ernst, et al., 2005; Geurts, et al., 2006).

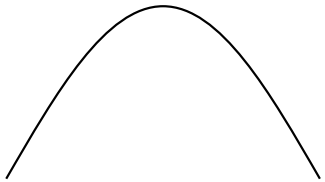
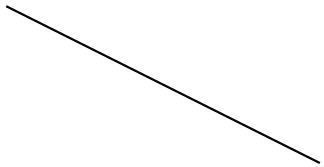
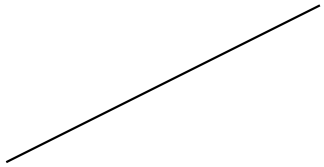
Example 1: NSCLC (ZZSK)

The clinical setting:

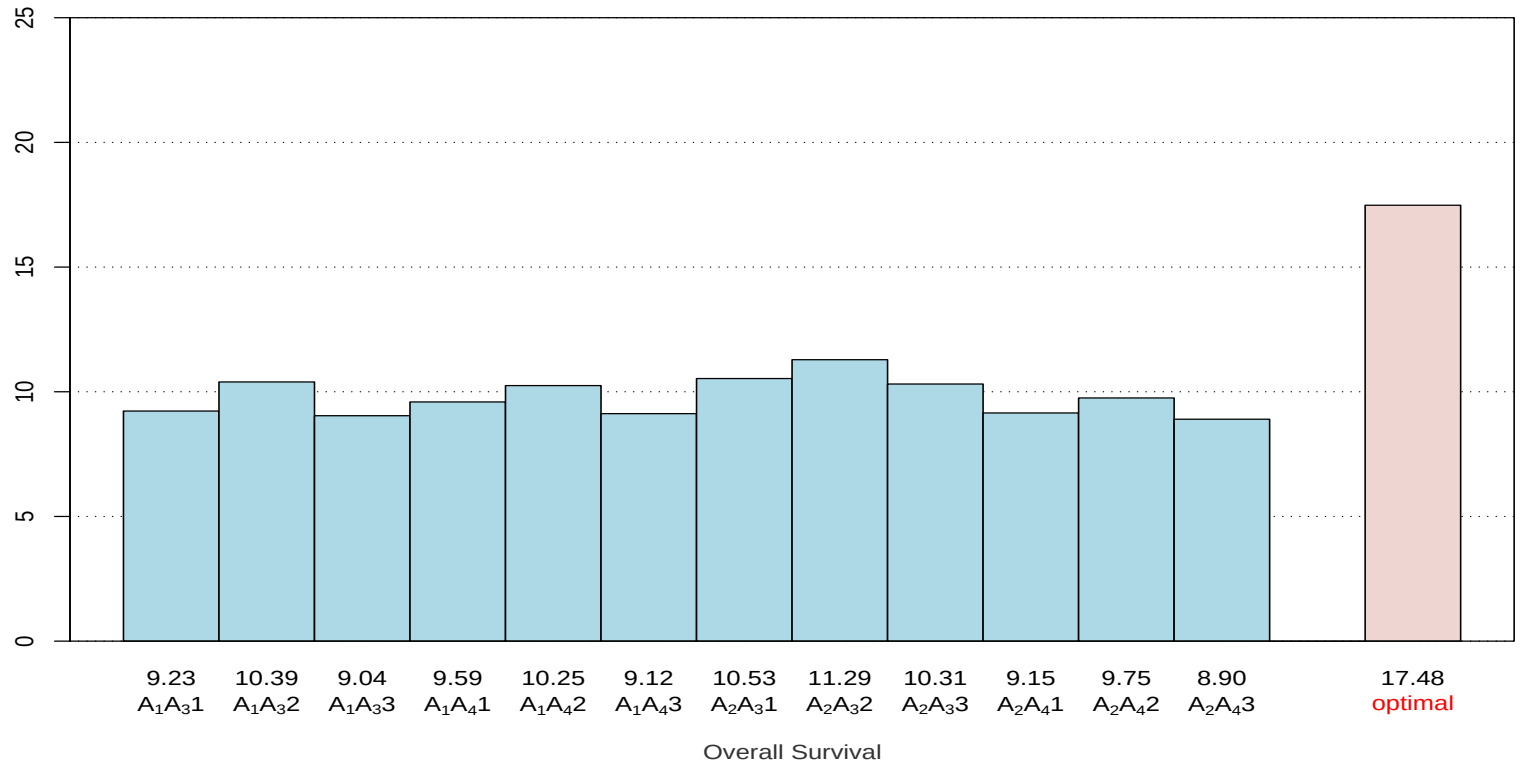
- There are two to three lines of therapy, but very few utilize three, and we will focus on two here.
- We need to make decisions at two treatment times: (1) at the beginning of the first line and (2) at the end of the first line.
- For time (1), we need to decide which of several agent options is best: we will only consider two options in the simulation.
- For time (2), we need to decide when to start the second line (out of three choices for simplicity) and which of two agents to assign.
- The reward function is overall survival which is right-censored.

Parameters of the clinical reinforcement trial:

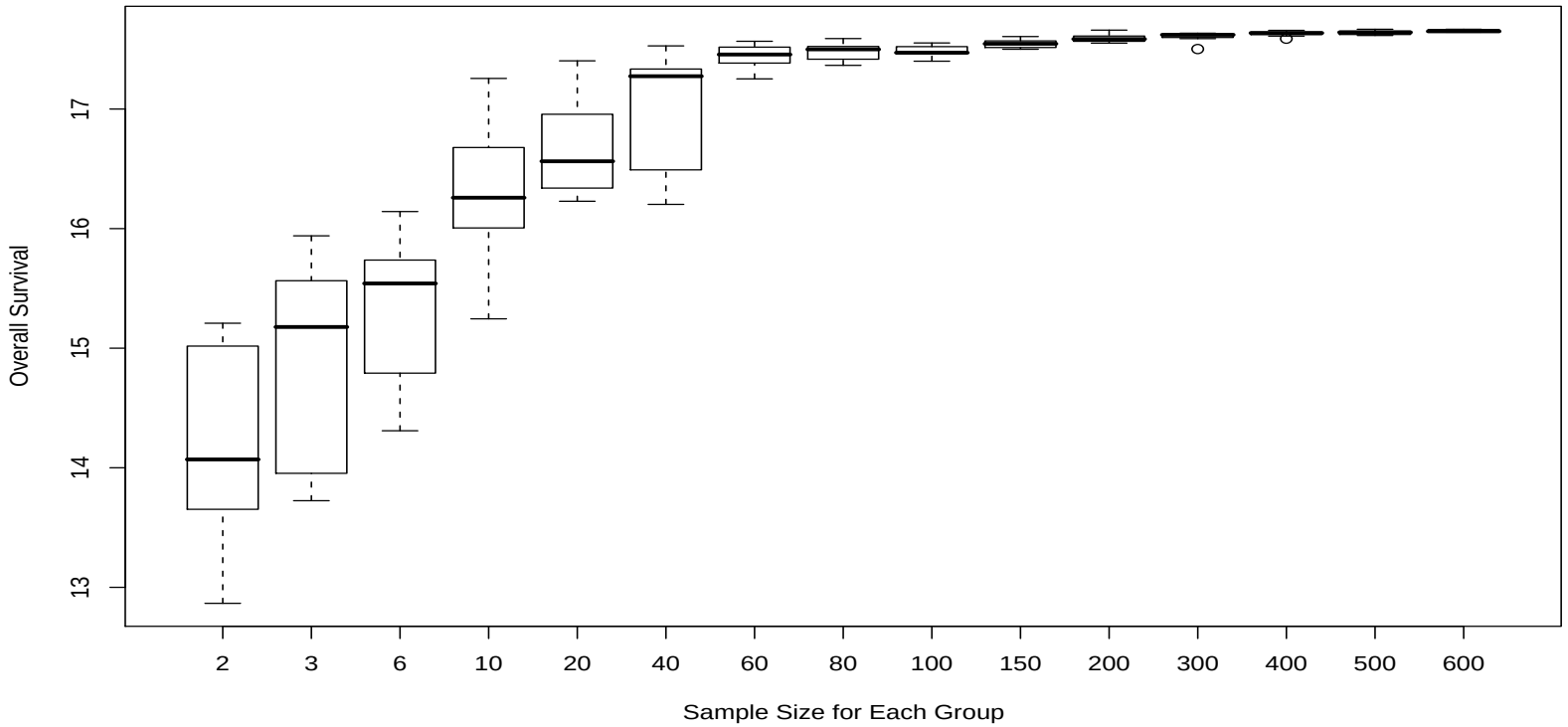
- 400 patients total: 100 each of 4 kinds of patients with different prognostic and treatment response relationships.
- The next table gives the general features of the four kinds of patients.
- A confirmatory phase III trial is conducted with 1300 patients, 100 assigned to the optimal regimen and 100 to each of the 12 possible fixed treatments.
- Reward is overall survival.
- Several levels of right censoring are considered.

Group	State Variables Status		Timing (h)	Optimal Regimen
1	$W_1 \sim N(0.25, \sigma^2)$ $M_1 \sim N(0.75, \sigma^2)$	$W_1 \downarrow M_1 \uparrow$		$A_1 A_3 2$
2	$W_1 \sim N(0.75, \sigma^2)$ $M_1 \sim N(0.75, \sigma^2)$	$W_1 \uparrow M_1 \uparrow$		$A_1 A_4 1$
3	$W_1 \sim N(0.25, \sigma^2)$ $M_1 \sim N(0.25, \sigma^2)$	$W_1 \downarrow M_1 \downarrow$		$A_2 A_3 3$
4	$W_1 \sim N(0.75, \sigma^2)$ $M_1 \sim N(0.25, \sigma^2)$	$W_1 \uparrow M_1 \downarrow$		$A_2 A_4 2$

Performance of optimal personalized regimen versus the 12 fixed combinations under no censoring.



Predicted optimal treatment survival probability versus trial sample size per group.



Example 2: Cystic Fibrosis

The clinical setting:

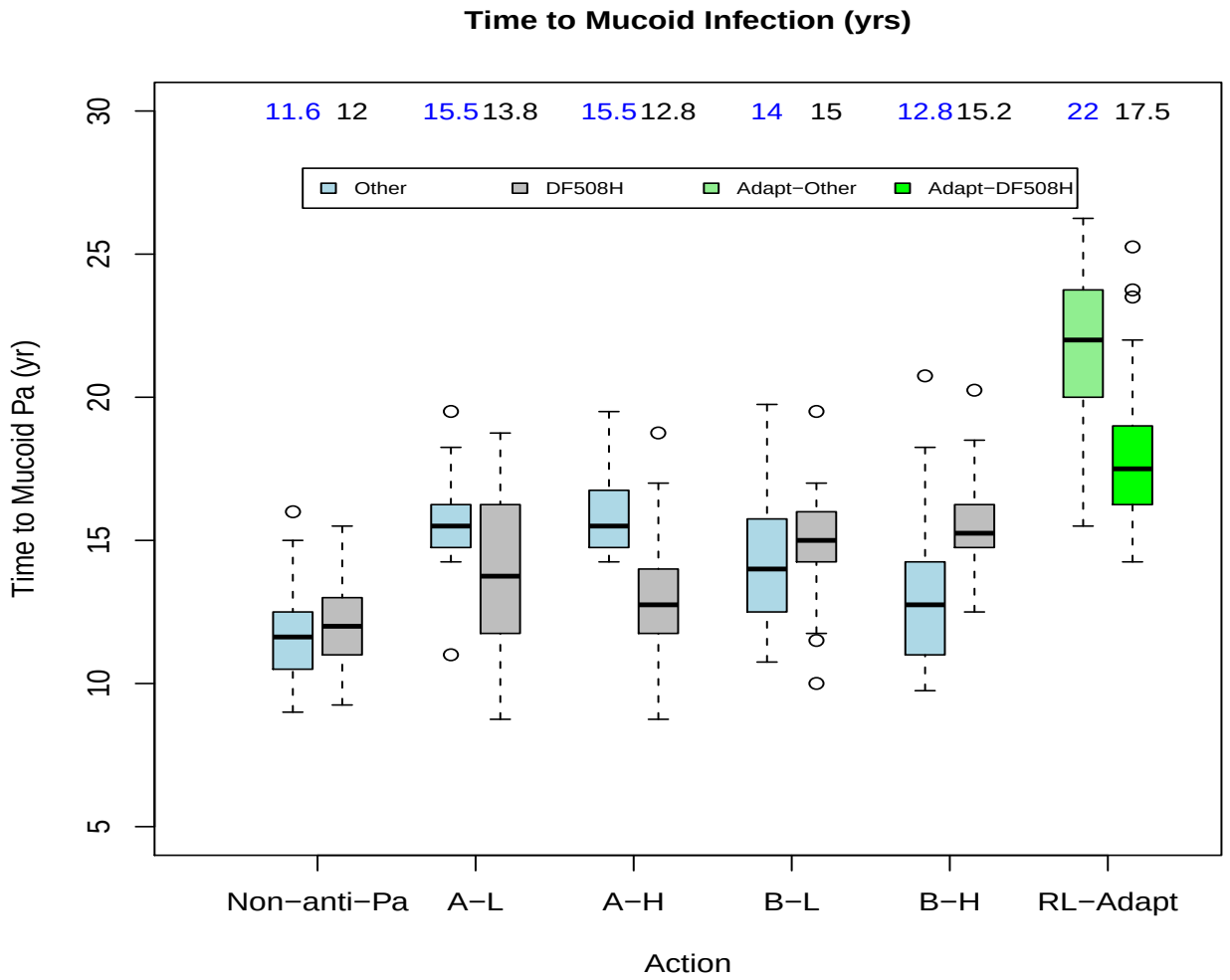
- Cystic fibrosis (CF) is a homozygous recessive genetic disorder that affects predominantly the lung and digestive tract.
- Children with CF are at higher risk for lung infections and pneumonia than normal children.
- The most serious lung pathogen is Pseudomonas aeruginosa (Pa) which does not usually infect non-CF children but can have very serious consequences in CF children.

- Pa infections are usually intermittent at first but eventually become chronic, leading to mucoid Pa infection, after which lung function decline is precipitous.
- There is a belief that if Pa infections can be eradicated rapidly, then the mucoid stage can be delayed significantly.
- Our goal is to find the best choice of treatment each time a patient is infected with CF to yield the longest mucoid-free survival.

Parameters of the clinical reinforcement trial:

- We need to recruit patients with age ranging from 0–20 years old and follow for about 2 years.
- For each episode of Pa infection, we will randomize to one of 5 treatments: placebo, AL, AH, BL and BH.
- Which treatments are acceptable will depend on patient prognostic data, including age.
- After trial completion, we will use Q-learning for an “infinite horizon” to estimate the optimal, personalized treatment choice as a function of prognostic values.
- A phase III randomize trial will then be conducted to verify superiority of the personalized treatment compared to fixed, standard-of-care approaches.

Comparison of time-to-mucoid infection between optimal personalized treatment and fixed treatments when genetics factor is included.



Open Questions

- We need better machine learning tools for predicting conditionally expected survival times under right-censoring.
- We need more evaluation of the properties of the proposed weighted classification.
- We need to come up with more refined technical tools for weak convergence so that inference and sample size calculations can be conducted for the value function and test error.
- We need to allow for varying numbers of decisions times and sizes of time intervals between decision times across individuals since this is a realistic feature in medicine.
- There remains much to do in both theory and applications.