Introduction to Design and Analysis of SMARTs

Michael R. Kosorok

September 2013

Outline

- What are SMARTs and Why Do We Care?
- Necessary Non-Standard Analytical Tools
- Some Illustrative Examples
- Overview of Statistical Issues
- Outcome Weighted Learning
- Multi-Decision Outcome Weighted Learning
- Open Questions
- Preparing Protocols
- Closing Comments

Why Do We Care?

We want to make the best treatment decisions based on available data.

- The single-decision setting:
  - A patient presents with a disease and we need to decide what treatment to give from a list of choices.
  - We want to make the best decision (treatment regimen) based on baseline data.
- The multi-decision setting:
  - We wish to treat a patients for a disease with multiple treatment decisions spread out over time.
  - We want to make the best decision based on historical data at each decision time (dynamic treatment regimen).
  - This could include information on responses to previous treatments.
  - The best decisions take into account delayed effects of treatment.

What are SMARTs?

SMARTs are Sequential Multiple Assignment Clinical Trials. The key features are:

- Randomizations are done at multiple decision time points.
- Choice of "treatments" to randomize can depend on success or failure of previously randomized treatments.
- These trials mimic (but with built-in foresight) what happens in real clinical settings since treatments are sometimes switched if they are not working or are too toxic.
- These trials are challenging to analyze (require non-traditional approaches) but we know how to do it.
- They have been used in mental health and addiction research but less so in other health arenas.

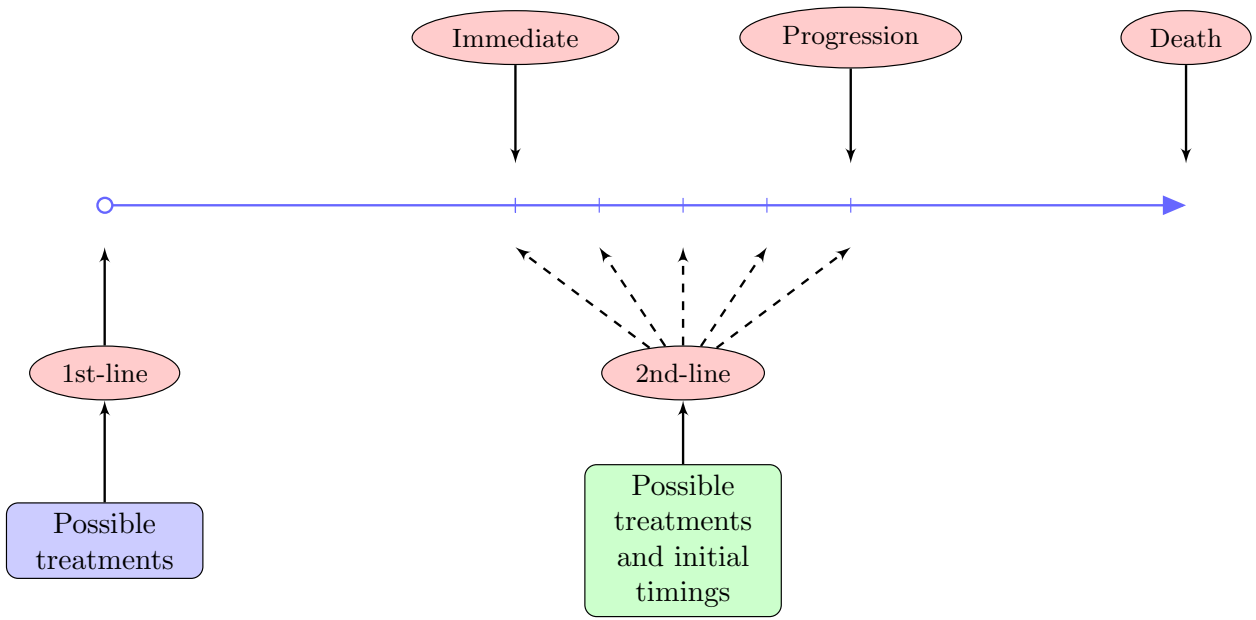
Necessary Non-Standard Analytical Tools

The statistical tools needed for estimation of optimal treatment regimens are non-standard.

- The single-decision setting:
  - Need both regression and maximization.
  - Frequentist and Bayesian methods can be used.
  - Machine learning (which can be either frequentist or Bayesian) are useful here.
  - The choice of frequentist versus Bayesian is not as important as the operational properties of the approach (e.g., consistency).
- The multi-decision setting:
  - Need reinforcement learning, which combines regression and maximization in a dynamic fashion, to account for delayed effects.
  - Q-learning, A-learning and G-estimation are popular choices.
  - Both frequentist and Bayesian approaches are available.

Example 1: Non-Small Cell Lung Cancer

In treating advanced non-small cell lung cancer, patients typically experience two or more lines of treatment.



Problem of Interest

Can we improve survival by personalizing the treatment at each decision point (drug at both and timing at second) based on prognostic data?

Example 1: Non-Small Cell Lung Cancer

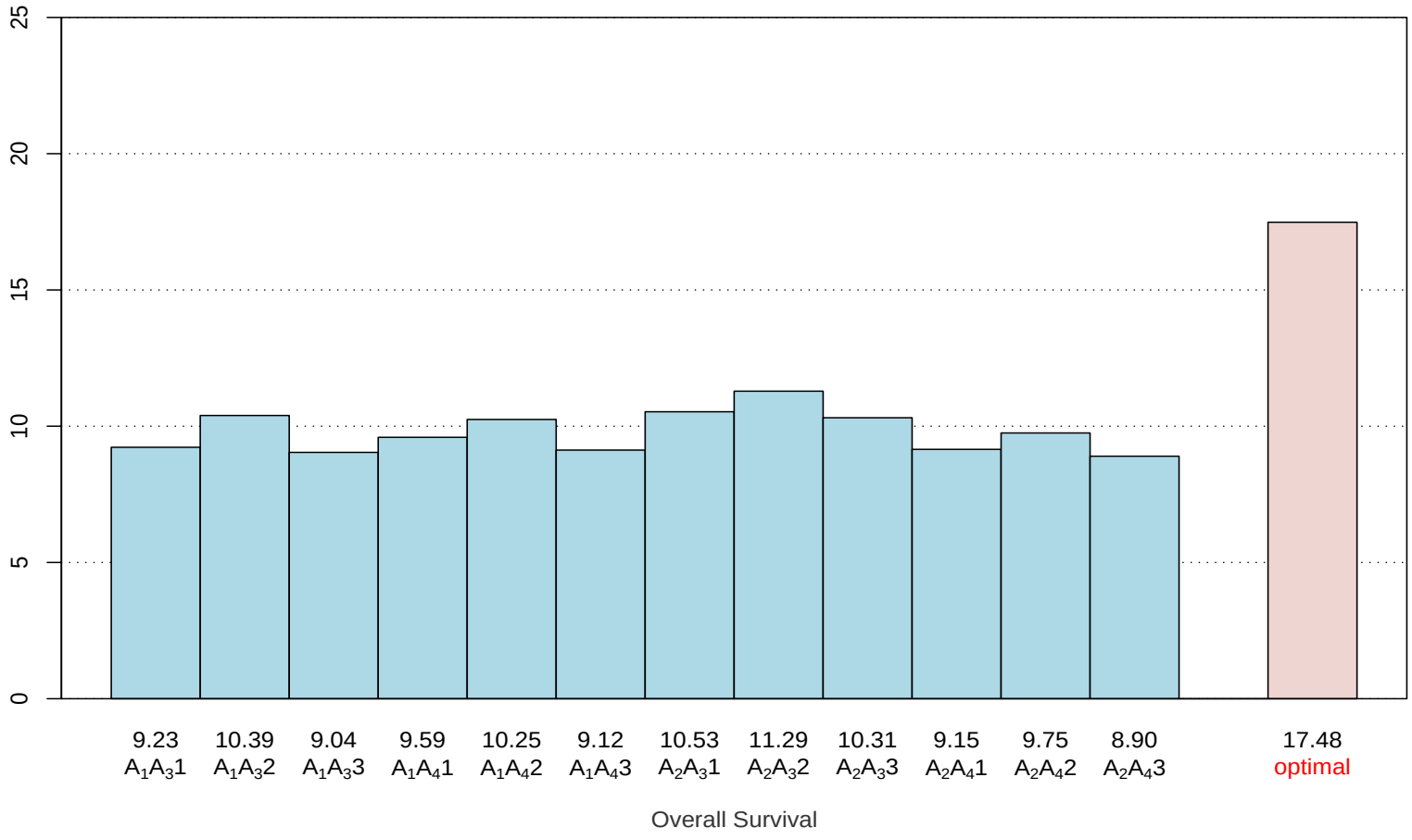
The clinical setting:

- For most patients, there are two lines of therapy.
- Choice of treatment at beginning (1) and end (2) of first line.
- The reward function is overall survival which is right-censored.

Realistic simulated patients (Zhao, et al., 2011):

- Difference equations used to generate patient trajectories for two clinical measures: tumor size and quality of life.
- A SMART trial was simulated (12 different treatment paths).
- Q-learning was used to estimate decision rules.

Example 1: Non-Small Cell Lung Cancer



Example 1: Non-Small Cell Lung Cancer

Some statistical issues:

- Machine learning is very useful for handling
  - nonlinear structure,
  - complex interactions, and
  - large numbers of variables.
- Machine learning tools for censored data are very limited (almost nonexistent) and appropriate extensions are needed.
- Complex treatment decisions (involving multiple drugs and/or timing) are new challenges for statistical learning.

Example 2: Palliative Care Treatment

The clinical setting:

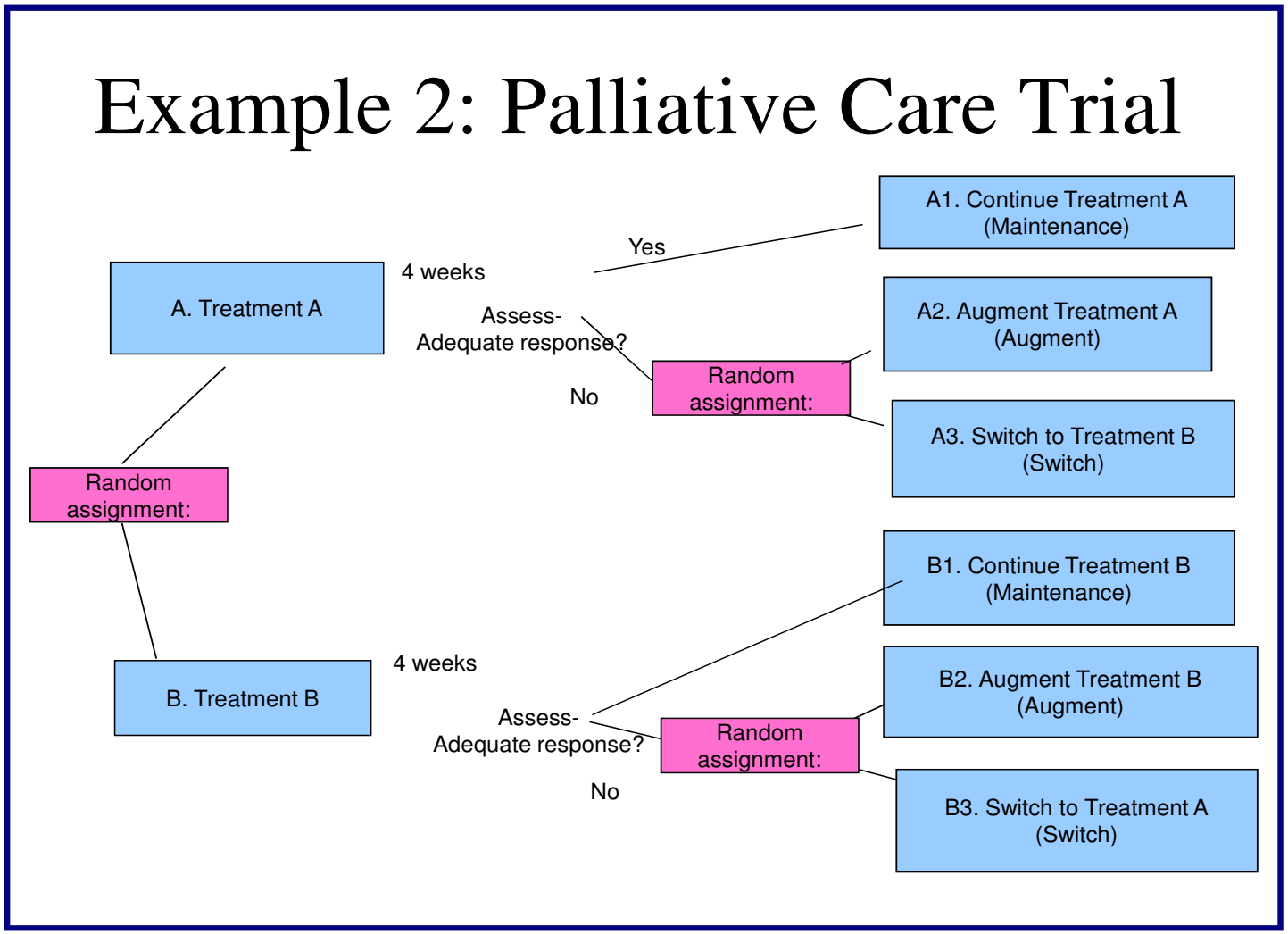
- In palliative care settings, the goal is patient comfort and reduced stress for care givers.
- Suppose we wish to know which of two fundamentally different approaches to palliative care is best initially.
- Suppose we also wish to know what best alternative to give if patients do not respond to initial treatment.

Scientific and statistical issues:

- This formalizes an intent-to-treat analysis where we explicitly assess options for non-responders.
- One advantage of a SMART design is the increased power for assessing treatment strategies since responders can be pooled with non-responders for some comparisons.
- We also have the ability to incorporate tailoring variables.

Example 2: Palliative Care

Potential palliative care trial design:



Example 3: Bronchopulmonary Dysplasia in Infants

The clinical setting:

- Sildenafil is effective for preventing bronchopulmonary dysplasia-associated pulmonary hypertension.
- Crucial to know what dose to use with which patients.
- We designed a Phase II individualized dose finding study.

Scientific and statistical issues:

- The investigators would like the design to be adaptive so that ineffective or harmful doses are discarded early.
- A statistical challenge is that dose is continuous.
- Since the methodology is complicated, how do we frame the proposal to satisfy reviewers and obtain approval?

Example 4: Cystic Fibrosis

The clinical setting:

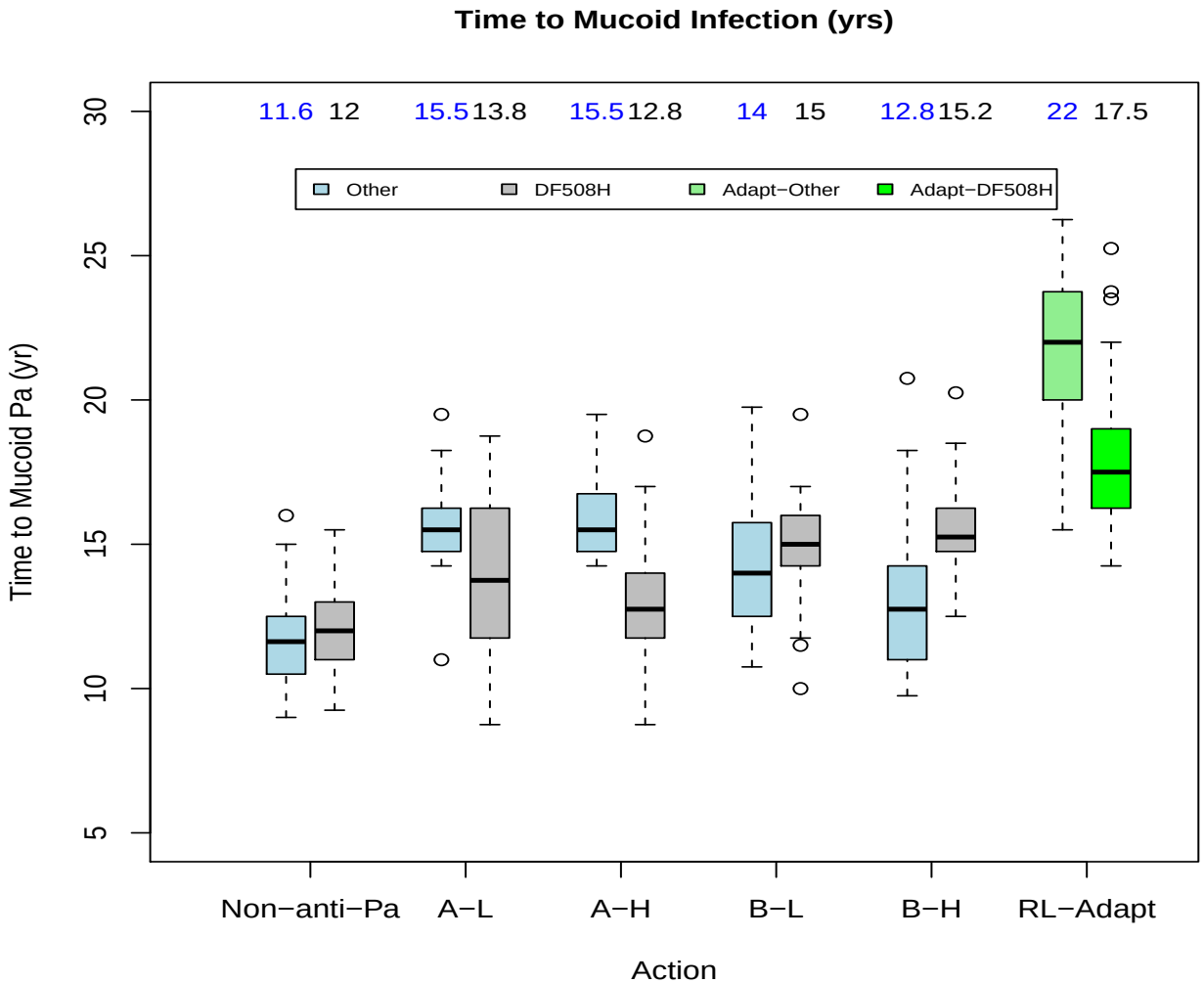
- Cystic fibrosis (CF) is a genetic disease.
- The major pathogen in CF is Pseudomonas aeruginosa (Pa).
- Pa lung infections are at first intermittent but eventually chronic, leading to harmful mucoid Pa by the late teens.
- Our goal is to find the best treatment each time a patient is infected to maximize mucoid-free survival from birth.

Realistic simulated patients and trial (Tang, et al., 2012):

- We recruit patients with ages 0–20 years old and follow for about 2 years for Phase II SMART trial.
- For each episode of Pa infection, we randomize to one of 5 treatments: placebo, AL, AH, BL and BH.
- After SMART trial completion, we use Q-learning to estimate optimal, personalized treatment choice.

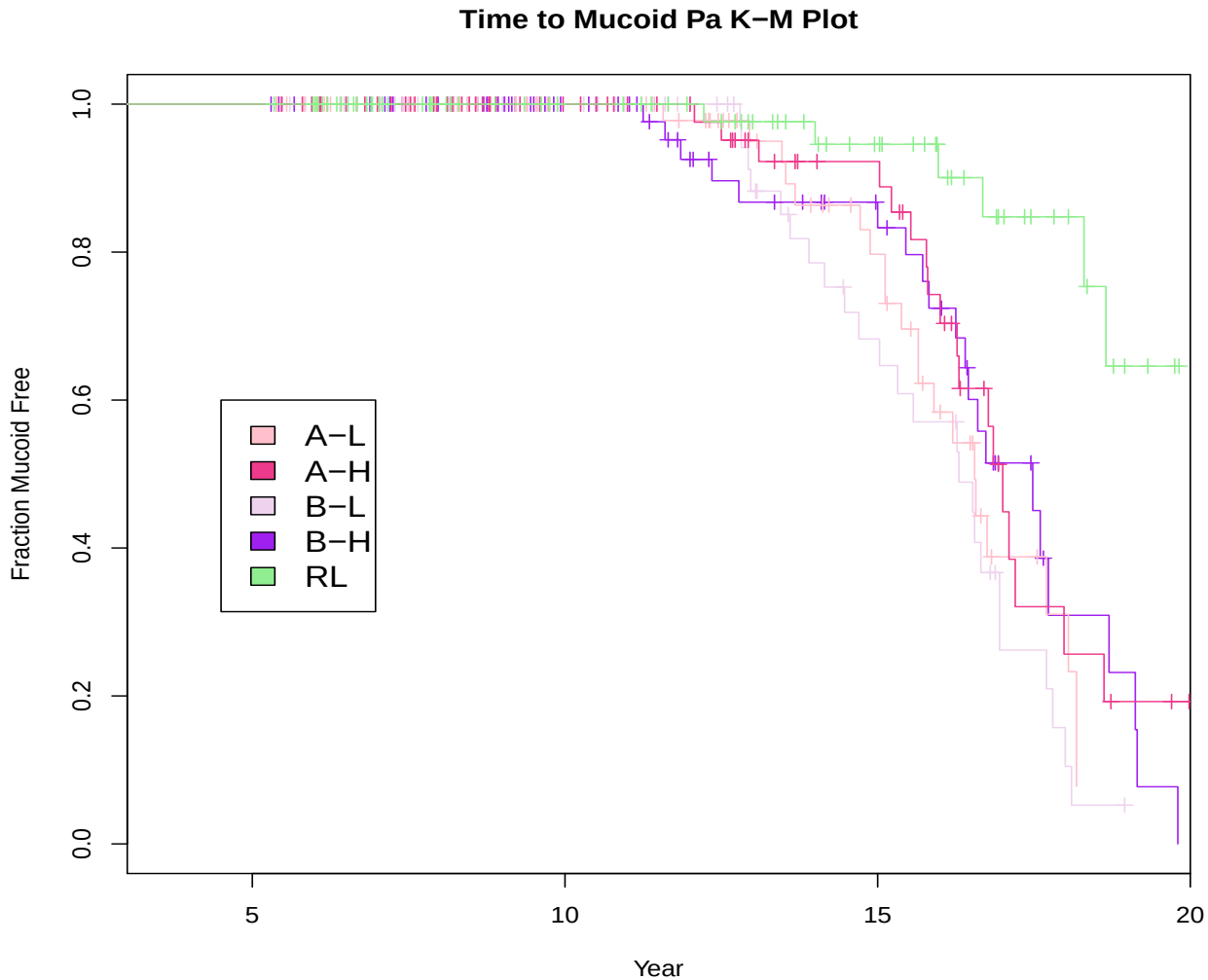
Example 4: Cystic Fibrosis

Comparison of time-to-mucoid infection between optimal personalized treatment and fixed treatments from SMART trial:



Example 4: Cystic Fibrosis

Kaplan-Meier plots from 5 year confirmatory Phase III trial of optimal versus fixed regimens:



Example 4: Cystic Fibrosis

Some scientific and statistical issues:

- Construction of primary clinical outcome (utility) as a composite of several outcomes was highly non-trivial.
- The fact that the disease course is so lengthy raises clinical trial design and regimen estimation challenges.
- The way we addressed this:
  - 2-year SMART Phase II trial with variety of ages.
  - 5-year Confirmatory Phase III trial also with variety of ages.
  - Careful selection of utility to include short time outcomes predictive of mucoid PA as well as mucoid PA.
  - Judicious use of a constant, infinite horizon Q-function.

Example 5: Laser Treatment of Burn Scars

The clinical setting:

- North Carolina Jaycee Burn Center is one of the top burn centers internationally
- An important issue is the treatment of severe scaring from burns
- New laser treatments (pulse dye and CO_2) available but unsure of best practice and comparison to standard-of-care (compression bandage treatment)
- We plan on randomizing patients to sequences of three treatments to determine best ordering
- Very novel SMART design (tranditional within SMART), IRB approved and starting any day.

Overview of Statistical Issues

- Non-standard statistical methods must be used, although weighted linear regression can be used in some settings.
- The best approach is reinforcement learning from artificial intelligence and related methods.
- The most popular form of reinforcement learning in this context is Q-learning.
- Q-learning involves multiple stages of regression (one for each decision time) and allows for many choices at each stage, including
  - linear regression,
  - support vector regression (including non-linear),
  - random forests, and many other approaches.

Overview of Statistical Issues, cont.

- Complexity of Q-learning and associated Q-functions
  - Nonlinearity
  - Complicated interactions
  - High dimensional data
- Complex decision making
  - Drug (treatment) choice
  - Timing of treatment
  - Dose level
- Censoring (time-to-event limited by end-of-study time)
- Clinical trial design challenges
- Focus in this talk on "frequentist" approaches.

Machine Learning

Machine (Statistical) learning consists of data driven tools for regression, classification and for other facets of decision making.

Many approaches originated in computer science (artificial intelligence and machine learning) but have more recently become part of statistical science (statistical learning).

Examples include:

- Support vector machines (SVM)
- Support vector regression (SVR)
- Random forests
- Reinforcement learning
- Q-learning and A-learning

Incorporating Censoring

Basic issue

The basic issue is that in estimating Q-functions where the outcome Y is a failure time, we are interest in a conditional expectation rather than the more standard hazard function in survival analysis (i.e., we can't use Cox regression).

Ad hoc approaches

- Censoring is almost never encountered in computer science based artificial intelligence approaches.
- One could throw out the censored observations.
- Another approach for SVR is to not penalize if the prediction is above the censored observation and only penalize if below: this is better than the above but still has significant bias.

Incorporating Censoring

Progress for single decision setting:

- Successfully developed new random forest approach for censored data, "Recursively Imputed Survival Trees" (Zhu and Kosorok, 2012).
- The above approach is very computationally efficient and avoids inverse weighting.
- Extended support vector regression to survival data using inverse probability of censoring weighting (Goldberg and Kosorok, 2012a).
- The above approach is consistent, with good error rates, and performs well, but the inverse weighting requires additional modeling of censoring.

Incorporating Censoring

Progress for multiple decision setting:

- Ad hoc approach based on decreased penalization for censored observations performed reasonably well in two-stage Q-learning for treating non-small cell lung cancer (Zhao, et al., 2011).
- However, theoretically, the above ad hoc approach can potentially have unbounded bias.
- Successfully developed Q-learning for right censored data using inverse probability of censoring weighting (Goldberg and Kosorok, 2012b).
- The approach is known to be asymptotically unbiased with good error rates and is computationally reasonable.

Outcome Weighted Learning

1. Let X be the vector of tailoring variables (baseline, biomarkers, etc.), A be the choice of treatment given, and Y be the clinical outcome (assuming larger is better for now). The "standard" approach to finding the optimal decision function $d(X)$ is to first estimate the Q-function

$$Q(x, a) = E\left[\, Y \middle|\, X = x, A = a\right]$$

   through regression of Y on (X, A), yielding the estimated decision function $\hat{d}(x)$ which assigns the treatment a which maximizes $\hat{Q}(x, a)$ for the patient's value of x.

2. The value function of $d(X)$ (Qian & Murphy, 2011) is the expected value $\mathcal{V}(d)$ of Y given that the decision rule $d(X)$ is applied to the given population of patients.

Outcome Weighted Learning (OWL)

Optimal Individualized Treatment Rule d^*

- Surprisingly, finding the decision rule $d(X)$ which maximizes $\mathcal{V}(d)$, where $\mathcal{V}^*$ denotes the maximum, can be shown to be equivalent to a weighted classification problem where
  - we classify patients to treatment choice A
  - as a function $d(X)$ of tailoring variables X
  - but weight by clinical outcome Y divided by propensity score $P(A|X)$ (which is known in randomized trials).
- For any rule d, $d(X) = \text{sign}(f(X))$ for some function f.
- These simple observations allows us to utilize machine learning techniques for classification relatively directly.

OWL Results

- Fisher consistent and asymptotically consistent.
- Risk bounds and convergence rates similar to those observed in SVM literature (Tsybakov, 2004).
- Excellent simulation results.
- Promising performance overall (Zhao, et al., 2012a).
- Opens door to application of statistical learning techniques to personalized medicine.
- Successfully applied to the Nefazodone-CBASP clinical trial on chronic depression (Keller et al., 2000).

Multi-Decision Outcome Weighted Learning

The data in the multi-decision setting consists of observations of (X, A, Y) at each of T decision times:

- At each decision time $1 \leq t \leq T$, we replace X_t with H_t which consists of X_t combined with all of the prior history contained in (X_j, A_j), $1 \leq j < t$.
- $H_1 = X_1$ since there is no history prior to $t = 1$.
- *Goal:* Estimate optimal decisions $D = \{d_1, \ldots, d_T\}$, where $d_j = d_j(H_j)$, using these data.
- Not just a direct extension of Q function estimation: need to take into account long term effects.
- A popular approach to obtaining D is through Q-learning (other approaches include G-estimation, A-learning, and Bayesian alternatives using the Bellman equation).

Multi-Decision Outcome Weighted Learning

- Can we extend classification to sequential treatments?
- Target value function directly as in OWL.
- Idea: compute the value functions directly as before and convert the problem to a series of weighted classification problems so that machine learning techniques can be used at each decision time starting at the end and going backwards.
- At each decision time, only patients following the optimal decision for future times are included in the sample.
- We call this Backward Outcome Weighted Learning (BOWL).

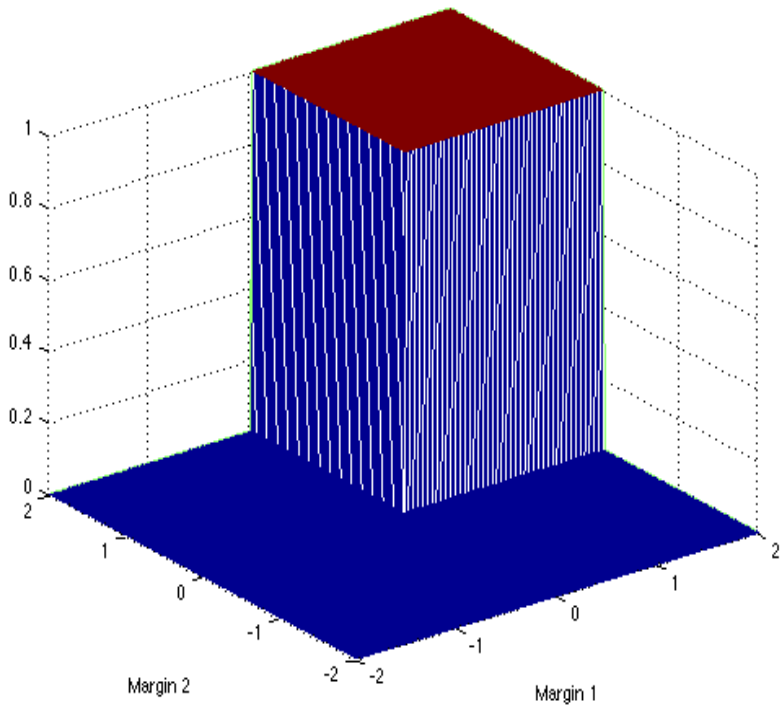
Iterative Outcome Weighted Learning (IOWL)

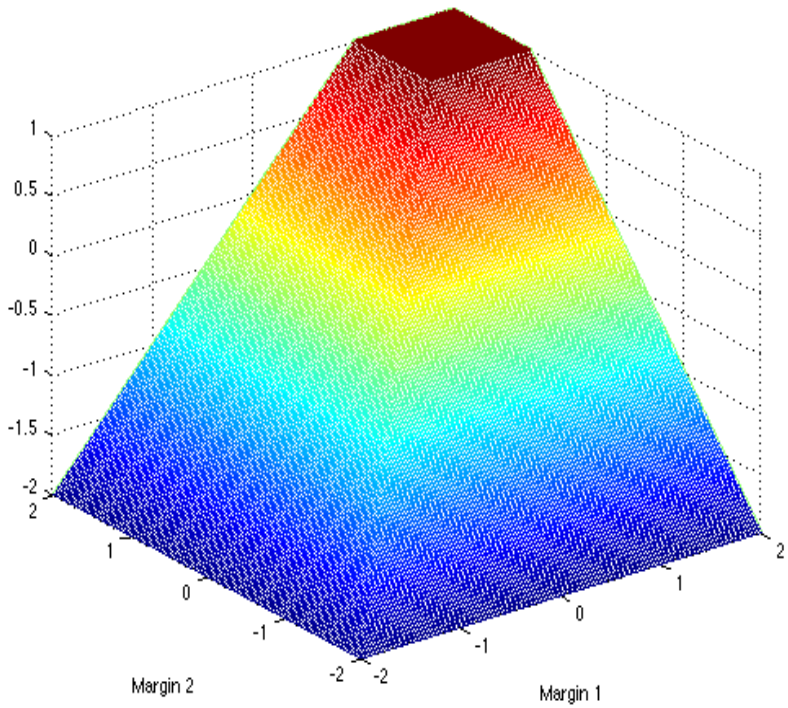
- Concern: only partial data are used in DTR estimation.
- Fix: an iterative procedure for two stage setup:
  1. Update $\hat{d}_2$ based on the subset where subjects take the estimated optimal treatment in stage 1.
  2. Similarly, update $\hat{d}_1$ restricted to the subjects who are assigned with $\hat{d}_2$ obtained in Step 1.
  3. Iteration between 1 and 2 stops upon stabilization of the value functions.
- Includes more subjects that can contribute to stage 1 estimation.
- In theory, both BOWL and IOWL lead to the optimal DTR.
- IOWL works slightly better when implemented.

Simultaneous Outcome Weighted Learning (SOWL)

- Ideally, we want to learn the optimal regimens at all stages simultaneously.
- To do this, we had to create a fundamentally new kind of machine learning technique.
- The basic ingredient is a multi-dimensional hinge (on the next page).
- We can employ quadratic programming via dual problem.
- Easy to generalize to nonlinear decision rules using kernel trick.

SOWL: Two Dimensional Hinge





Properties of BOWL (IOWL) and SOWL

- Fisher consistency.
- Optimal Value Consistency
  - Estimation of optimal DTR within an RKHS.
  - The value of constructed DTR via BOWL/SOWL converges to the best achievable value by any regimens restricted within the selected RKHS.
- Asymptotic consistency.
- Risk bound

$$\mathcal{V}^* - \mathcal{V}(\hat{D}) \leq O(n^{-\gamma}).$$

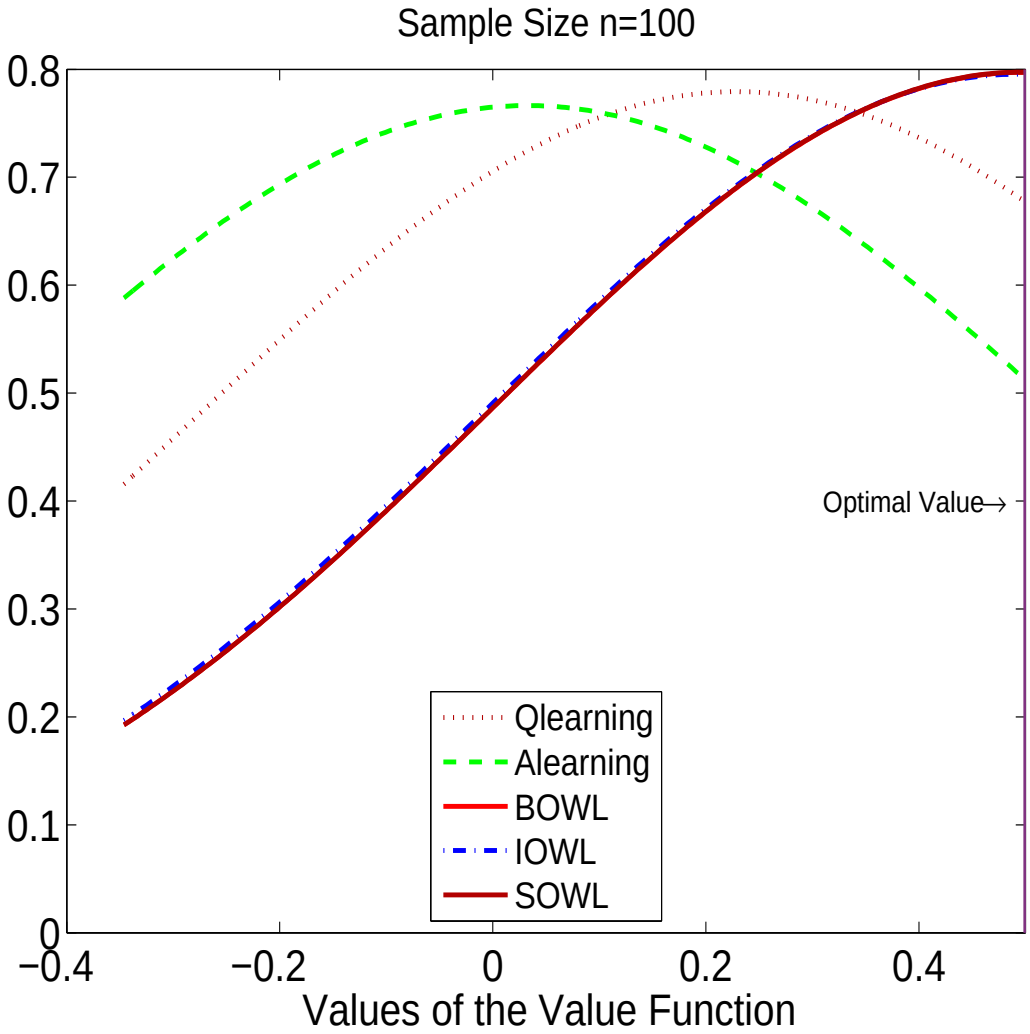
  - Bound on both estimation error and approximation error.
  - Under certain margin conditions (differential treatment effects), $\gamma = 1$.

Simulation Studies

- Methods: BOWL/IOWL/SOWL with Linear kernel; Q-learning with linear regression; A-learning with linear basis for regret function.
- Three two-stage scenarios:
  1. Linear but $Y_1 = 0$ (hard for Q-learning).
  2. Nonlinear.
  3. Nonlinear and $Y_1 = 0$.
- Training data sample size $n = 100, 200, 400$.
- Testing data sample size 10000.
- 500 replications.

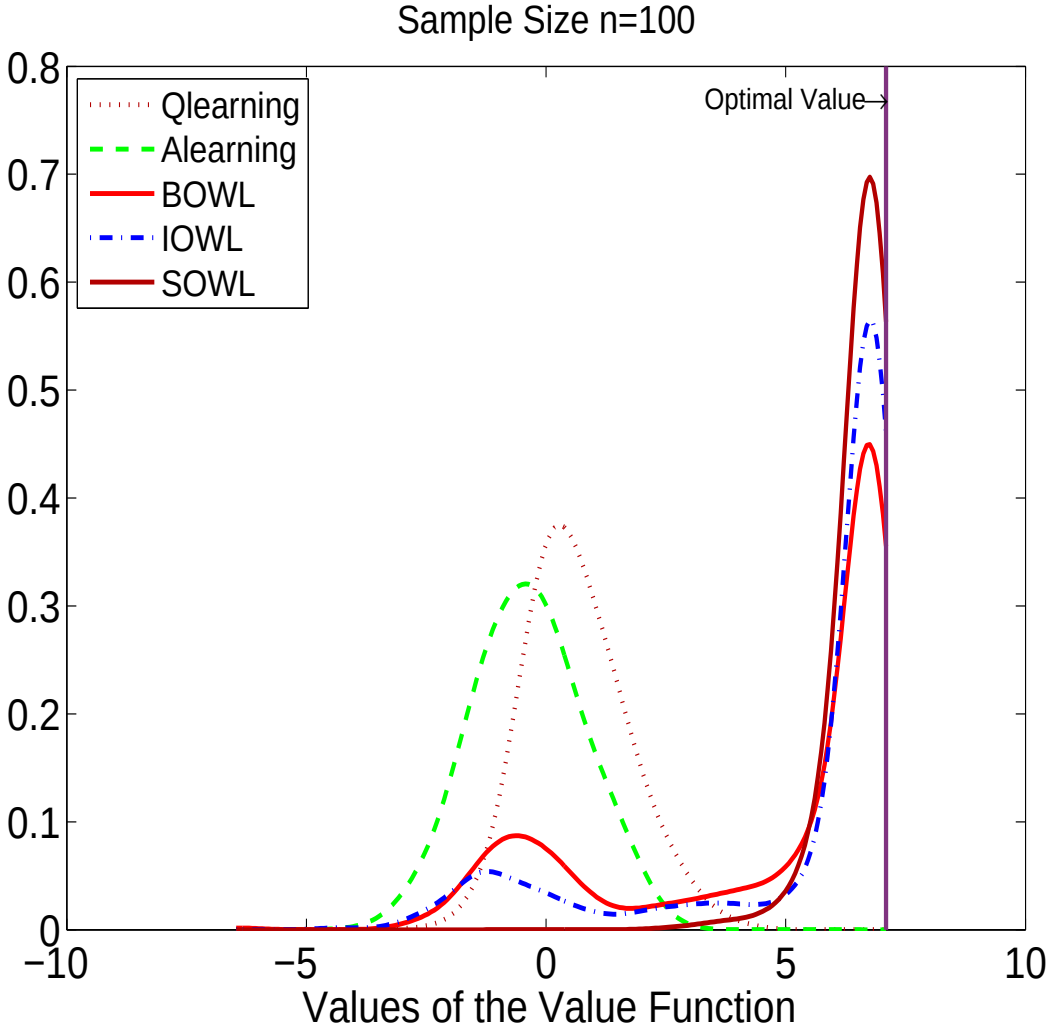
Simulation Results: Scenario 1

Figure: Smoothed Histograms of Values of Estimated DTRs ($n = 100$)



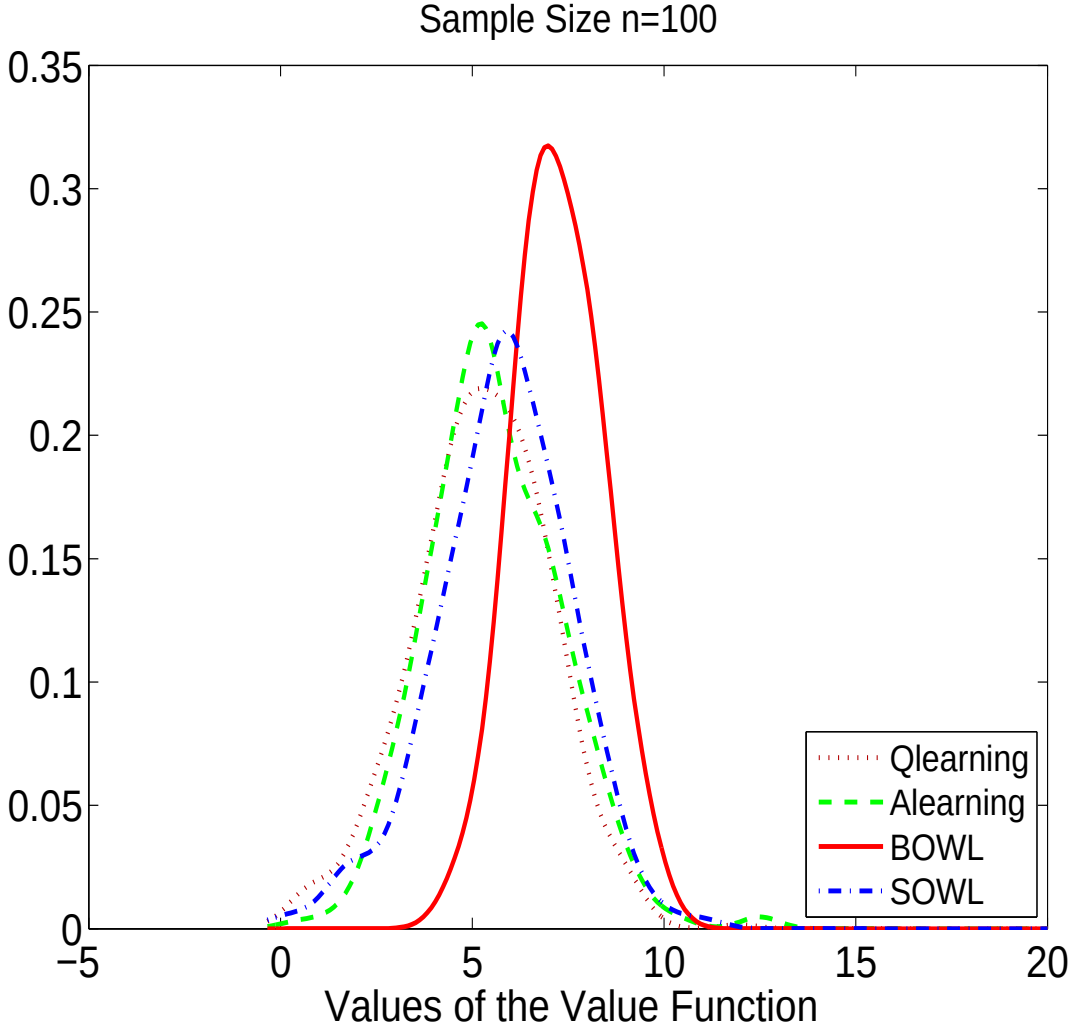
Simulation Results: Scenario 2

Figure: Smoothed Histograms of Values of Estimated DTRs ($n = 100$)



Simulation Results: Scenario 3

Figure: Smoothed Histograms of Values of Estimated DTRs ($n = 100$)



Smoking Cessation Study

- Stage I: Project Quit,
  - Baseline variables: Age, Gender, Education, Race, Baseline motivation to quit smoking, etc.
  - Treatment: highly or lowly tailored Story (denoted by 1 or -1).
- Stage II: Forever Free,
  - Intermediate variables: Motivation to quit smoking and Self efficacy at 6 months.
  - Treatment or not (denoted by 1 or -1).
- Outcome Y_{Q1} (Y_{Q2}): Quit status at j^{th} stage, $j = 1, 2$.
- Sample size: 281 subjects completed the entire study.

Smoking Cessation Study

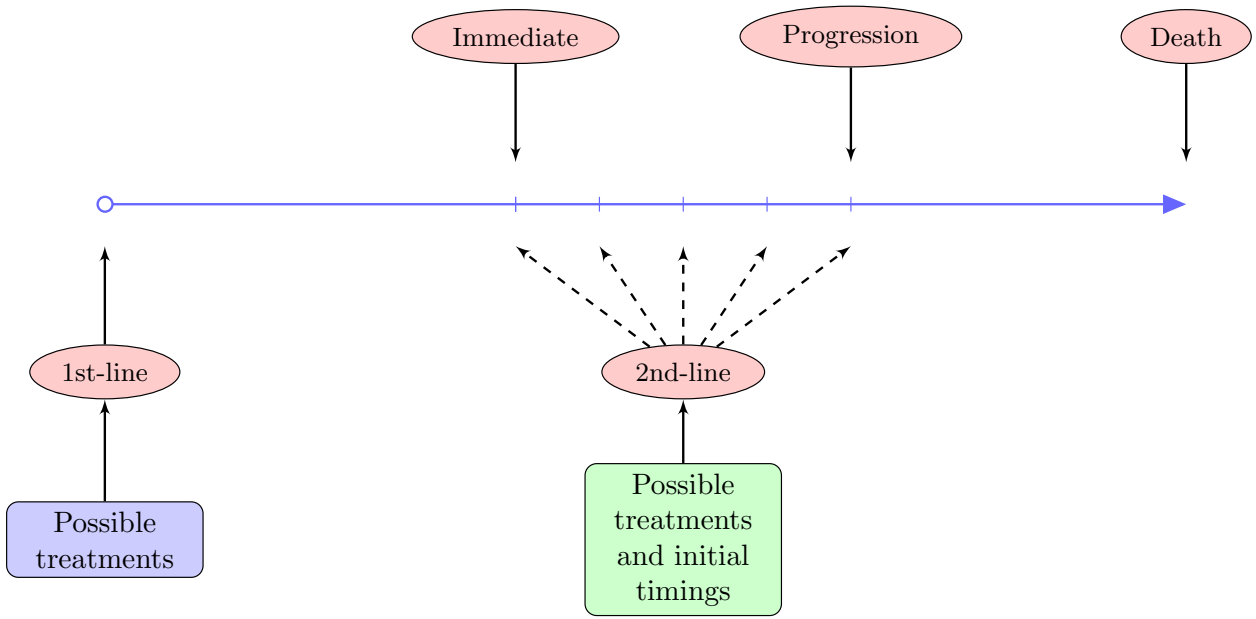
- Cross validation type analysis.
- Additional outcomes
  - $Y_{Sj} = 1$ if the level of satisfaction with the smoking cessation program is high and $Y_{Sj} = 0$ otherwise, $j = 1, 2$;
  - $Y_{Nj} = 0$ if the patient had zero abstinent months; $Y_{Nj} = 1$ if 1-3 abstinent months; and $Y_{Nj} = 2$ if 4 or more abstinent months during the j^{th} stage.

Table: Mean Cross Validated Values using Different Methods

	Mean Cross Validated Values				
Outcome	BOWL	IOWL	SOWL	Q-learning	A-learning
Y_Q	0.747 (0.010)	0.768 (0.010)	0.755 (0.012)	0.692 (0.008)	0.709 (0.008)
Y_N	1.550 (0.031)	1.534 (0.040)	1.500 (0.026)	1.487 (0.020)	1.453 (0.023)
Y_S	1.262 (0.009)	1.288 (0.013)	1.203 (0.015)	1.216 (0.008)	1.183 (0.007)

Open Questions

- Survival outcomes
- Multicategory/Continuous treatments.
  - Multiple therapies.
  - Continuous range of dose levels.
- Optimize timing to switch treatments in multi-stage trials.



Other Open Questions

- Development of meaningful inference tools: this is hard even for linear regression in Q-learning.
- Develop sample size algorithms or formulas.
- When should parametric or semiparametric approaches be used instead of machine learning approaches?
- How to design trials for long-term chronic diseases.
- How to elicit and formulate outcomes (utility).
- How to handle continuing reassessment so that previously developed regimens could be enlarged to include new and emerging treatments.

Preparing Protocols

- Each setting seems to be unique.
- Often best to frame the trial first as a traditional trial with randomized treatments and then add personalized medicine and dynamic treatment regimen aspects as later aims.
- There are ways to frame dynamic treatment regimen estimation, in some cases, as weighted linear regression.
- Sample sizes roughly correspond to large traditional Phase II (or small Phase III) designs for SMART trials.
- We are working on sample size software for OWL studies.
- We have completed or are working on about 5 such trials.

Closing Comments

- We know that these methods work and that they are ready to be utilized in clinical research.
- The methods have the potential to improve patient care dramatically without full understanding of underlying mechanisms.
- We know how to prepare protocols and believe that generally they are approvable by IRBs and funding agencies.
- These approaches require specialized biostatistical knowledge to design and analyze.
- We invite interested researchers to work with us in designing and implementing these kinds of studies.
- We are available to help.

References:

- Chakraborty, B., Murphy, S. A., and Strecher, V. (2009). Inference for non-regular parameters in optimal dynamic treatment regimens. *Statistical Methods in Medical Research* 19:317–343.
- Goldberg, Y., and Kosorok, M. R. (2012a). Support Vector Regression for Right Censored Data. Submitted.
- Goldberg, Y., and Kosorok, M. R. (2012b). Q-learning with censored data. *Annals of Statistics* 40:529–560.
- Qian, M., and Murphy, S. A. (2011). Performance guarantees for individualized treatment rules. *Annals of Statistics* 39:1180–1210.
- Tang, Y., and Kosorok, M. R. (2012). Developing adaptive personalized therapy for cystic fibrosis using reinforcement learning. UNC-Chapel Hill Biostatistics Technical Report 30.

References, cont.:

- Tsybakov, A. B. (2004). Optimal aggregation of classifiers in statistical learning. *Annals of Statistics* 32:135–166.
- Zhao, Y., Zeng, D., Socinski, M. A., and Kosorok, M. R. (2011). Reinforcement learning strategies for clinical trials in non-small cell lung cancer. *Biometrics* 67:1422–1433.
- Zhao, Y., Zeng, D., Rush, A. J., and Kosorok, M. R. (2012a). Estimating individualized treatment rules using outcome weighted learning. *JASA* 107:1106-1118.
- Zhao, Y., Zeng, D., Laber, E. B., and Kosorok, M. R. (2012b). New statistical learning methods for estimating optimal dynamic treatment regimens. Submitted.
- Zhu, R., and Kosorok, M. R. (2012). Recursively imputed survival trees. *JASA* 107:331–340.