

Personalized Medicine and Artificial Intelligence

Michael R. Kosorok, Ph.D.

Department of Biostatistics
University of North Carolina at Chapel Hill

Summer, 2012

Outline

- ① Overview of Personalized Medicine
 - Introduction
 - Current Approaches
- ② Progress on Single-Decision Regime Discovery
 - Methodology
 - Theoretical Results
 - Simulation Studies and Data Analysis
 - Comments
- ③ Progress on Multi-Decision (Dynamic) Regime Discovery
 - Framework
 - Example
 - New Developments
- ④ Overall Conclusions and Open Questions

Part I

Overview of Personalized Medicine

Personalized Medicine

- **What is Personalized Medicine?**
Customized healthcare decisions and practices for the individual patient.
- **Why Do We Need Personalized Medicine?**



- Multiple active treatments available.
- Heterogeneity in responses:
 - ① **Across patients:** what works for one may not work for another.
 - ② **Within a patient:** what works now may not work later.

Nonpsychotic Chronic Major Depressive Disorder (Single-Decision)

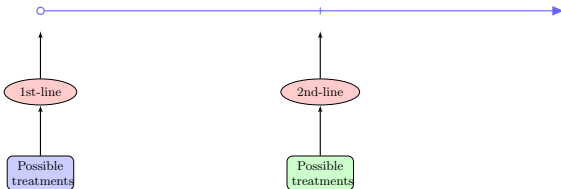
- The goal of the Nefazodone-CBASP clinical trial (Keller et al., 2000) is to determine the best treatment choice among
 - **Pharmacotherapy** (nefazodone).
 - **Psychotherapy** (cognitive behavioral-analysis system of psychotherapy (CBASP)).
 - **Combination of both.**
- 681 patients, with 50 prognostic variables measured on each patient.

Further Goal

Can we reduce depression by creating individualized treatment rules based on prognostic data?

Late Stage Non-Small Cell Lung Cancer (Multi-Decision)

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 (at the beginning of a treatment line) based on prognostic data?

The Basic Process

Current approaches to developing personalized medicine typically includes five key elements:

- obtaining patient genetic/genomic data using array and other high throughput technology;
- identifying one or more biomarkers;
- developing new or selecting available therapies;
- measuring the relationship between biomarkers and clinical outcomes, including prognosis and response to therapy; and
- verifying the relationship in a prospective randomized clinical trial.

Review of Personalized Medicine (2006-2010)

- We now summarize studies on personalized medicine published in six high-impact journals — *Journal of the American Medical Association*, *Journal of the National Cancer Institute*, *Lancet*, *Nature*, *Nature Medicine*, and *the New England Journal of Medicine* — from 2006 to 2010.
- All papers were manually selected and reviewed based on specified inclusion and exclusion criteria.

- 76 articles were selected meeting the above criteria, but two have since been retracted and were not included, resulting in 74 articles for our sample, 53 of which were cancer-related.
- In all 74, a biomarker was used to stratify patients for differential treatment.

Data Driven versus Knowledge Driven

- Because of the so-called “curse of dimensionality,” identifying potential biomarkers from patient genomic profiles is a tremendous challenge.
- In the studies reviewed, two main approaches were uncovered for identifying the needed biomarkers:
 - a data-driven approach using primarily empirical methods and
 - a knowledge-driven approach using existing biological knowledge about functions of genes, proteins, pathways and mechanisms.
- 56 papers developed new biomarkers: 16 based on data-driven approach, 36 knowledge driven, 4 hybrid.

Prognostic vs. Predictive Biomarkers

Two types of relationships between biomarkers and clinical outcomes were observed in the reviewed studies:

- association between biomarkers and patient prognosis (*prognostic biomarkers*) and
- association between biomarkers and response to treatment (*predictive biomarkers*).

In the reviewed studies:

- 19 compared different treatments for one patient group;
- 33 studied the same therapy across different groups; and
- 16 made both types of comparisons.

Reliability and Reproducibility

- A continuing controversy of personalized medicine focuses on its reliability and reproducibility (two of the studies reviewed were retracted because of non-replicability).
- The complexity of the data and statistical analyses involved make study of reproducibility of results both difficult and important:
 - datasets must be made publicly available for verification;
 - biomarkers need to be validated in a different group of patients;
 - quality data management is another important issue;
 - creative statistical methods are needed.

Several recommendations regarding these issues have been made and more are to come.

Statistical and Computational Task and Challenges

Task

Develop statistically efficient clinical trial designs and analysis methods for discovering individualized treatment rules.

- **Predictors:** Medical records, Diagnostic test, Demographics, Imaging, Genetics, Genomics, Proteomics

Challenges

- Identify the optimal individualized treatment rule using training data where **optimal treatment is unknown**.
- High-dimensional predictors; arbitrary order nonparametric interactions.
- Longitudinal data: sequentially dependent.

Part II

Progress on Single-Decision Regime Discovery

Single Decision: Data and Goal

- Observe independently and identically distributed training data $(X_i, A_i, R_i), i = 1, \dots, n$.
 - X : baseline variables, $X \in \mathbb{R}^d$,
 - A : binary treatment options, $A \in \{-1, 1\}$,
 - R : outcome (larger is better), $R \in \mathbb{R}^+$, R is bounded.
- Randomized study with known randomization probability of the treatment.
- Construct individualized treatment rule (ITR)

$$\mathcal{D}(X) : \mathbb{R}^d \rightarrow \{-1, 1\}.$$

Goal

Maximize the expected outcome if the ITR is implemented in the future.

Standard Approach and Challenges

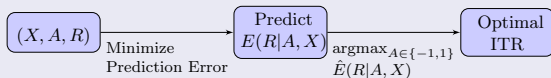
- Standard approach:
 - Use regression and/or machine learning (e.g., support vector regression (SVR)) to estimate

$$Q(x, a) = E(R|X = x, A = a)$$

- $\hat{D}_n(x) = \operatorname{argmax}_a \hat{Q}_n(x, a)$.
- Issues:
 - For right-censored outcomes, we developed improved random forrests (Zhu and Kosorok, 2012, *JASA*) and SVR (Goldberg and Kosorok, 2012, Submitted).
 - The current approach is indirect, since we must estimate $Q(x, a)$ and invert to estimate $\mathcal{D}(x)$.

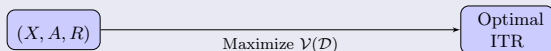
Optimal Individualized Treatment Rule Discovery

Traditional approach: regression-based



Problem: mismatch between minimizing the prediction error and maximizing the value function.

Our approach



Can we directly estimate the decision rule which maximizes the value function?

Value Function and Optimal Individualized Treatment Rule

- Let P denote the distribution of (X, A, R) , where treatments are randomized, and $P^{\mathcal{D}}$ denoted the distribution of (X, A, R) , where treatments are chosen according to \mathcal{D} . The **value function of \mathcal{D}** (Qian & Murphy, 2011) is

$$\mathcal{V}(\mathcal{D}) = E^{\mathcal{D}}(R) = \int R dP^{\mathcal{D}} = \int R \frac{dP^{\mathcal{D}}}{dP} dP = E \left[\frac{I(A = \mathcal{D}(X))}{P(A|X)} R \right].$$

- Optimal Individualized Treatment Rule:**

$$\mathcal{D}^* \in \underset{\mathcal{D}}{\operatorname{argmax}} \mathcal{V}(\mathcal{D}).$$

$$E(R|X, A = 1) > E(R|X, A = -1) \Rightarrow \mathcal{D}^*(X) = 1$$

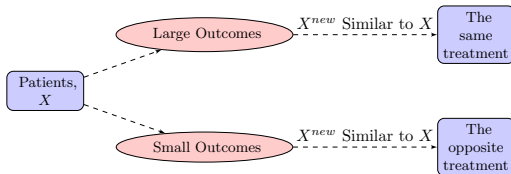
$$E(R|X, A = 1) < E(R|X, A = -1) \Rightarrow \mathcal{D}^*(X) = -1$$

Classification Perspective

Intuition: Classification (Artificial Intelligence and Statistical Learning)

Given a new observation X^{new} , predict the class label $\mathcal{D}^{*,\text{new}}$.

- No direct information on the true class labels, \mathcal{D}^* .
- Can we assign the right treatment based on the observed information?



Outcome Weighted Learning (OWL)

Optimal Individualized Treatment Rule \mathcal{D}^*

Maximize the value

$$E \left[\frac{I(A = \mathcal{D}(X))}{P(A|X)} R \right]$$

Minimize the risk

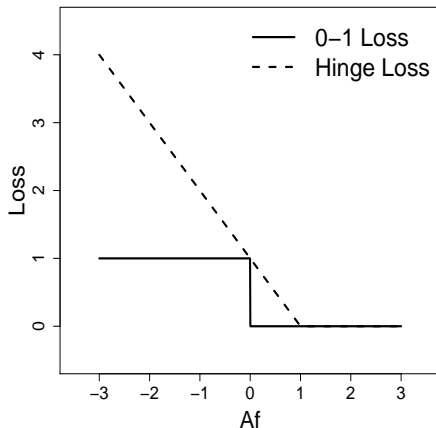
$$E \left[\frac{I(A \neq \mathcal{D}(X))}{P(A|X)} R \right]$$

- For any rule \mathcal{D} , $\mathcal{D}(X) = \text{sign}(f(X))$ for some function f .
- Empirical approximation to the risk function:

$$n^{-1} \sum_{i=1}^n \frac{R_i}{P(A_i|X_i)} I(A_i \neq \text{sign}(f(X_i))).$$

- **Computation challenges:** non-convexity and discontinuity of 0-1 loss.

Convex Surrogate Loss: Hinge Loss



Hinge Loss: $\phi(Af(X)) = (1 - Af(X))^+$, where $x^+ = \max(x, 0)$

Outcome Weighted Support Vector Machine (SVM)

Objective Function: Regularization Framework

$$\min_f \left\{ \frac{1}{n} \sum_{i=1}^n \frac{R_i}{P(A_i|X_i)} \phi(A_i f(X_i)) + \lambda_n \|f\|^2 \right\}. \quad (1)$$

- $\|f\|$ is some norm for f , and λ_n controls the severity of the penalty on the functions.
- A linear decision rule: $f(X) = X^T \beta + \beta_0$, with $\|f\|$ as the Euclidean norm of β .
- **Estimated individualized treatment rule:**

$$\hat{D}_n = \text{sign}(\hat{f}_n(X)),$$

where \hat{f}_n is the solution to (1).

Computation and Kernel Trick

- The dual problem is a convex optimization problem.
- Quadratic programming; Karush-Kuhn-Tucker conditions.
- Linear decision rules may be insufficient.
- Kernel trick, $k : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$.
- Nonlinear decision rule with $f(x) = \beta k(\cdot, x) + \beta_0$.
- Reproducing kernel Hilbert space (RKHS) \mathcal{H}_k with norm denoted by $\|\cdot\|_k$:

$$\mathcal{H}_k = \left\{ g(x) = \sum_{i=1}^m \alpha_i k(x_i, x) \right\}.$$

- A linear kernel yields a linear decision rule.

Risk Bound and Convergence Rates of the OWL Estimator

- Understand the accuracy of OWL procedure.
- Fisher consistent, consistent, and general risk bounds.
- Precise risk bound under certain regularity conditions.
- The value converges surprisingly fast to the optimal, almost as fast as n^{-1} .
- Similar to rate results in SVM literature (Tsybakov, 2004).

Empirical Study

- OWL with Gaussian kernel: two tuning parameters
 - λ_n : the parameter for penalty.
 - σ_n : the inverse bandwidth of the kernel.
- Methods for comparison:
 - OWL with Linear kernel.
 - Regression based methods:
 - h_1 penalized least squares (h_1 -PLS) (Qian & Murphy, 2011) with basis function $(1, X, A, XA)$.
 - Ordinary Least Squares (OLS) with basis function $(1, X, A, XA)$.
- Evaluation of values in terms of mean squared error (MSE).
 - 1000 replications; each training data set is of size 100, 200, 400 or 800.
 - Independent validation set of size 10000.

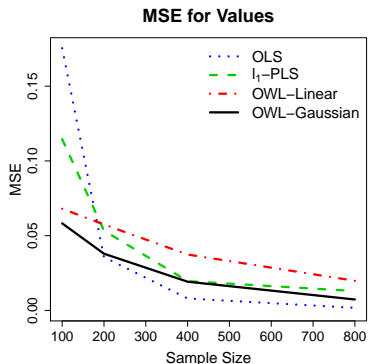
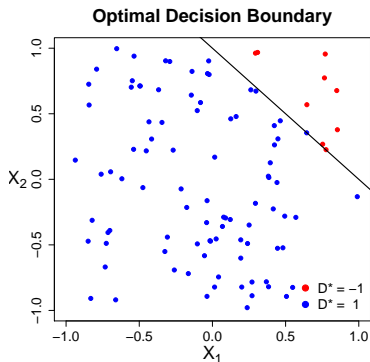
Data Generation

- $X = (X_1, \dots, X_{50}) \sim U[-1, 1]^{50}$.
- $A \in \{-1, 1\}$, $P(A = 1) = P(A = -1) = 0.5$.
- The response $R \sim N(Q_0, 1)$, where

$$Q_0 = 1 + 2X_1 + X_2 + 0.5X_3 + T_0(X, A).$$

- 1 $T_0(X, A) = 0.442(1 - X_1 - X_2)A.$
- 2 $T_0(X, A) = (X_2 - 2X_1^3 - 0.1) A.$
- 3 $T_0(X, A) = (0.5 - X_1^2 - X_2^2) (X_1^2 + X_2^2 - 0.3) A.$

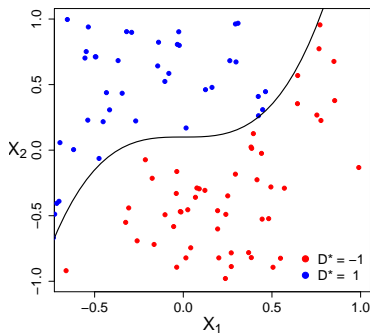
Simulation Results



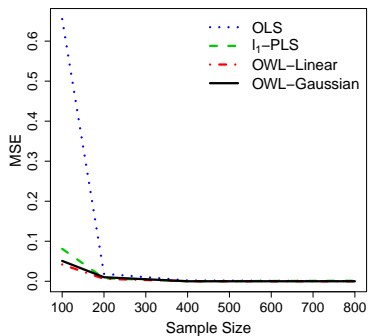
Scenario 1: $T_0(X, A) = 0.442(1 - X_1 - X_2)A$

Simulation Results

Optimal Decision Boundary



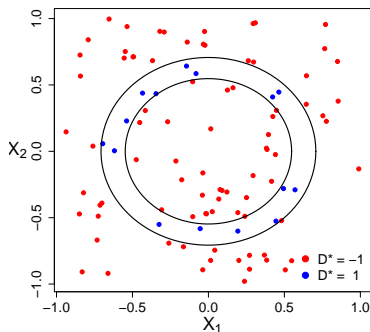
MSE for Values



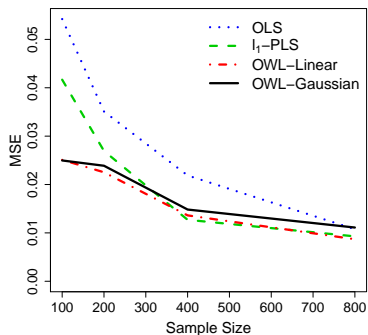
Scenario 2: $T_0(X, A) = (X_2 - 2X_1^3 - 0.1) A$

Simulation Results

Optimal Decision Boundary

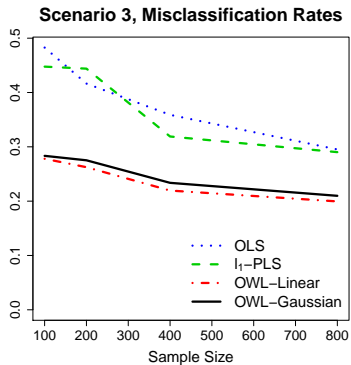


MSE for Values



Scenario 3: $T_0(X, A) = (0.5 - X_1^2 - X_2^2) (X_1^2 + X_2^2 - 0.3) A$

Simulation Results: Misclassification



Nefazodone-CBASP clinical trial (Keller et al., 2000)

- 681 patients with non-psychotic chronic major depressive disorder (MDD).
- Randomized in a 1:1:1 ratio to either nefazodone, cognitive behavioral-analysis system of psychotherapy (CBASP) or the combination of nefazodone and psychotherapy.
- Primary outcome: score on the 24-item Hamilton Rating Scale for Depression (HRSD); the lower the better.
- 50 baseline variables: demographics, psychological problem diagnostics etc.

Nefazodone-CBASP clinical trial (Keller et al., 2000)

Pairwise Comparison:

- OWL: gaussian kernel.
 l_1 -PLS and OLS: $(1, X, A, XA)$.
- Value calculated with a 5-fold cross validation type analysis.

Table 1: Mean HRSD (Lower is Better) from Cross Validation Procedure with Different Methods

	OLS	l_1 -PLS	OWL
Nefazodone vs CBASP	15.87	15.95	15.74
Combination vs Nefazodone	11.75	11.28	10.71
Combination vs CBASP	12.22	10.97	10.86

Comments

The Outcome Weighted Learning procedure

- Discovers an optimal individualized therapy to improve expected outcome.
- Nonparametric approach sidesteps the inversion step and invokes statistical learning techniques directly.

Some open questions:

- How to handle censoring?
- How to generate sample size formulas to enable practical Phase II design?

Part III

Progress on Multi-Decision (Dynamic) Regime Discovery

Dynamic Treatment Regimes (DTR)

Observe data on n individuals, T stages for each individual,

$$X_1, A_1, X_2, A_2, \dots, X_T, A_T, X_{T+1}$$

X_t : Observation available at the t^{th} stage.

A_t : Treatment at the t^{th} stage, $A_t \in \{-1, 1\}$.

H_t : History available at the t^{th} stage, $H_t = \{X_1, A_1, X_2, \dots, A_{t-1}, X_t\}$.

R_t : Outcome following the t^{th} stage, $R_t = r_t(H_{t+1})$.

A DTR is a sequence of decision rules:

$$\mathcal{D} = (\mathcal{D}_1(H_1), \dots, \mathcal{D}_T(H_T)), \mathcal{D}_t(H_t) \in \{-1, 1\}.$$

Goal

Maximize the **expected sum of outcomes** if the DTR is implemented in the future.

Value Function and Optimal DTR for Two Stages

- The value function: $\mathcal{V}(\mathcal{D}) = E^{\mathcal{D}}(R_1 + R_2)$.
- Optimal DTR: $\mathcal{D}^* = \operatorname{argmax}_{\mathcal{D}} \mathcal{V}(\mathcal{D})$.
- Constructing Optimal DTRs based on Q functions:

$$Q_2(h_2, a_2) = E(R_2 | H_2 = h_2, A_2 = a_2)$$

$$\mathcal{D}_2^*(h_2) = \operatorname{argmax}_{a_2} Q_2(h_2, a_2)$$

$$Q_1(h_1, a_1) = E(R_1 + \max_{a_2} Q_2(H_2, a_2) | H_1 = h_1, A_1 = a_1)$$

$$\mathcal{D}_1^*(h_1) = \operatorname{argmax}_{a_1} Q_1(h_1, a_1)$$

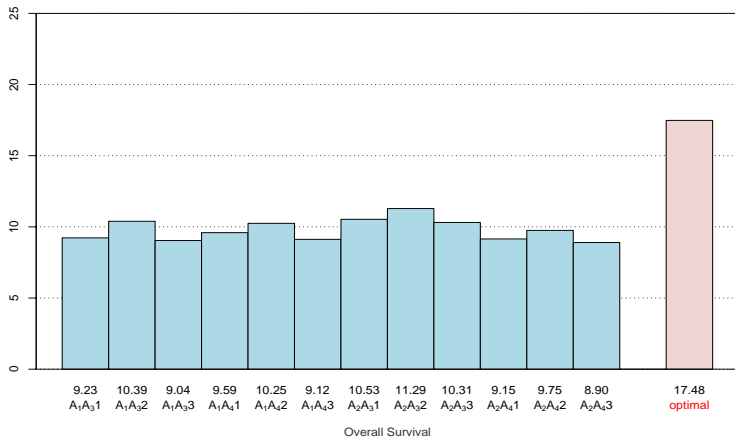
- Q learning with regression: estimate the Q-functions from data using regression and then find the optimal DTR.

Non-Small Cell Lung Cancer (Yufan Zhao et al., 2011)

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.

Performance of Optimal Personalized Versus Fixed Regimens



Standard Approach and Challenges

- Standard approach:
 - Use regression and/or machine learning (e.g., SVR) to estimate the Q-functions sequentially backwards.
 - At time t , use as outcome the estimated pseudo-value

$$R_t + \max_{a_{t+1}} \hat{Q}_{t+1}(H_{t+1}, a_{t+1}).$$

- Issues:
 - For right-censored outcomes, we developed Q-learning for censored data and possibly irregular number and spacing of decision times (Goldberg and Kosorok, 2012, AOS).
 - As before, the current approach is indirect, since we must estimate $Q_t(h, a)$ and invert to estimate $\mathcal{D}_t(h)$.

Backwards Outcome Weighted Learning (BOWL)

Problem with Q learning

- Mismatch exists between estimating the optimal Q function and the goal of maximizing the value function (Murphy, 2005).
- Non-smooth maximization operation.
- High dimensional covariate space.

BOWL

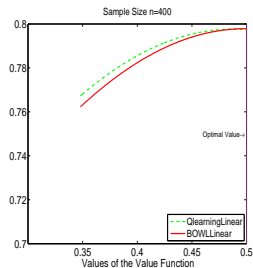
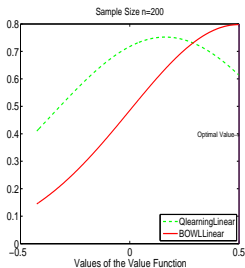
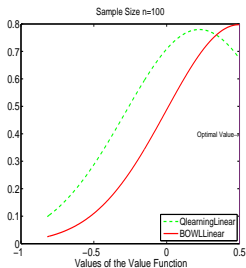
- Generalization of OWL to multi-decision setup.
- Find the optimal decision rule by **directly maximizing the value function** for each stage backwards repeatedly.
- Consistency and risk bound of BOWL estimator.

Simulation Study

Generative Model (Chakraborty et al., 2010)

- $X_1 \sim U[-1, 1]^{50}$, $X_2 = X_1$.
 - $A_1, A_2 \in \{-1, 1\}$, $P(A_1 = 1) = P(A_2 = 1) = 0.5$.
 - $R_1 = 0$, $R_2|H_2, A_2 \sim N(-0.5A_1 + 0.5A_2 + 0.5A_1A_2, 1)$.
-
- Training data sample size $n = 100, 200, 400$.
 - Testing data sample size 10000.
 - 500 replications.
 - Methods: BOWL with Gaussian/Linear kernel; Q learning with linear regression.

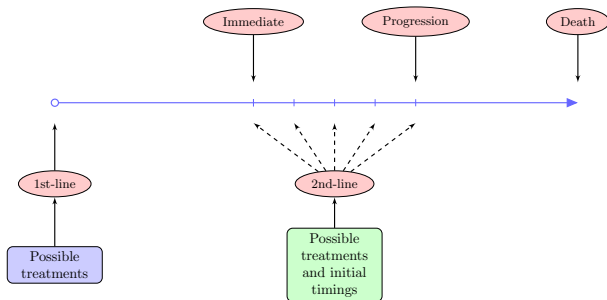
Simulation Results



Note: Q learning encounters difficulties with small sample sizes.

Open Issues for BOWL

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



Part IV

Overall Conclusions and Open Questions

Conclusions

- Single- and multi- decision personalized medicine trials can discover effective individualized regimens that improve significantly over standard approaches.
- Artificial intelligence and statistical learning tools play a significant role in new developments.
- The sample sizes required are usually reasonable.
- For the multi-decision setting, good dynamic models (both mechanistic and stochastic) are needed to construct virtual patients and virtual trials before designing trials.
- The advantage is the discovery of effective new treatments that could be missed by conventional approaches.

Open Questions

- Better tools for high-dimensional data: interpretability and simplicity.
- Inference for individualized treatment regimes: limiting distribution of the value function and sample size formula in both single-decision and multi-decision setup.
- Survival data (for OWL and BOWL, etc.).
- Missing data.
- Observational studies.

Acknowledgments

- Yingqi Zhao
- Yufan Zhao
- Zheng Ren
- Yair Goldberg
- Donglin Zeng
- Eric Laber
- Mark A. Socinski, A. John Rush and Richard M. Goldberg
- Marie Davidian and Stephen L. George
- Fred A. Wright and Anastasios A. Tsiatis
- Min Qian and Lacey Gunter

References

- Chakraborty et al. (2010). Inference for non-regular parameters in optimal dynamic treatment regimes. *Statistical Methods in Medical Research* 19:317 - 343.
- Goldberg, Y., & Kosorok, M. R. (2012). Q-learning with censored data. *Annals of Statistics* 40:529-560.
- Keller, M. B. et al. (2000). A Comparison of Nefazodone, The Cognitive Behavioral-Analysis System of Psychotherapy, and Their Combination for the Treatment of Chronic Depression. *NEJM* 342(20):1462-1470.
- Murphy S.A. (2005). A Generalization Error for Q-Learning. *Journal of Machine Learning Research* 6:1073-1097.

References (continued)

- Qian, M., & Murphy, S. A. (2011). Performance Guarantees for Individualized Treatment Rules. *Annals of Statistics* 39:1180-1210.
- Tsybakov, A. B. (2004). Optimal Aggregation of Classifiers in Statistical Learning. *Annals of Statistics* 32:135-166.
- Zhao, Yingqi, et al. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, In press.
- Zhao, Yufan, et al. (2011). Reinforcement learning strategies for clinical trials in non-small cell lung cancer. *Biometrics*, 67:1422 - 1433.
- Zhu, R., & Kosorok, M. R. (2012). Recursively imputed survival trees. *Journal of the American Statistical Association* 107:331-340.