

Recent Advances in Outcome Weighted Learning for Precision Medicine

Michael R. Kosorok

Department of Biostatistics
Department of Statistics and Operations Research
University of North Carolina at Chapel Hill

September 15, 2017

Acknowledgments

- ▶ Many co-authors and collaborators helped with this work.
- ▶ Especially Guanhua Chen (UW-Madison), Yifan Cui (PhD student in Statistics and Operations Research at UNC-Chapel Hill), Anna Kahkoska (MD-PhD student at UNC-Chapel Hill), Eric Laber (NCSU), Daniel Lockett (PhD student in Biostatistics at UNC-Chapel Hill), David Maahs (Stanford), Elizabeth Mayer-Davis (UNC-Chapel Hill), Donglin Zeng (UNC-Chapel Hill), and Ruoqing Zhu (UI-Urban-Champaign).
- ▶ This project was funded in part by US NIH grants P01 CA142538 and UL1 TR001111, and US NSF grant DMS-1407732, among others.

Outline

Introduction

Outcome weighted learning

A Tree Based Approach for Censored Data

Dose finding

Precision medicine in mHealth

Overall conclusions and future work

Personalized vs. precision medicine

- ▶ Personalized medicine
 - ▶ The idea of developing targeted treatments that take into account patient heterogeneity
 - ▶ Clinicians have been doing this for centuries: *not really new*
- ▶ Precision medicine
 - ▶ The idea of developing targeted treatments that take into account patient heterogeneity
 - ▶ Empirically based, scientifically rigorous, reproducible, and generalizable (i.e., will work with future patients): *very new*

Ingredients of precision medicine

- ▶ Incorporates data on:
 - ▶ Genetic makeup, proteomics, metabolomics, metabiomics, etc.
 - ▶ Environmental factors, demographics, etc.
 - ▶ Lifestyle, phenotypic information, etc.
- ▶ Scientific tools:
 - ▶ Biomedical knowledge
 - ▶ Data (potentially integrated across many platforms)
 - ▶ Knowledge driven vs. data driven
 - ▶ Computational, mathematical and statistical tools

Clinical focus

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

- ▶ *The single-decision setting:*
 - ▶ A patient presents with a disease and we need to decide what treatment (or dose) to give from a list of choices.
 - ▶ We want to make the best decision (treatment regimen) based on available baseline patient-level feature data.
- ▶ The multi-decision setting:
 - ▶ Treat patients for diseases with multiple treatment decisions occurring over time.
 - ▶ Make the best decision based on historical patient-level data at each decision time (dynamic treatment regimen).
 - ▶ The best decisions take into account delayed effects.
- ▶ Real time decision making in mHealth:
 - ▶ A large number of decisions need to be made in real time
 - ▶ Technical and practical challenges for implementing

Outline

Introduction

Outcome weighted learning

A Tree Based Approach for Censored Data

Dose finding

Precision medicine in mHealth

Overall conclusions and future work

Single decision setting

- ▶ Let X be the vector of patient tailoring variables, A the choice of treatment given, and R the clinical outcome (with larger being better).
- ▶ The “standard” approach is to first estimate the Q-function

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

through regression of R on (X, A) , and invert to obtain

$$\hat{d}(x) = \underset{a}{\operatorname{argmax}} \hat{Q}(x, a).$$

- ▶ Issue: This approach is indirect, since we must estimate $Q(x, a)$ and invert.

Value function and optimal individualized treatment rule

- ▶ Let P be the distribution of (X, A, R) , with treatments randomized via $\pi(A|X)$, and P^d the distribution of (X, A, R) , with treatments chosen according to d . The value function of d (Qian & Murphy, 2011) is

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

- ▶ Optimal Individualized Treatment Rule:

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

$$E(R|X, A=1) > E(R|X, A=-1) \Rightarrow d^*(X) = 1$$

$$E(R|X, A=1) < E(R|X, A=-1) \Rightarrow d^*(X) = -1$$

Outcome weighted learning (OWL or O-learning)

Optimal Individualized Treatment Rule d^*

Maximize the value

$$E \left[\frac{I(A = d(X))}{\pi(A|X)} R \right]$$

Minimize the risk

$$E \left[\frac{I(A \neq d(X))}{\pi(A|X)} R \right]$$

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

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

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

Using a support vector machine (SVM) approach

Objective Function: Regularization Framework

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

- ▶ $\phi(u) = (1 - u)^+$ is the hinge loss surrogate, $\|f\|$ is some norm for f , and λ_n controls the penalty on f .
- ▶ 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).

Outcome Weighted Learning (OWL or O-Learning)



Figure 1: Zhao YQ, Zeng D, Rush AJ, and Kosorok MR (2012). Estimating individualized treatment rules using outcome weighted learning. *JASA* 107:1106–1118.

Results for O-Learning

- ▶ Can use kernel trick to extend to nonparametric decision rule (e.g., the Gaussian kernel).
- ▶ Fisher consistent, asymptotically consistent, and model robust.
- ▶ Risk bounds and convergence rates similar to those observed in SVM literature (Tsybakov, 2004).
- ▶ Excellent simulation results and data analysis of Nefazodone-CBASP clinical trial (Keller et al., 2000).
- ▶ Promising performance overall (Y.Q. Zhao, et al., 2012).
- ▶ An example of a direct value function approach (see also B. Zhang, et al., 2012).
- ▶ Opens door to a unique application of machine learning techniques to personalized medicine.

O-Learning Extensions

- ▶ Multiple decision times:
 - ▶ Zhao YQ, Zeng D, Laber EB, and Kosorok MR (2015). New statistical learning methods for estimating optimal dynamic treatment regimes. *JASA* 110:583–598.
 - ▶ Also works well, but a concern is the low fraction of data being used.
- ▶ Issue: O-learning is not invariant under constant shifts in outcome R .
 - ▶ Zhou X, Mayer-Hamblett N, Khan U, and Kosorok MR (2017). Residual weighted learning for estimating individualized treatment rules. *JASA* 112:169–187.

O-Learning Extensions

- ▶ What about observational data?
 - ▶ One method that seems to work is to replace the known propensity score $\pi(A_i|X_i)$ in (1) with an estimate $\hat{\pi}(A_i|X_i)$.
- ▶ More than two treatment options
 - ▶ Ordinal treatment options (e.g., organized by dose level) (in progress)
 - ▶ Nonordinal treatments (in progress)

O-Learning Extensions

- ▶ Censored data:
 - ▶ Zhao YQ, Zeng D, Laber EB, Song R, Yuan M, and Kosorok MR (2015). Doubly robust learning for estimating individualized treatment with censored data. *Biometrika* 102:151-168.
 - ▶ Two basic approaches:
 - ▶ Inverse weighting: requires correct modeling of conditional distribution function of censoring time given covariates.
 - ▶ Double robust: correct modeling of one of either conditional censoring or conditional survival distribution functions (inverse weighting still used).
 - ▶ Works quite well, but a concern is the need for inverse probability of censoring weighting.

O-Learning Extensions

- ▶ Continuous treatment options (e.g., dose or time)
 - ▶ Chen G, Zeng D, and Kosorok MR (2016). Personalized dose finding using outcome weighted learning (with discussion and rejoinder). *Journal of the American Statistical Association* 111:1509-1547.
- ▶ V-learning for continuous time and mHealth (in progress)

Outline

Introduction

Outcome weighted learning

A Tree Based Approach for Censored Data

Dose finding

Precision medicine in mHealth

Overall conclusions and future work

A Tree Based Approach for Censored Data

- ▶ Cui Y, Zhu R, and Kosorok MR (In press). Tree based weighted learning for estimating individualized treatment rules with censored data. *Electronic Journal of Statistics*.
- ▶ Recall that (X, A, R) is the observed data in outcome weighted learning:
 - ▶ X is the vector of patient-level features.
 - ▶ A is the treatment indicator.
 - ▶ R is the positive outcome: in survival data this is the event time which is censored.

A Tree Based Approach for Censored Data

Recall for outcome weighted learning, the targeted ITR satisfies

$$\mathcal{D}^*(X) = \arg \min_d E \left[\frac{R1\{A \neq d(X)\}}{\pi(A|X)} \right], \quad (2)$$

where we use the hinge loss surrogate to ensure convexity.

However, due to censoring, we do not observe $R = T$ but only observe $Y = T \wedge C$ and $\delta = 1\{Y = T\}$. We will use two nonparametric estimates of T based on censored data:

- ▶ $R_1 = \hat{E}(T|X, A)$ and
- ▶ $R_2 = 1\{\delta = 1\}Y + 1\{\delta = 0\}\hat{E}(T|X, A, Y, T > Y),$

where the \hat{E} terms are estimated via random forests.

Recursively Imputed Survival Trees

We use recursively imputed survival trees (RIST) to estimate \hat{E} :

- ▶ Zhu R and Kosorok MR (2013). Recursively imputed survival trees. *JASA* 107:331–340.

The basic ingredients of RIST:

- ▶ RIST is a random forest method (see, e.g., Breiman, 2001, *Machine Learning*) for right censored data (slightly modified).
- ▶ RIST uses a form of Monte Carlo EM (see, e.g., Wei and Tanner, 1990, *JASA*) to avoid inverse probability of censoring weighting (IPCW) as done in Hothorn et al (2006, *Biostatistics*) or single-step hazard estimation in terminal nodes (see, e.g., Ishwaran et al, 2008, *Annals of Applied Statistics*).

Theoretical Results

Theorem

Under reasonable regularity conditions, there exist sequences $r_n, w_n > 0$ converging to zero such that for each covariate value x and treatment choice a ,

$$\text{pr}\left\{|\widehat{E}(T | x, a) - E(T | x, a)| \leq r_n\right\} \geq 1 - w_n$$

and

$$\begin{aligned} \text{pr}\left\{|\widehat{E}(T | x, a, T > Y, Y) - E(T | x, a, T > Y, Y)| \leq r_n\right\} \\ \geq 1 - 2w_n. \end{aligned}$$

Theoretical Results

Let $d^*(x) = \text{sign}(f^*(x))$ be the decision function which optimizes the value function $V(d)$ over all d and $\hat{d}_n(x) = \text{sign}(\hat{f}_n(x))$ be the estimated decision function.

Theorem

Under reasonable regularity conditions and assuming that the sequence $\lambda_n > 0$ satisfies $\lambda_n \rightarrow 0$ and $\lambda_n \ln n \rightarrow \infty$, there exist sequences $\epsilon_n, w_n > 0$ converging to zero such that

$$pr \left\{ V(f^*) \leq V(\hat{f}_n) + \epsilon_n \right\} \geq 1 - w_n,$$

Simulation Results

We considered several simulation scenarios, but present two here:

- ▶ Scenario 3: Failure time is non-linear with 5 covariates while censoring time follows a Cox model.
- ▶ Scenario 4: Failure time is AFT with 10 covariates while censoring time follows a Cox model.

Other features of both scenarios:

- ▶ 45% censoring rate with $n = 200$ sample size
- ▶ 500 replications and predictions based on log failure time
- ▶ Compare five methods: (1) oracle, (2) RIST- R_1 , (3) RIST- R_2 , (4) ICO, (5) DR, and (6) Cox
- ▶ Both linear and Gaussian kernels used (except for Cox).

Simulation Results

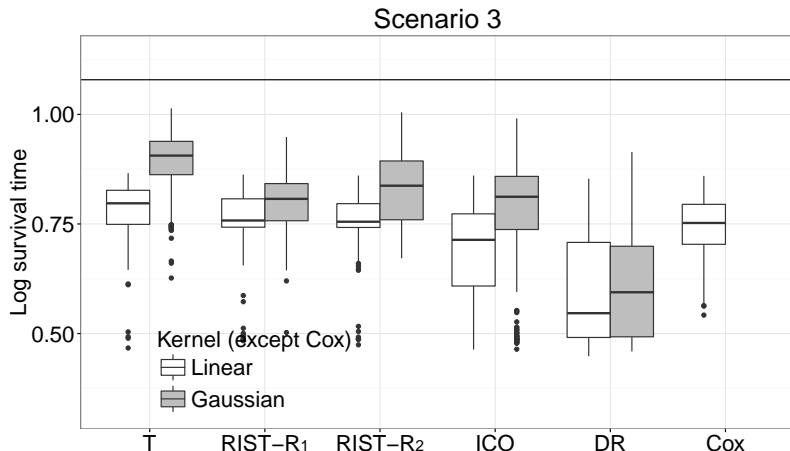


Figure 2: Boxplots of mean log survival time for different treatment regimes. The black horizontal line is the theoretical optimal value.

Simulation Results

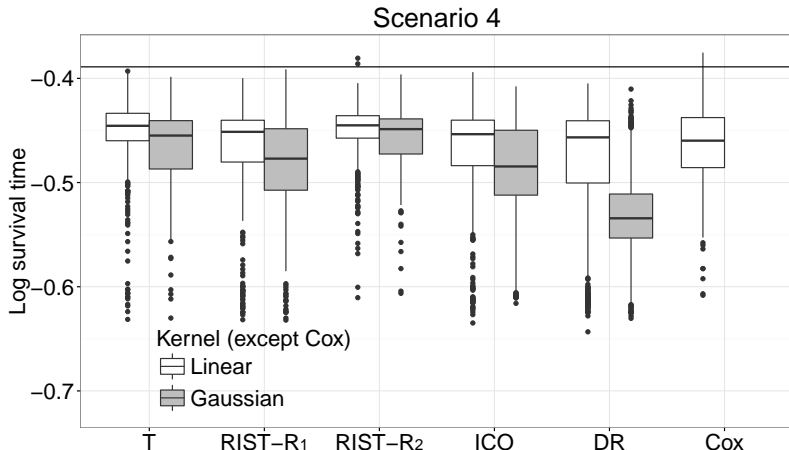


Figure 3: Boxplots of mean log survival time for different treatment regimes. The black horizontal line is the theoretical optimal value.

Non-Small Cell Lung Cancer (NSCLC) Trial Data

We apply the proposed methods to the NSCLC randomized clinical trial data described in Socinski, et al (2002, *Journal of Clinical Oncology* 29:1335–1343):

- ▶ Two treatment arms with 114 patients each (228 total)
- ▶ Study duration of 104 weeks
- ▶ Five covariates:
 - ▶ performance status: ranging from 70% to 100%
 - ▶ cancer stage: 3 or 4
 - ▶ race: black, white, other
 - ▶ gender
 - ▶ age: ranging from 31 to 82
- ▶ We use the methods from the simulations with value function based on $\log(T)$

NSCLC Trial Data: Results

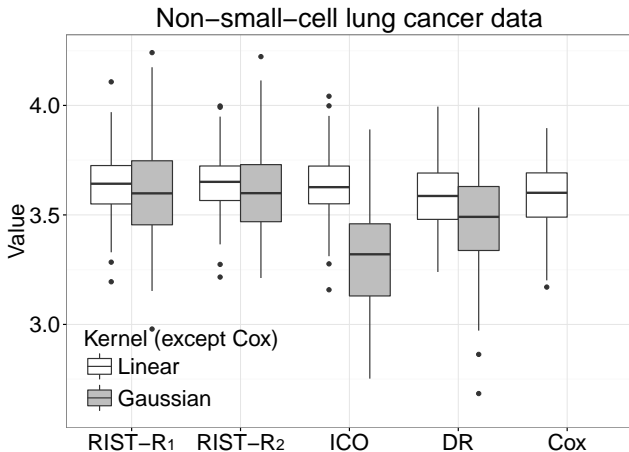


Figure 4: Boxplots of cross-validated (CV) value of survival weeks on the log scale. Four-fold CV randomly repeated 100 times.

Overall Results for Tree Based Approach

- ▶ The proposed outcome-weighted learning using RIST- R_2 works well over a range of right-censored survival settings.
- ▶ The idea of using random forests to impute before downstream analyses seems beneficial.
- ▶ There remain several important open research questions about inference in this context.

Outline

Introduction

Outcome weighted learning

A Tree Based Approach for Censored Data

Dose finding

Precision medicine in mHealth

Overall conclusions and future work

Motivating example

Bronchopulmonary Dysplasia

The most common morbidity associated with premature birth is bronchopulmonary dysplasia (BPD).

BPD $\xrightarrow{20\%}$ pulmonary arterial hypertension $\xrightarrow{40\%}$ die.

Sildenafil

Sildenafil is a potent inhibitor of type 5 phosphodiesterase.

Sildenafil has been used for adults with pulmonary arterial hypertension.

Challenge: Researchers want to know the best dose of sildenafil for premature infants with BPD as a function of infant characteristics.

Traditional dose trial for sildenafil

- ▶ 100 patients randomized to:

(mg/kg/dose)	0	0.75	1.50	2.25	3
Rate of developing hypertension	20%	15%	10%	5%	10%

- ▶ But 2.25mg/kg/dose may not be optimal for all babies:
 - ▶ a very early preterm baby with moderate BPD: 1.5 mg/kg/dose.
 - ▶ an older baby with severe BPD : 5 mg/kg/dose.

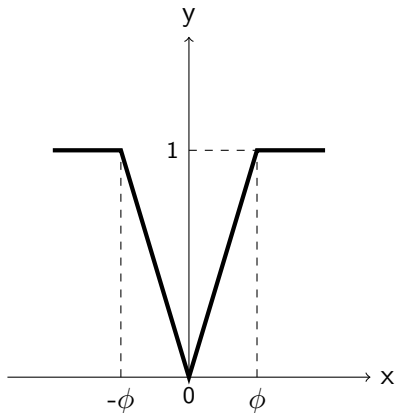
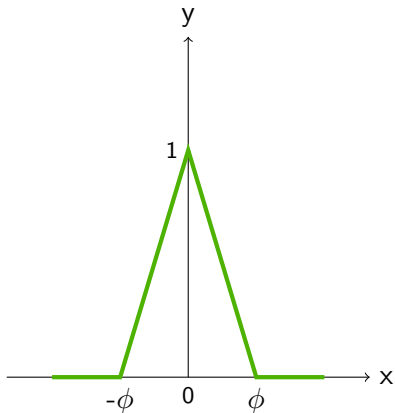
What is the challenge?

- ▶ For continuous dose $P(A = f(X)) = 0$.
- ▶ Thus O-learning cannot be applied directly.
- ▶ Solution: Chen G, Zeng D, and Kosorok MR (2016). Personalized dose finding using outcome weighted learning (with discussion and rejoinder). *JASA* 111:1509-1547.
- ▶ Note that maximizing the value function is approximately equivalent to minimizing:

$$\mathcal{R}_\phi(f) = E \left(\frac{R \ell_\phi(A - f(X))}{\phi p(A|X)} \right),$$

where $\ell_\phi(A - f(X)) = \min(|A - f(X)|/\phi, 1)$.

The l_ϕ Loss



- ▶ Connection to Kernel estimator: Triangular Kernel (left).
- ▶ l_ϕ is a bounded loss and brings robustness.

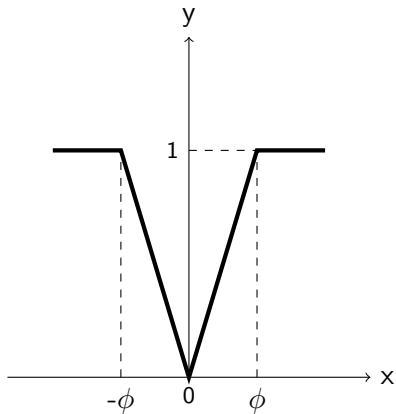
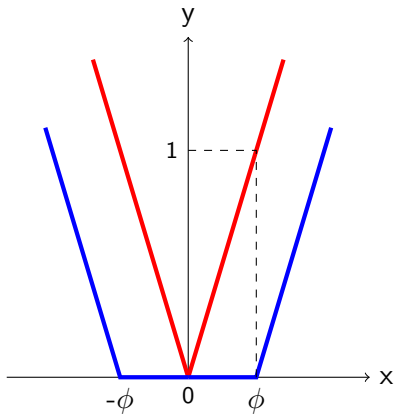
Outcome weighted learning (O-Learning)

Objective Function: Loss + Penalty

$$\min_f \left\{ \frac{1}{n} \sum_{i=1}^n \frac{R_i \ell_\phi(A_i - f(X_i))}{2\phi p(A_i|X_i)} + \lambda_n \|f\|^2 \right\}. \quad (3)$$

- ▶ $\|f\|$ is some norm for f , and λ_n controls the penalty.
- ▶ A linear decision rule: $f(X) = X^T \mathbf{w} + b$, with $\|f\|$ being the Euclidean norm of \mathbf{w} .
- ▶ The kernel trick generalizes to nonlinear/nonparametric rules.
- ▶ **Estimated individualized treatment rule:** \hat{f}_n solves (3).
- ▶ (2) is not convex but is the difference of two convex functions, so we use the DC algorithm.

l_ϕ as D.C.



Results for Dose Finding

- ▶ Consistent, with convergence rate able to get close to $n^{-1/4}$: can be strengthened to nearly $n^{-1/2}$ under a modified kernel and additional assumptions (Luedtke and van der Laan, 2016).
- ▶ In contrast, unbounded loss functions such as the absolute deviation loss will not yield consistent result.
- ▶ Generalizes to observational data via inverse propensity score weighting.
- ▶ Good simulation results: improves over other methods, especially indirect methods.
- ▶ Works well when applied to a Warfarin dosing data set.

Outline

Introduction

Outcome weighted learning

A Tree Based Approach for Censored Data

Dose finding

Precision medicine in mHealth

Overall conclusions and future work

Precision medicine in mHealth

Overall research goal:

- ▶ Develop estimation techniques (using data collected with mobile devices) for dynamic treatment regimes (which can be implemented as personalized mHealth interventions)

Motivating example: type 1 diabetes

- ▶ Understand type 1 diabetes (T1D) and how it is managed
- ▶ Develop tailored mHealth interventions for T1D management

What is T1D?

- ▶ Autoimmune disease wherein the pancreas produces insufficient insulin
- ▶ Daily regimen of monitoring glucose and replacing insulin
- ▶ Hypo- and hyperglycemia result from poor management
- ▶ Glucose levels affected by insulin, diet, and physical activity

The glucose-insulin dynamical system

A day in the life of a T1D patient:

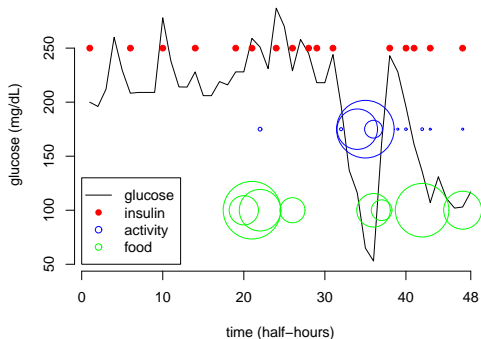


Figure 5: Plot of glucose, insulin, physical activity, and food intake.

Mobile technology in T1D care

Mobile devices can be used to administer treatment and assist with data collection in an outpatient setting, including

- ▶ Continuous glucose monitoring
- ▶ Accelerometers to track physical activity
- ▶ Insulin pumps to administer and log injections automatically

These technologies can be incorporated using mobile phones.

Research goals

Methodological goals:

- ▶ Estimate dynamic treatment regimes for use in mobile health
- ▶ Infinite time horizon, minimal modeling assumptions
- ▶ Observational data with minute-by-minute observations
- ▶ Online estimation to facilitate real-time decision making

Clinical goals:

- ▶ Provide patients information on the best actions to stabilize glucose
- ▶ Recommendations that are dynamic and personalized to the patient

Conceptual framework

- ▶ We use a Markov decision process (MDP) context
- ▶ One potential approach is to use infinite horizon Q-learning (models state-value as a function of action assuming all future actions are optimal):
 - ▶ Ertefaie A (2014). Constructing dynamic treatment regimes in infinite-horizon settings. *arXiv preprint arXiv:1406.0764*.
- ▶ We developed V-learning which uses a policy learning approach (models state-value as a function of policy):
 - ▶ Lockett DJ, Laber EB, Kahkoska AR, Maahs DM, Mayer-Davis E, Kosorok MR (2016). Estimating dynamic treatment regimes in mobile health using V-learning. *arXiv preprint arXiv:1611.03531*.

Summary for V-Learning

- ▶ Advantages of V-learning include
 - ▶ Flexibility in choosing a model for $V(\pi, \mathbf{s}; \theta^\pi)$
 - ▶ Online estimation, randomized decision rules
 - ▶ Flexibility in specifying reference distribution
 - ▶ Parametric value estimates
- ▶ A tailored treatment regime delivered through mobile devices may help to reduce hypo- and hyperglycemia in T1D patients

Outline

Introduction

Outcome weighted learning

A Tree Based Approach for Censored Data

Dose finding

Precision medicine in mHealth

Overall conclusions and future work

Overall Conclusions and Future Work

- ▶ This is an exciting time for precision medicine at the confluence of machine learning, Big Data, and statistics.
- ▶ Outcome weighted learning, or O-learning, V-learning and related methods, bridge between machine learning and precision medicine.
- ▶ There are numerous open questions.
- ▶ This work is part of the emergence of a new (or renewed) discipline focused on data driven decision making and precision medicine, and statistics is playing a key role.