Adaptive Semi-Supervised Inference for Optimal Treatment Decisions with Electronic Medical Record Data

Kevin Gunn, Wenbin Lu and Rui Song*

Abstract

A treatment decision is a rule that assigns a treatment to a patient based on their observed clinical information. The treatment decision that yields the greatest overall expected clinical benefit to the entire patient population is called the optimal treatment decision. We consider estimation of the optimal treatment decision within the restricted class of linear decision rules under semi-supervised settings, where the data consists of a set of 'labeled' patients and a much larger set of 'unlabeled' patients. This paper proposes an imputation-based semi-supervised method, SQ-SNP, utilizing 'unlabeled' individuals into the linear decision rule to offer a more efficient estimator. An inference procedure and results for asymptotic normality and consistency are provided. We present simulation studies to assess its performance relative to a fully supervised method and the effects of misspecification for the proposed linear decision rule. Afterward, application to an electronic medical records (EMR) study on the treatment of hypotensive episodes during intensive care unit (ICU) stays is discussed.

Keywords: Electronic medical records data; Kernel regression; Optimal treatment decision; Semi-supervised learning.

*Kevin Gunn is graduate student, Wenbin Lu is Professor (E-mail: lu@stat.ncsu.edu), and Rui Song is Associate Professor, Department of Statistics, North Carolina State University, Raleigh, NC 27695.

1 Introduction

Precision medicine, which is focused on creating treatment decisions for a patient based on his/her clinical information, has earned considerable interest. A treatment decision is a rule which, given a patient's observed clinical information, assigns a treatment to the patient. The objective is to choose the treatment, from the set of all possible treatments, that maximizes the patient's expected outcome. This creates a treatment tailored to each patient in the population of interest. The treatment decision that yields the greatest overall expected clinical benefit to the entire patient population is called the optimal treatment decision (OTD).

A recent area of interest in optimal treatment decision estimation has been the utilization of Electronic Medical Record (EMR) data. These data allow researchers to explore optimal treatment decisions in specific clinical scenarios not feasible in a randomized clinical trial (RCT) such as patients suffering from sepsis (Raghu et al. 2017). EMR data can also provide guidance where there is little previous research conducted, including second line treatment choices for type 2 diabetes episodes (Wang et al. 2016). Personalized medicine is particularly suited to the ICU due to patient heterogeneity among those treated for diseases like sepsis and acute kidney injury where a "one-size-fits-all" approach is not appropriate (Maslove et al. 2017). The recent shift in hospitals and other healthcare organizations to store clinical information in EMRs provide investigators with an immense amount of detailed information on the health status and clinical process of patients during interactions with their healthcare system. With databases such as MIMIC-III, an openly available critical care data set, these data are more readily accessible (Johnson et al. 2016). MIMIC-III encompasses medical record chart data recorded by clinicians such as; demographics, medical diagnoses, medications, vital signs, laboratory measurements (obtained in and out of the hospital), progress notes, imaging results, and billing information. The wealth of detailed clinical data provides the opportunity to enhance clinical decisions for accurate personalized medicine.

There is a great deal of work on statistical techniques to estimate the OTD from RCT data or observational study data under fully supervised settings, where a single decision

or a series of sequential decisions may be of interest (Murphy 2003; Moodie et al. 2007; Robins 2004; Zhao et al. 2012; Zhang et al. 2012). The purpose of this paper is to address the estimation of the OTD under an assumed linear *working* model, where the majority of individuals are missing response or treatment information, i.e., 'unlabeled' subjects. In other words, a semi-supervised learning (SSL) problem. In SSL, the knowledge gained through $\mathbb{P}_{\mathbf{X}}$ (distribution of $\mathbf{X}$) from the 'unlabeled' individuals is incorporated to improve inference on $\mathbb{P}_{\mathbf{Y}|\mathbf{X}}$ (Chapelle et al. 2006; Zhu and Goldberg 2009). In the OTD setting, we extend this to include treatment information, $\mathbf{A}$, into the inference procedure by trying to estimate $\mathbb{P}_{\mathbf{Y}|\mathbf{X},\mathbf{A}}$ using $\mathbb{P}_{\mathbf{A}|\mathbf{X}}$ and $\mathbb{P}_{\mathbf{X}}$. The proposed semi-supervised (SS) method is a three-step semi-nonparametric estimator to find the target parameter, $\boldsymbol{\beta}$, from the assumed linear *working* model. We combine SSL with Q-learning (Schulte et al. 2014; Nahum-Shani et al. 2012; Watkins and Dayan 1992) to estimate the OTD at a single decision point. Q-learning posits a model, denoted as $Q(\mathbf{X}, A)$, for the outcome of interest given the subject's information for each available treatment. The OTD finds the treatment that maximizes the Q-function, $Q(\mathbf{X}, A)$. The proposed SS Q-learning method is an imputation based semi-nonparametric estimator (SQ-SNP) based on kernel regression that is flexible to the underlying distribution, $\mathbb{P}_{\mathbf{Y}|\mathbf{A},\mathbf{X}}$. SQ-SNP utilizes all the available information on $\mathbb{P}_{\mathbf{X}}$ to provide a more robust estimator of parameters of interest that depend on $\mathbf{X}$. The 'unlabeled' subjects have their outcome imputed for each treatment, followed by a linear regression of the contrast among imputed outcomes under each treatment against $\mathbf{X}$ to obtain the OTD. SQ-SNP's incorporation of all available data into the estimation procedure intends to reduce the bias and the standard error for the parameters of interest.

Our work also discusses a fully supervised linear estimator, titled transformed response ordinary least squares (TR-OLS), for estimation of the OTD. TR-OLS regresses an inverse propensity score weighted estimator of the contrast among treatments on $\mathbf{X}$ to directly estimate $\boldsymbol{\beta}$. The focus of both approaches is on providing linear decision rules in low-dimensional settings due to their simple interpretations. Our method, SQ-SNP, is intended to be adaptive to the true underlying distribution making it robust to model misspecification. The TR-OLS estimator in comparison with SQ-SNP measures the benefit of including

the covariate information from the 'unlabeled' subjects into the estimation procedure for the OTD.

This article is organized as follows; Section 2 provides an overview of previous work in semi-supervised learning for EMR data and a situation where SSL optimal treatment decision making is applicable. Section 3 discusses the assumptions and framework for optimal treatment decisions. The estimation procedure and asymptotic theory results are presented in Section 4 for TR-OLS and Section 5 for SQ-SNP, respectively. Section 6 provides simulation studies to compare the empirical performances of both methods. Section 7 demonstrates an application to a study of patients in the intensive care unit (ICU) undergoing a hypotensive episode. Afterward, a discussion about SQ-SNP and future work within the area of optimal treatment decision estimation with EMR data is given.

2 Semi-Supervised Learning in Electronic Medical Records

Secondary analysis of EMR data can serve as an important tool in improving patient outcomes, but a major challenge is the lack of complete records for a patient's hospital stay (Weiskopf et al. 2013; Beaulieu-Jones et al. 2018). When the amount of missing information far exceeds the amount of available information, semi-supervised learning techniques have been used for prediction and inference with EMR data. Chakrabortty and Cai (2018) discuss a SS technique to improve estimation efficiency for parameters in a linear model, and apply it to an EMR study to measure the value of a biomarker commonly used to determine subtypes of rheumatoid arthritis. Beaulieu-Jones et al. (2016), Garla et al. (2013) and Hong et al. (2019) present semi-supervised methods for more accurate phenotype prediction when only a small subset of the patients underwent a manual chart review by a physician. Our work continues the interest of EMR data in semi-supervised settings by creating a SS Q-learning algorithm to provide an OTD rule in a study of IV fluid resuscitation and vasoactive therapy for hypotensive patients in MIMIC-III.

Given the situation where a large portion of the outcome of interest and treatment information is unavailable, estimation of the OTD is a challenging problem. Lee et al. (2012) perform a retrospective study on patients in MIMIC-II that underwent a single

hypotensive episode during their ICU stay to examine outcomes of patients receiving IV fluid resuscitation and vasoactive therapy. In their study, a majority of subjects received neither treatment option, despite both treatments typically used as a first line of therapy against hypotension (Kellum et al. 2012). Additionally, serum creatinine was a common contributor to missing data within this patient population, even though it is a marker for kidney damage, a concern for patients with hypotension. Our method will focus on overcoming the challenges of estimating of the OTD when missingness occurs in the response or treatment for a majority of subjects undergoing a hypotensive episode in the ICU.

3 Framework and Assumptions

Treatments received by patients will be denoted $A \in \mathcal{A}$, where $\mathcal{A}$ is the set of possible treatments. Let $\mathbf{X} \in \mathcal{X} \subseteq \mathbb{R}^p$ be a vector of subject characteristics ascertained prior to treatment. We will further assume $\mathcal{X}$ is a compact set, and $Var(\mathbf{X})$ is positive definite. Let $Y \in \mathbb{R}$ denote the observed response variable of interest and consider a larger value of Y as the better outcome. Suppose we have a fully 'observed' subject group, $\mathcal{O} = [\mathcal{O}_i = \{Y_i, A_i, \mathbf{X}_i\} : i = 1, ...n \text{ are i.i.d.}]$, and an 'unobserved' subject group where the treatment and/or response are not available, $\mathcal{U} = [\mathbf{X}_j : j = n+1, ...N \text{ are i.i.d.}]$, where $N >> n$. Assume $\mathcal{O}$ has finite 2^{nd} moments and all observations, $\mathcal{O} \cup \mathcal{U}$, are drawn from the same underlying distribution, $\mathbb{P} = \left(\mathbb{P}_{\mathbf{Y}|\mathbf{A},\mathbf{X}}, \mathbb{P}_{\mathbf{A}|\mathbf{X}}, \mathbb{P}_{\mathbf{X}}\right)$, i.e., the response and treatment information is missing completely at random (MCAR). The mathematical formulation of the optimal treatment decision will be defined to clarify the estimation procedure.

Consider a disease with two treatment options, each patient is assigned one treatment and the patient's response to that treatment is observed, but it is unknown how the patient would respond to the other choice of treatment. This motivates the concept of *potential* responses to treatment. The *potential* response based on each treatment is the response value that would be observed if a patient were given treatment 1 or 0, respectively. It will be denoted as $Y^*(1)$ and $Y^*(0)$. Assume the patient's response to the given treatment is constructed as, $Y_i = A_i Y^*(1) + (1 - A_i)Y^*(0)$ (Rubin 1978). This assumption is referred to as the *consistency* assumption, which states the observed response is equal to the potential

response under the treatment actually received by the patient. Furthermore, it is assumed that there are *no unmeasured confounders*, i.e., $(Y^*(1), Y^*(0)) \perp A \mid \mathbf{X}$ (Rosenbaum and Rubin 1983). The last assumption made is the *positivity* assumption, which states, the probability of receiving a treatment $a \in \mathcal{A}$ is greater than zero. Let $\pi(\mathbf{X})$ denote the propensity score, i.e. $\pi(\mathbf{X}) = P(A = 1 \mid \mathbf{X})$. It holds that $0 < \pi(\mathbf{X}) < 1$ by the *positivity* assumption.

Formally, a treatment decision is a rule that maps the covariates onto the treatment space, $\delta : \mathcal{X} \to \mathcal{A}$. A decision rule is implemented as a treatment, $Y^*(\delta) = \delta(\mathbf{X})Y^*(1) + (1 - \delta(\mathbf{X}))Y^*(0)$. The optimal treatment decision is denoted as δ^{opt}, and is defined as $\delta^{opt} = \arg\max_{\delta \in \Delta} E\left[Y^*(\delta)\right]$, where Δ is the class of all treatment decisions. Under this framework, $E\left[Y^*(\delta)\right] = E_{\mathbf{X}}\left\{\delta(\mathbf{X})E\left[Y \mid \mathbf{X}, A = 1\right] + (1 - \delta(\mathbf{X}))E\left[Y \mid \mathbf{X}, A = 0\right]\right\}$. Hence, the OTD becomes, $\delta^{opt} = I\left(E\left[Y \mid \mathbf{X}, A = 1\right] - E\left[Y \mid \mathbf{X}, A = 0\right] \geq 0\right)$.

Suppose the following model for the rest of the article, $E\left[Y \mid \mathbf{X}, A\right] = \mu(\mathbf{X}) + AC(\mathbf{X})$. The function, $\mu(\mathbf{X})$, represents the baseline effects of $\mathbf{X}$ on Y, and the contrast function, $C(\mathbf{X}) = E\left[Y \mid \mathbf{X}, A = 1\right] - E\left[Y \mid \mathbf{X}, A = 0\right]$, is the expected difference between treatments on the individual given his/her covariates. We further consider a linear contrast function as our assumed *working* model, i.e. $E\left[Y \mid \mathbf{X}, A\right] = \mu(\mathbf{X}) + A(\boldsymbol{\beta}'\widetilde{\mathbf{X}})$, where $\widetilde{\mathbf{X}} = (1, \mathbf{X}')'$. The OTD, δ^{opt}, arising from a posited linear model with parameter, $\boldsymbol{\beta}$, will belong to the class of estimators, $\Delta_{\boldsymbol{\beta}}$, and be denoted as $\delta^{opt}_{\boldsymbol{\beta}}$. Our estimator for the OTD within $\Delta_{\boldsymbol{\beta}}$ can be deduced to the function, $\widehat{\delta}^{opt}_{\boldsymbol{\beta}}(\mathbf{X}) = I(\widehat{\boldsymbol{\beta}}'\widetilde{\mathbf{X}} > 0)$, where $\widehat{\boldsymbol{\beta}}$ is the solution to the normal equations. This class of estimators is chosen to produce decision rules that are easily interpretable due to their simplicity.

4 Transformed Response Ordinary Least Squares

The OTD is a function of $C(\mathbf{X})$, and as a result it is not necessary to estimate the baseline effect, $\mu(\mathbf{X})$. To reduce the possibility of model misspecification, we estimate $\delta^{opt}_{\boldsymbol{\beta}}$ without fully specifying $E\left[Y \mid \mathbf{X}, A = a\right]$ for $a \in \{0, 1\}$. The transformed response ordinary least squares (TR-OLS) method regresses $C(\mathbf{X})$ directly onto the covariates to provide a fully supervised linear decision rule for the OTD.

TR-OLS uses an inverse propensity weighted (IPW) estimator of $C(\mathbf{X})$ as responses in a linear model. An unbiased IPW estimator for $C(\mathbf{X})$ is,

$$\widetilde{Y} = \frac{Y(A - \pi(\mathbf{X}))}{\pi(\mathbf{X})(1 - \pi(\mathbf{X}))}.$$

Therefore, the transformed response model is, $\widetilde{Y} = C(\mathbf{X}) + \epsilon$, whereby under the assumed *working* model $C(\mathbf{X})$ is equivalent to $\boldsymbol{\beta}'\widetilde{\mathbf{X}}$ and $E[\epsilon \mid \mathbf{X}] = 0$. The least squares is minimized to obtain $\widehat{\boldsymbol{\beta}}_{TR}$, the regression parameters for the observed cases,

$$\widehat{\boldsymbol{\beta}}_{TR} = \arg\min_{\beta} \sum_{i=1}^{n} \left[\widetilde{Y}_i - \boldsymbol{\beta}'\widetilde{\mathbf{X}}_i\right]^2. \tag{1}$$

Under the assumptions given in section 3 and as $n \to \infty$,

$$\sqrt{n}\left(\widehat{\boldsymbol{\beta}}_{TR} - \boldsymbol{\beta}\right) = n^{-1/2} \sum_{i=1}^{n} \Psi_{TR}(\mathbf{O}_i) + o_p(1) \xrightarrow{d} N_{p+1}\left(0, \mathcal{V}_{\boldsymbol{\beta}_{TR}}\right), \tag{2}$$

where $\Psi_{TR}(\mathbf{O}) = \Lambda^{-1}\widetilde{\mathbf{X}}\left(\widetilde{Y} - \widehat{\boldsymbol{\beta}}'_{TR}\widetilde{\mathbf{X}}\right)$ and $\Lambda = E[\widetilde{\mathbf{X}}\widetilde{\mathbf{X}}']$. The asymptotic variance, $\mathcal{V}_{\boldsymbol{\beta}_{TR}}$, is equivalent to $E[\Psi_{TR}(\mathbf{O})\Psi'_{TR}(\mathbf{O})]$. Consistent estimators for $\mathcal{V}_{\boldsymbol{\beta}_{TR}}$ and Λ are given by $\widehat{\mathcal{V}}_{\boldsymbol{\beta}_{TR}} = \frac{1}{n}\sum_{i=1}^{n} \Psi_{TR}(\mathbf{O}_i)\Psi'_{TR}(\mathbf{O}_i)$ and $\Lambda_n = n^{-1}\sum_{i=1}^{n} \widetilde{\mathbf{X}}_i\widetilde{\mathbf{X}}'_i$, respectively.

The supervised estimator, $\widehat{\boldsymbol{\beta}}_{TR}$, uses data only from $\mathcal{O}$. By including the data from $\mathcal{U}$, the SS estimator intends to create a more efficient estimator for $\boldsymbol{\beta}$ relative to $\widehat{\boldsymbol{\beta}}_{TR}$. The parameter, $\widehat{\boldsymbol{\beta}}_{TR}$, gauges the improvement offered by incorporating $\mathcal{U}$ into the estimation procedure discussed in Sections 5.1 and 5.2. An increase in estimation efficiency for the parameters, $\boldsymbol{\beta}$, in the linear decision rule employed by the OTD is the primary goal, especially under model misspecification, as it signifies a more robust estimate of $\delta^{opt}_{\boldsymbol{\beta}}$. A discussion of Q-learning for a single decision point is needed to establish the difference in the SSL approach.

5 Semi-Supervised Q-Learning for a Single Decision Point

Q-learning at a single decision point aims to estimate the expected outcome of interest conditioned on treatment and covariates, $Q(\mathbf{x}, a) = E[Y \mid \mathbf{X} = \mathbf{x}, A = a]$. In the context of optimal treatment decision making, Q-functions typically model $Q(\mathbf{x}, a)$ using all the observed data, $\mathcal{O}$, with a parametric or semi-parametric estimator, $Q(\mathbf{x}, a, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a finite-dimensional parameter (Schulte et al. 2014). Using the estimator, $Q(\mathbf{x}, a, \boldsymbol{\theta})$, the OTD is defined as $I(Q(\mathbf{x}, 1, \boldsymbol{\theta}) - Q(\mathbf{x}, 0, \boldsymbol{\theta}) \geq 0)$ assuming larger Y is preferred. Given the assumed *working* model, the Q-functions are expressed as, $Q(\mathbf{x}, a) = \mu(\mathbf{x}) + a(\boldsymbol{\beta}'\widetilde{\mathbf{x}})$, and the OTD is equivalent to $\delta^{opt}_{\boldsymbol{\beta}}(\mathbf{x}) = I(\boldsymbol{\beta}'\widetilde{\mathbf{x}} \geq 0)$. If the data available consists of sets $\mathcal{O}$ and $\mathcal{U}$, then Q-functions should use all available data to make treatment decisions. We propose a three-step SS imputation procedure as follows: (i) impute $Q(\mathbf{X}_j, a)$ for $a \in \{0, 1\}$ and $\forall j = n+1, ..., N$, (ii) compute the imputed contrast functions, $C(\mathbf{X}_j) = Q(\mathbf{X}_j, 1) - Q(\mathbf{X}_j, 0)$ for $\forall j = n+1, ..., N$, (iii) regress the imputed contrast functions on $\mathbf{X}$ to obtain $\widehat{\boldsymbol{\beta}}$ for $\widehat{\delta}^{opt}_{\boldsymbol{\beta}}$.

5.1 Fully Nonparametric Imputation of the Q-Function

If all the Y and A in $\mathcal{U}$ were actually observed, we could obtain an estimate for $\boldsymbol{\beta}$ utilizing the entire data set with the TR-OLS approach discussed in Section 4. Instead, we must use a different approach. We present a fully non-parametric estimator based on kernel regression (KR). KR is a locally weighted average function with no parametric assumptions. The kernel function is denoted as $W(\cdot)$ with $W : \mathbb{R}^p \to \mathbb{R}$. The bandwidth, h, is a function of n and greater than zero. In practice, h is often estimated using $\mathcal{K}$-fold cross validation. Q-functions as KR estimators are defined as,

$$Q^{(np)}(\mathbf{x}, 1) = \frac{\sum_{i=1}^{n} W\left(\frac{\mathbf{x}-\mathbf{X_i}}{h}\right) A_i Y_i}{\sum_{i=1}^{n} W\left(\frac{\mathbf{x}-\mathbf{X_i}}{h}\right) A_i} \tag{3}$$

and,

$$Q^{(np)}(\mathbf{x}, 0) = \frac{\sum_{i=1}^{n} W\left(\frac{\mathbf{x}-\mathbf{X_i}}{h}\right) (1 - A_i) Y_i}{\sum_{i=1}^{n} W\left(\frac{\mathbf{x}-\mathbf{X_i}}{h}\right) (1 - A_i)}. \tag{4}$$

The first step is to impute both $Q^{(np)}(\mathbf{x},0)$ and $Q^{(np)}(\mathbf{x},1)$, followed by the computation of $\widehat{C}^{(np)}(\mathbf{x})$. The set $\mathcal{U}$ with imputed data is $\left[\left\{\widehat{C}^{(np)}(\mathbf{X}_j),\mathbf{X}_j\right\} : j=n+1,...,N\right]$, and the solution the normal equations is obtained by minimizing the argument,

$$\widehat{\boldsymbol{\beta}}_{np} = \arg\min_{\beta} \sum_{j=n+1}^{N} \left[\widehat{C}^{(np)}(\mathbf{X}_j) - \boldsymbol{\beta}'\widetilde{\mathbf{X}}_j\right]^2. \tag{5}$$

The new nonparametric SS linear decision rule is $\widehat{\delta}^{opt}_{\boldsymbol{\beta}_{np}} = I(\boldsymbol{\beta}'_{np}\widetilde{\mathbf{X}} > 0)$, and utilizes $\mathbb{P}_{\mathbf{X}}$ from $\mathcal{U}$.

The subsequent assumptions are needed for Theorem 1 given below. Most of these assumptions are fairly standard (Fan 1992; Hansen 2008; Newey 1994), but are adapted slightly to the framework for OTDs. (1) $h \equiv h(n) = o(1)$, (2) $Q^{(np)}(\mathbf{x},a)$, $\pi(\mathbf{x})$, and $f(\cdot)$ are r times continuously differentiable with bounded r^{th} derivatives on some open set within $\mathcal{X}$. (3) $W(\cdot)$ is a symmetric r^{th} order kernel for some integer $r \geq 2$. $W(\cdot)$ is Lipschitz continuous and has bounded support that it shares with $\pi(\cdot)$, $\mathcal{W} \subseteq \mathbb{R}^p$. (4) $E(|AY|^s) < \infty$ and $E(|(1-A)Y|^s) < \infty$ for some $s > 2$. (5) $E(|AY|^s \mid \mathbf{X}=\mathbf{x})f(\mathbf{x}) = E(|Y|^s \mid \mathbf{X}=\mathbf{x})\pi(\mathbf{x})f(\mathbf{x})$ and $f(\mathbf{x})$ are bounded on $\mathcal{X}$ and $\inf_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}) > 0$.

Theorem 1. *Suppose* $\mathrm{n}^{1/2}\mathrm{h}^{\mathrm{r}} \to 0$ *and* $\sqrt{\frac{\ln n}{nh^p}} \to 0$ *as* $n \to \infty$ *and* $N >> n$ *such that* $n/N \to 0$. *Then, under the required assumptions (1)-(5),*

$$n^{1/2}\left(\widehat{\boldsymbol{\beta}}_{np} - \boldsymbol{\beta}\right) = n^{-1/2}\sum_{i=1}^{n}\Psi_{np}(\mathbf{O}_i) + o_p(1) \xrightarrow{d} N_{p+1}\left(0, \mathcal{V}_{\boldsymbol{\beta}_{np}}\right)$$

where $\Psi_{np}(\mathbf{O}) = \left\{\frac{A}{\pi(\mathbf{X})} - \frac{1-A}{1-\pi(\mathbf{X})}\right\}\Lambda^{-1}\widetilde{\mathbf{X}}\left[Y - Q^{(np)}(\mathbf{X},A)\right]$ *and asymptotic variance,* $\mathcal{V}_{\boldsymbol{\beta}_{np}} = E\left[\Psi_{np}(\mathbf{O})\,\Psi'_{np}(\mathbf{O})\right]$.

There are drawbacks to using KR owing to the *curse of dimensionality* and inherent *finite sample bias*. Fan (1992) discusses *finite sample bias* for KR has order of h^2 for 2^{nd} order kernels, and through Taylor expansion it is clear the bias has order of h^r for r^{th} order kernels. For this article, the goal is developing an efficient low-dimensional linear decision rule, so the *curse of dimensionality* is not of concern. However, if the sample size within $\mathcal{O}$ is not large enough, $\boldsymbol{\beta}_{np}$ is a biased estimator. In Section 5.2, we discuss a cross-validation

technique to correct for bias with a linear term by regressing the residuals, $Y - Q^{(np)}(\mathbf{x}, a)$, against the covariates.

5.2 Semi-Nonparametric Imputation of the Q-Function

We overcome the issue of *finite sample bias* in nonparametric estimation by adapting the work of Chakrabortty and Cai (2018) on semi-nonparametric imputation based estimators into the SS OTD framework. The new semi-nonparametric (SQ-SNP) estimator, $Q^{(SS)}(\mathbf{x}, a, \boldsymbol{\theta}_a)$, builds upon the estimator in Section 5.1, but produces a more efficient estimator for $\boldsymbol{\beta}$ by reducing bias. SQ-SNP achieves this through a *linear refitting* step.

SQ-SNP is composed of two steps, where the first step is to fit the KR estimator, $\widehat{Q}^{(np)}(\mathbf{x}, a)$, on the data in $\mathcal{O}$, and the second step (*linear refitting* step) corrects for the bias of $\widehat{Q}^{(np)}(\mathbf{x}, a)$ with a linear model. The *linear refitting* step applies an inverse propensity score weighted least squares regression of the residuals, $Y - Q^{(np)}(\mathbf{X}, a)$, on $\mathbf{X}$ to obtain estimates for the parameters of interest, $\boldsymbol{\theta}_a$ for $a \in \{0, 1\}$. Namely, $\boldsymbol{\theta}_a$ for $a \in \{0, 1\}$, are the solutions to minimizing the weighted least squares equations,

$$\widehat{\boldsymbol{\theta}}_1 = \arg\min_{\boldsymbol{\theta}_1} \sum_{k=1}^{\mathcal{K}} \sum_{i \in \mathcal{O}_k} \frac{A_i}{\pi(\mathbf{X}_i)} \left(Y_i - \widehat{Q}_k^{(np)}(\mathbf{X}_i, A_i) - \boldsymbol{\theta}_1' \widetilde{\mathbf{X}}_i \right)^2 \tag{6}$$

and,

$$\widehat{\boldsymbol{\theta}}_0 = \arg\min_{\boldsymbol{\theta}_0} \sum_{k=1}^{\mathcal{K}} \sum_{i \in \mathcal{O}_k} \frac{1 - A_i}{1 - \pi(\mathbf{X}_i)} \left(Y_i - \widehat{Q}_k^{(np)}(\mathbf{X}_i, A_i) - \boldsymbol{\theta}_0' \widetilde{\mathbf{X}}_i \right)^2 . \tag{7}$$

The *linear refitting* step utilizes $\mathcal{K}$-fold cross-validation to obtain an unbiased estimator of $\boldsymbol{\theta}_a$ for $a \in \{0, 1\}$. Cross-validation separates $\mathcal{O}$ into $\mathcal{K}$ partitions of equal size, where both treatment cohorts have their populations split equally among the folds. The partitions are denoted as $\mathcal{O}_k$ for $k \in \{1, 2, .., \mathcal{K}\}$. Cross-validation overcomes underestimation of the true residuals caused by over-fitting during the kernel regression step. The final imputation function is $Q^{(SS)}(\mathbf{x}, a, \boldsymbol{\theta}_a) = Q^{(np)}(\mathbf{x}, a) + \boldsymbol{\theta}_a' \mathbf{x}$, which is estimated as:

$$\widehat{Q}^{(SS)}(\mathbf{x}, a, \widehat{\boldsymbol{\theta}}_a) = \frac{1}{\mathcal{K}} \sum_{k=1}^{\mathcal{K}} \widehat{Q}_k^{(np)}(\mathbf{x}, a) + \widehat{\boldsymbol{\theta}}_a' \widetilde{\mathbf{x}} , \tag{8}$$

for $a \in \{0,1\}$. We compute $\widehat{C}^{(SS)}(\mathbf{x}) = \widehat{Q}^{(SS)}(\mathbf{x},1,\widehat{\boldsymbol{\theta}}_1) - \widehat{Q}^{(SS)}(\mathbf{x},0,\widehat{\boldsymbol{\theta}}_0)$, and obtain $\widehat{\boldsymbol{\beta}}_{SS}$ similarly to Equation 5. The OTD based on SQ-SNP is $\widehat{\delta}^{opt}_{\boldsymbol{\beta}_{SS}} = I(\widehat{\boldsymbol{\beta}}'_{SS}\widetilde{\mathbf{x}} > 0)$.

We need an additional condition for the influence function given in Theorem 2 to be unbiased and consistent: *condition* (1) let $\widehat{\mathbf{R}}_{k,a}(\widetilde{\mathbf{X}}) = \widetilde{\mathbf{X}}[\widehat{Q}^{(np)}_k(\mathbf{X},a) - Q^{(np)}(\mathbf{X},a)]$ and $\bar{\mathbf{R}}_{k,a}(\widetilde{\mathbf{X}}) = E[\widehat{\mathbf{R}}_{k,a}] - \widehat{\mathbf{R}}_{k,a}$, $\mathcal{R}_{n,\mathcal{K},a} = n^{-1/2}\sum_{k=1}^{\mathcal{K}}\sum_{i\in\mathcal{O}_k}\bar{\mathbf{R}}_{k,a}(\widetilde{\mathbf{X}}_i)$. Then for $\mathcal{K} \geq 2$, $\mathcal{R}_{n,\mathcal{K},a} = o_p(1)$, $\forall a \in \mathcal{A}$ by *Lemma A.1* in Chakrabortty and Cai (2018).

Theorem 2. *Suppose the KR estimator, $Q^{(np)}(\mathbf{X},a)$, satisfies assumptions (1) through (5) and condition (1) holds,*

$$n^{1/2}\left(\widehat{\boldsymbol{\beta}}_{SS} - \boldsymbol{\beta}\right) = n^{-1/2}\sum_{i=1}^{n}\Psi_{SS}(\mathbf{O}_i) + o_p(1) \xrightarrow{d} N_{p+1}\left(0, \mathcal{V}_{\boldsymbol{\beta}_{SS}}\right)$$

where $\Psi_{SS}(\mathbf{O}) = \left\{\frac{A}{\pi(\mathbf{X})} - \frac{1-A}{1-\pi(\mathbf{X})}\right\}\Lambda^{-1}\widetilde{\mathbf{X}}\left[Y - Q^{(SS)}(\mathbf{X},A,\boldsymbol{\theta}_A)\right]$ *and asymptotic variance,* $\mathcal{V}_{\boldsymbol{\beta}_{SS}} = E\left[\Psi_{SS}(\mathbf{O})\Psi'_{SS}(\mathbf{O})\right]$.

The proof of Theorems 1 and 2 are outlined in the Appendix.

In an observational study, $\pi(\mathbf{X})$ must be estimated in the TR-OLS estimator described in Section 4, and during the *linear refitting* step of our SQ-SNP method. A logistic regression model is often a reasonable choice in practice for estimation of $\pi(\mathbf{X})$. Section 6 demonstrates simulation studies that require estimation of $\pi(\mathbf{X})$ to understand how our method will perform in a setting similar to that encountered in EMR data.

5.3 Inference for Semi-Supervised *Q*-Functions

The SQ-SNP Q-functions, $\widehat{Q}^{(SS)}(\mathbf{x},a,\widehat{\boldsymbol{\theta}}_a)$, over-fit Y in $\mathcal{O}$ due to the *linear refitting* step. Due to this issue, $\widehat{\mathcal{V}}_{\boldsymbol{\beta}_{SS}} = n^{-1}\sum_{i=1}^{n}\Psi_{SS}(\mathbf{O}_i)\Psi'_{SS}(\mathbf{O}_i)$ underestimates $\mathcal{V}_{\boldsymbol{\beta}_{SS}}$ in practice. A 'double-CV' step is proposed to reduce bias during the variance estimation procedure (Chakrabortty and Cai 2018). We create $\mathcal{K}$ folds, $\mathcal{O}_k$ for $k \in \{1,..,\mathcal{K}\}$, then construct $\mathcal{K}$ distinct estimates of $\boldsymbol{\theta}_a$, $\left\{\boldsymbol{\theta}_{a,(k)} : k = 1,..,\mathcal{K}\right\}$. Let $\mathcal{O}^*_k$ represent the data not included in $\mathcal{O}_k$, where $\mathcal{O}^*_k \perp \mathcal{O}_k$ and $\mathcal{O}^*_k \bigcup \mathcal{O}_k = \mathcal{O}$. Fold $\mathcal{O}_k$ is not involved in estimating $\left\{\widehat{Q}^{(np)}(\mathbf{X}_i,A_i)\right\}_{i\in\mathcal{O}^*_k}$ as well as the parameter, $\boldsymbol{\theta}_{a,(k)}$, to mirror imputation of missing

response information in the SS setting. The cross-validated estimate $\boldsymbol{\theta}_{a,(k)}$ is the solution to:

$$\widehat{\boldsymbol{\theta}}_{1,(k)} = \arg\min_{\boldsymbol{\theta}_1} \sum_{i \notin \mathcal{O}_k} \frac{A_i}{\pi(\mathbf{X}_i)} \left(Y_i - \widehat{Q}_k^{(np)}\left(\mathbf{X}_i, A_i\right) - \boldsymbol{\theta}_1' \mathbf{X}_i \right)^2 \tag{9}$$

$\forall k \in \{1, ..., \mathcal{K}\}$, where $\boldsymbol{\theta}_{0,(k)}$ is estimated analogously. The 'double-CV' imputation function becomes, $\widehat{Q}_k^{(SS)}\left(\mathbf{x}, a, \boldsymbol{\theta}_{a,(k)}\right) = \widehat{Q}_k^{(np)}\left(\mathbf{x}, a\right) + \widehat{\boldsymbol{\theta}}_{a,(k)}' \widetilde{\mathbf{x}}$. We substitute $\widehat{Q}_k^{(SS)}\left(\mathbf{x}, a, \boldsymbol{\theta}_{a,(k)}\right)$ for $\widehat{Q}^{(SS)}\left(\mathbf{x}, a, \boldsymbol{\theta}_a\right)$ in the corresponding influence functions from Theorem 2. The estimators $\Lambda_N = N^{-1} \sum_{j=n+1}^{N} \widetilde{\mathbf{X}}_i \widetilde{\mathbf{X}}_i'$ and Λ_{n+N} are consistent estimators of Λ that use the available data in $\mathcal{U}$. The asymptotic variance, $\mathcal{V}_{\boldsymbol{\beta}_{SS}}$, has the consistent 'double-CV' estimator:

$$\widehat{\mathcal{V}}_{\boldsymbol{\beta}_{SS,(\mathcal{K})}} = \frac{1}{n} \sum_{i=1}^{n} \Psi_{SS,(k)}\left(\mathbf{O}_i\right) \Psi_{SS,(k)}'\left(\mathbf{O}_i\right), \tag{10}$$

which allows us to establish standard error and confidence intervals estimates for $\widehat{\boldsymbol{\beta}}_{SS}$.

6 Simulation Analysis

We conducted simulation studies to examine the finite-sample performance of SQ-SNP and TR-OLS on estimating the OTD. Our simulation study design considers varying degrees of model misspecification in $C(\mathbf{X})$ to better understand the advantages of SQ-SNP when estimating $\delta_{\boldsymbol{\beta}}^{opt}$. During each simulation replication, the propensity score will be estimated to better reflect the analysis of an observation study with EMR data. We present the percent of correct decisions (PCD) and value function to illustrate the benefits of semi-supervised prediction on assessing $\delta_{\boldsymbol{\beta}}^{opt}$. Furthermore, we study the bias, empirical standard error (ESE), asymptotic standard error (ASE), and the relative efficiency of $\widehat{\boldsymbol{\beta}}_{SS}$ with respect to $\widehat{\boldsymbol{\beta}}_{TR}$.

6.1 Data Generation

In this simulation, n is the amount of subjects in $\mathcal{O}$ and N is the amount of subjects in $\mathcal{U}$. The N unobserved responses were simulated to be missing completely at random

(MCAR), where the response and treatment are not observed. Let $\mathbf{X} \sim N(\mathbf{0}_p, I_p)$, where $\mathbf{X}$ is contained in $[-5,5]^p$ to ensure it is in a compact set, and $\mathcal{A} \in \{0,1\}$. The following models are tested with $p = 2$, $n = 500$ and $N = 5000$. We examine three choices for $C(\mathbf{X})$,

- *Model 1 (Linear)*: $\mathbf{Y} = \mu(\mathbf{X}) + \mathbf{A}(\boldsymbol{\eta}'\mathbf{X}) + \epsilon$
- *Model 2 (Cubic)*: $\mathbf{Y} = \mu(\mathbf{X}) + \mathbf{A}(\boldsymbol{\gamma}'\mathbf{X})^{\mathbf{3}} + \epsilon$
- *Model 3 (Sine)*: $\mathbf{Y} = \mu(\mathbf{X}) + \mathbf{A}(\sin(\boldsymbol{\eta}'\mathbf{X})) + \epsilon$,

where $\boldsymbol{\eta} = \mathbf{1_2}$, $\boldsymbol{\gamma} = (0.3, 0.6)'$, and $\epsilon \sim \mathcal{N}(0,1)$. Two baseline functions were studied. The first one simulates a situation where the baseline effect for subjects is small and the second function proposes a larger baseline effect,

1. $\mu(\mathbf{X}) = (\boldsymbol{\omega}'\mathbf{X})^3$,
2. $\mu(\mathbf{X}) = (\boldsymbol{\alpha}'\mathbf{X})(1 + \boldsymbol{\omega}'\mathbf{X})$.

We study two different baseline functions to understand the flexibility of SQ-SNP in estimating $C(\mathbf{X})$. The models were studied with $\boldsymbol{\omega} = (0.5, 0.5)'$ and $\boldsymbol{\alpha} = (0.75, 0.75)'$. We simulate a propensity score of $\pi(\mathbf{X}) = logit(0.5X_1 - 0.5X_2)$, and an allocation of treatment as $\mathbf{A} \sim bernoulli(p = \pi(\mathbf{X}))$. The propensity scores were generated to simulate observational data where the allocation of treatment depends on a patient's health status information. The true values, $\boldsymbol{\beta}_0$, are estimated by simulating a fully observed Monte Carlo data set of size 500,000.

6.2 Simulation Results

The results of interest are percent of correct decisions (PCD), value function (V), and component-wise relative efficiency (RE) over 500 replications. The component-wise RE for each estimator was calculated as $\sum_{i=1}^{500} \|\widehat{\boldsymbol{\beta}}_{TR-OLS,i,j} - \boldsymbol{\beta}_{0,j}\|^2 / \sum_{k=1}^{500} \|\widehat{\boldsymbol{\beta}}_{SQ-SNP,j} - \boldsymbol{\beta}_{0,j}\|^2$, where $j = 1, .., p+1$. During each replication, new vectors $\mathbf{Y}$, $\mathbf{A}$ and matrix $\mathbf{X}$ are simulated. The percent of correct decisions (PCD) was calculated for each method,

$$PCD_i = 1 - \frac{1}{5500} \sum_{k=1}^{5500} \left| I(\widehat{\boldsymbol{\beta}}'\widetilde{\mathbf{X}}_k > 0) - I(\boldsymbol{\beta}_{\mathbf{0}}'\widetilde{\mathbf{X}}_k > 0) \right|,$$

and then averaged over all 500 simulations, $PCD = (1/500)\sum_{i=1}^{500} PCD_i$. The true value function is calculated with the Monte Carlo data set of sample size 500,000 as,

$$V_0 = \frac{1}{500,000} \sum_{m=1}^{500,000} \left\{ \mu(\mathbf{X}_m) + I(C(\mathbf{X}_m) > 0) C(\mathbf{X}_m) \right\}.$$

Estimation of the value function using TR-OLS or SQ-SNP is calculated as,

$$\widehat{V}_i = \frac{1}{500,000} \sum_{m=1}^{500,000} \left\{ \mu(\mathbf{X}_m) + \widehat{\delta}^{opt}_{\boldsymbol{\beta}} C(\mathbf{X}_m) \right\},$$

on the same Monte Carlo data set and then averaged over 500 replications. Since we assume a larger response to treatment is preferred, a larger value function signifies an OTD that on average provides the best treatment allocation to the population of interest. A Gaussian kernel was chosen for $W(\cdot)$, and a value of $\mathcal{K} = 5$ was chosen for cross-validation. The value for h was estimated through cross-validation.

Table 1 summarizes the value function and PCD results under all 6 settings. Across all settings, we observe an improvement in the average value function, average PCD, and their respective empirical SE's while using the SQ-SNP method compared to TR-OLS. Table 2 presents the bias, empirical SE, asymptotic SE, component-wise RE, and component-wise coverage probabilities of the 95% confidence intervals for SQ-SNP and TR-OLS under both settings for $\mu(\mathbf{X})$. SQ-SNP demonstrates its ability to reduce SE of the parameters in the linear model regardless of the true underlying distribution of $\mathbb{P}_{\mathbf{Y}|\mathbf{A},\mathbf{X}}$.

The results suggest SQ-SNP is a substantially more efficient estimator for $\boldsymbol{\beta}$. Robust utilization of the data in $\mathcal{U}$ through SQ-SNP provides more information on $\mathbb{P}_{\mathbf{X}}$, allowing our method to efficiently estimate $\boldsymbol{\beta}$ under departures from a linear contrast function. Consequentially, it produces a more robust estimator of $\delta^{opt}_{\boldsymbol{\beta}}$ relative to TR-OLS. SQ-SNP outperforms TR-OLS under all 6 settings, but the OTD estimator, $\widehat{\delta}^{opt}_{\boldsymbol{\beta}_{SS}}$, obtained through SQ-SNP is particularly advantageous when model misspecification occurs according to Table 1.

7 Application to an EMR Study

We apply our proposed method, SQ-SNP, to an EMR study on patients undergoing a hypotensive episode in the ICU within the MIMIC-III database. It is important to treat hypotensive episodes (HE's) in ICU patients to minimize end-organ damage. A marker of end-organ damage is an increase in serum creatinine, post-hypotensive episode (Lehman et al. 2010). We consider a baseline serum creatinine measurement observed no earlier than 24 hours before the start of the episode. Subsequently, a post-episode serum creatinine measurement was measured no more than 72 hours after the HE. Two available treatments for HE's include IV fluid resuscitation and vasoactive therapy (Lee et al. 2012). Our objective is to determine the optimal initial treatment for each individual between IV fluid resuscitation and vasopressor interventions to minimize any increase in serum creatinine 72 hours post-hypotensive episode.

Lee et al. (2012) defines the initiation of an HE as the occurrence of two successive mean arterial pressure (MAP) measurement values > 60 *mm Hg*, followed by two MAP measurements ≤ 60 *mm Hg*. The end of an HE is characterized by two successive MAP measurements ≤ 60 *mm Hg*, followed by two successive MAP measurements > 60 *mm Hg*. The study cohort includes subjects greater than or equal to 15 years old who experienced a single hypotensive episodes during their ICU stay, which amounts to 3,372 subjects.

To formulate this problem into the OTD framework, we must define Y as the negative of the difference between post-HE and pre-HE serum creatinine measurements. The treatment space is defined as $A = 1$ if the patient was given vasopressor treatment, and $A = 0$ if they received IV fluid resuscitation. Vasopressor treatments include dobutamine, dopamine, epinephrine, norepinephrine, phenylephrine, milrinone or vasopressin during the HE. IV fluid treatment is defined as receiving at least one infusion of colloids, or isotonic crystalloids of at least 250 ml (Lee et al. 2012). The predictors, $\mathbf{X}$, include normalized versions of mean MAP and mean heart rate in the 3 hour window immediately before the HE onset. A total of 945 subjects received either IV fluid resuscitation or vasopressor treatment as their sole treatment. The number of patients untreated with solely IV fluid rescuitation or vasoactive therapy and/or missing response information resulted in 2,427 subjects. For this

data analysis, we are focusing on the semi-supervised setting, therefore we randomly choose $n = 300$ out of 945 subjects to create the fully observed set, $\mathcal{O}$, and place the remaining 3,072 subjects into $\mathcal{U}$.

Recall SSL relies on the assumption all individuals are drawn from $\mathbb{P}$, which means the distribution for each of the predictors should come from $\mathbb{P}_{\mathbf{X}}$ regardless of set, $\mathcal{O}$ or $\mathcal{U}$, the individual belongs in. Table 3 provides tests for equality between the distributions of $\mathbf{X}$ between the two data sets.

Propensity scores were modeled with a logistic regression model that included the following covariates; baseline creatinine, age, gender, indicator function for patients that received surgical services, log transformation of the van Walraven et al. (2009) Elixhauser comorbidity score (method of categorizing comorbidities of patients based on their ICD-9 codes), log transformation of simplified acute physiologic score (SAPS) (scoring system to reflect the risk of death upon admission to ICU) (Le Gall et al. 1984), total urine output in the 3 hour window prior to the HE, average mean heart rate in the 3 hour window prior to the HE, and average mean arterial pressure (MAP) in the 3 hour window prior to the HE. We perform TR-OLS and SQ-SNP on the data set to obtain estimates and standard errors for $\boldsymbol{\beta}$. A Gaussian kernel was chosen for $W(\cdot)$ in Equations 3 and 4 and $\mathcal{K} = 5$ similar to Section 6. Table 4 presents the point estimates, estimated SE, and p-values for testing null effects. SQ-SNP reduces the estimated ASE for each predictor.

In Table 5, we demonstrate treatment allocation given by both methods on all subjects in the data set. The OTD estimated by SQ-SNP and TR-OLS produce similar treatment decisions for the vast majority of subjects. SQ-SNP assigns more patients to treatment with IV fluids than TR-OLS and is less likely to assign vasoactive therapy. The point estimates are relatively close, but the smaller ASE for SQ-SNP suggests it is the more reliable estimator. The appendix contains Tables 6 and 7 that demonstrate the point estimates and ASE for $\boldsymbol{\beta}$ using the TR-OLS method on the 945 subjects with treatment and response information as a "gold standard" for comparison with our SQ-SNP method.

8 Discussion

We proposed a new method for estimating the optimal treatment decision within a specified class of linear decision rules at a single decision point, where the class is chosen based on considerations of simplicity and interpretation. Extensive simulation shows the proposed method, SQ-SNP, outperforms a supervised regression method under the correct specification and misspecification of the assumed *working* model. An application to an EMR study illustrates SQ-SNP improvements in efficiency of the linear regression coefficients involved in the optimal treatment decision. The estimated OTD provides understanding into the patient's outcome and preferred treatment based on their individual attributes, but it must be interpreted with caution due to issues such as potential bias from confounding variables in observational studies with EMR data (Hersh et al. 2013).

This method does have limitations. It is not designed directly for high dimensional data as the *curse of dimensionality* is an issue with kernel regression estimators. Dimension reduction techniques, such as principal component analysis or sliced inverse regression, can be utilized to reduce information provided by all the covariates into a lower-dimensional subspace for accurate prediction with kernel regression estimators, but this eliminates the ability to make simple interpretations from the decision rules created by SQ-SNP. An appropriate incorporation of high-dimensional data would allow for full utilization all the available information from EMR studies. Conventional assumptions in SSL tacitly require Y and A in $\mathcal{U}$ to be MCAR, but an assumption of missing at random (MAR), may be more applicable in many situations with data from observational studies. Our SQ-SNP method described in this article may be adapted to incorporate the propensity of Y and A missing given the covariates, $\mathbf{X}$, but this adaptation requires non-trivial alterations to the estimator for ASE based on influence functions presented in Section 5.3. The MAR assumption also complicates the convergence rate of the SQ-SNP estimator owing to the fact n is a random quantity in this setting, i.e. $n = \sum_{i=1}^{M} R_i$ where M represents all observation in $\mathcal{O} \cup \mathcal{U}$, and R denotes an indicator of Y or A being observed.

Research into EMR and EHR data has immense potential to improve the quality of health care for patients during a hospital stay while also reducing costs. The data accessi-

ble through EMRs/EHRs allow investigators to determine clinical decision support tools, personalized medicine, disease monitoring, phenotype prediction for patients at risk of developing diseases such as cancer or diabetes (Raghupathi and Raghupathi 2014). In the context of personalized medicine, there are many directions for research with EMR data. Further work on SS Q-learning may be extended to create dynamic optimal treatment regimes utilizing electronic health record (EHR) data with information on patients with extended periods of stay or multiple visits to the ICU to determine the optimal sequence of successive treatments for patients with chronic diseases. Section 7 discusses estimation of δ^{opt} by using variables obtained from structured EMR data. MIMIC-III also contains unstructured data through progress reports and discharge summaries recorded during a patient's hospital stay. Clinical text data holds a great deal of valuable patient information worth extracting for covariates in a Q-learning or A-learning model. Clinical natural language processing tools (e.g. cTakes) provides investigators with tools to extract information from clinical text such as a patient's disease status, symptoms, prescribed treatments, and family history (Savova et al. 2010). Further, method development for OTDs under situations where $\mathcal{A}$ extends beyond two treatment options is an important next step in creating comprehensive treatment recommendations for ICU patients. Estimation of optimal dosing for treatments benefits from utilizing EMR data, especially in combinination with a patient's genetic information for pharmacogenomic modeling of drug response (Wilke et al. 2011). Moreover, investigation into optimal constrained treatment decision estimators (Linn et al. 2015; Wang et al. 2018), i.e. a decision rule that maximizes an individual's primary outcome while simultaneously controlling for risk factors, is relevant to ICU settings where clinicians typically balance several potentially competing outcomes across a heterogeneous patient population. Future work into developing estimators for OTD's tailored to the challenges of EMR data provides new opportunity for investigation into improving the clinical decision-making process.

A Proof of Theorem 1.

Let $\widehat{C}(\mathbf{x}) = \widehat{Q}^{(np)}(\mathbf{x}, 1) - \widehat{Q}^{(np)}(\mathbf{x}, 0)$.

$$\begin{aligned}\left(\widehat{\boldsymbol{\beta}}_{np}-\boldsymbol{\beta}\right) &= \Lambda_N^{-1}\left[N^{-1}\sum_{j=n+1}^{n+N}\widetilde{\mathbf{X}}_j\left\{\widehat{C}(\mathbf{X}_j)-\boldsymbol{\beta}'\widetilde{\mathbf{X}}_j\right\}\right] \\ &= \Lambda_N^{-1}\left[N^{-1}\sum_{j=n+1}^{n+N}\widetilde{\mathbf{X}}_j\left\{\widehat{C}(\mathbf{X}_j)-C(\mathbf{X}_j)\right\}\right]+\Lambda_N^{-1}\left[N^{-1}\sum_{j=n+1}^{n+N}\widetilde{\mathbf{X}}_j\left\{C(\mathbf{X}_j)-\boldsymbol{\beta}'\widetilde{\mathbf{X}}_j\right\}\right] \\ &= \Lambda^{-1}E\left[\widetilde{\mathbf{X}}\left\{\widehat{C}(\mathbf{X})-C(\mathbf{X})\right\}\right]+O_p(N^{-1/2}).\end{aligned}$$

The first step follows from the normal equations. The last step is due to the fact $\Lambda_N^{-1}\left[N^{-1}\sum_{j=n+1}^{n+N}\widetilde{\mathbf{X}}_j\left\{\widehat{C}(\mathbf{X}_j)-C(\mathbf{X}_j)\right\}\right]=\Lambda^{-1}E\left[\widetilde{\mathbf{X}}\left\{\widehat{C}(\mathbf{X})-C(\mathbf{X})\right\}\right]+o_p(1)$ by standard arguments involving the weak law of large numbers. According to the central limit theorem $N^{-1/2}\left[N^{-1/2}\sum_{j=n+1}^{N}\Lambda_n^{-1}\left\{C(\mathbf{X})_j-\boldsymbol{\beta}'\widetilde{\mathbf{X}}_j\right\}\right]=O_p(N^{-1/2})$. Multiplying both sides by $n^{1/2}$ we have,

$$\begin{aligned}n^{1/2}\left(\widehat{\boldsymbol{\beta}}_{np}-\boldsymbol{\beta}\right) &= n^{1/2}\Lambda^{-1}E\left[\mathbf{X}\left\{\widehat{C}(\mathbf{x})-C(\mathbf{X})\right\}\right]+O_p\left((n/N)^{\frac{1}{2}}\right) &(11)\\ &= n^{1/2}\Lambda^{-1}E\left[\widetilde{\mathbf{X}}\left\{\widehat{Q}^{(np)}(\mathbf{X},1)-Q^{(np)}(\mathbf{X},1)\right\}\right] &(12)\\ &- n^{1/2}\Lambda^{-1}E\left[\widetilde{\mathbf{X}}\left\{\widehat{Q}^{(np)}(\mathbf{X},0)-Q^{(np)}(\mathbf{X},0)\right\}\right]+O_p\left((n/N)^{\frac{1}{2}}\right). &(13)\end{aligned}$$

Note that $n/N\to 0$ implying $O_p\left((n/N)^{\frac{1}{2}}\right)\equiv o_p(1)$. Next, let $\tau(\mathbf{X})=\pi(\mathbf{X})f(\mathbf{X})$ and $\widehat{\tau}(\mathbf{X})=\frac{1}{nh^p}\sum_{i=1}^{n}A_iW_h(\mathbf{X}_i-\mathbf{X})$, where $W_h(\mathbf{X}_i-\mathbf{X})=W(\frac{\mathbf{X}_i-\mathbf{X}}{h})$. Let's rewrite $E\left[\widetilde{\mathbf{X}}\left\{\widehat{Q}^{(np)}(\mathbf{X},1)-Q^{(np)}(\mathbf{X},1)\right\}\right]$ as,

$$\begin{aligned}&= E\left\{\frac{\frac{1}{nh^p}\sum_{i=1}^{n}A_i\mathbf{X}W_h(\mathbf{X}_i-\mathbf{X})\left\{Y_i-Q^{(np)}(\mathbf{X},1)\right\}}{\tau(\mathbf{X})}\right\} &(14)\\ &+E\left\{\widetilde{\mathbf{X}}\left(\widehat{Q}^{(np)}(\mathbf{X},1)-Q^{(np)}(\mathbf{X},1)\right)\left\{\frac{\tau(\mathbf{X})-\widehat{\tau}(\mathbf{X})}{\tau(\mathbf{X})}\right\}\right\}=H_{n,1}^{(1)}+H_{n,2}^{(1)}. &(15)\end{aligned}$$

Then $H_{n,1}^{(1)}$ is equivalent to:

$$
\begin{aligned}
&= \frac{1}{nh^p}\sum_{i=1}^{n} A_i \int \widetilde{\mathbf{X}} \left\{ Y_i - Q^{(np)}(\mathbf{X}, 1) \right\} \frac{W_h(\mathbf{X} - \mathbf{X}_i)}{\tau(\mathbf{X})} f(\mathbf{X}) d\mathbf{X} \\
&= \frac{1}{nh^p}\sum_{i=1}^{n} A_i \int \widetilde{\mathbf{X}} \left\{ Y_i - Q^{(np)}(\mathbf{X}, 1) \right\} \frac{W_h(\mathbf{X} - \mathbf{X}_i)}{\pi(\mathbf{X})} d\mathbf{X} \\
&= \frac{1}{n}\sum_{i=1}^{n} A_i \int \left(\widetilde{\mathbf{X}}_i + h\mathbf{t}_i \right) \left\{ Y_i - Q^{(np)}(\mathbf{X}_i + h\mathbf{t}_i, 1) \right\} \frac{W(\mathbf{t}_i)}{\pi\left(\mathbf{X}_i + h\mathbf{t}_i\right)} d\mathbf{t}_i
\end{aligned}
$$

By assumptions (1) and (2), Taylor expansion in $h\mathbf{t}_i$ for sufficiently small h leads to,

$$
H_{n,1}^{(1)} = \frac{1}{n}\sum_{i=1}^{n} \frac{A_i}{\pi\left(\mathbf{X}_i\right)} \widetilde{\mathbf{X}}_i \left\{ Y_i - Q^{(np)}(\mathbf{X}_i, 1) \right\} + O_p(h^r).
$$

Since $n^{1/2}h^r \to 0$ as $n \to \infty$,

$$
n^{1/2}\Lambda H_{n,1}^{(1)} = \widetilde{H}_{n,1}^{(1)} = n^{-1/2}\sum_{i=1}^{n} \frac{A_i}{\pi\left(\mathbf{X}_i\right)} \Lambda\widetilde{\mathbf{X}}_i \left\{ Y_i - Q^{(np)}(\mathbf{X}_i, 1) \right\} + o_p(1). \tag{16}
$$

Let $q(\mathbf{X}) = \widehat{Q}^{(np)}(\mathbf{X}, 1) - Q^{(1)}(\mathbf{X}, 1)$, $l(\mathbf{X}) = \dfrac{\tau(\mathbf{X}) - \widehat{\tau}(\mathbf{X})}{\tau(\mathbf{X})} = 1 - \dfrac{\widehat{\tau}(\mathbf{X})}{\tau(\mathbf{X})}$, and $Q^{(np)}(\mathbf{X}, 1) = \alpha(\mathbf{X})/\tau(\mathbf{X})$, where $\alpha(\mathbf{X})$ is the numerator in Equation 14. It follows that $H_{n,2}^{(1)} = E\left\{ \widetilde{\mathbf{X}} q(\mathbf{X}) l(\mathbf{X}) \right\}$ and,

$$
E\left\{ \widetilde{\mathbf{X}} q(\mathbf{X}) l(\mathbf{X}) \right\} \leq \sup_{\mathbf{x}\in\mathcal{X}} \left\{ \|\widetilde{\mathbf{X}}\| \; |q(\mathbf{X})| \; |l(\mathbf{X})| \right\} = o_p(1). \tag{17}
$$

Equation 17 requires that $\mathbf{X}$ is bounded, $\sqrt{\frac{\ln n}{nh^p}} \to 0$ as $n \to \infty$, as well as assumptions (3), (4) and (5). It follows by similar argument of *Lemma B.1* in Newey (1994) combined with Taylor series expansion that $\sup_{\mathbf{x}\in\mathcal{X}} |\widehat{\tau}(\mathbf{X}) - \tau(\mathbf{X})| = \sup_{\mathbf{x}\in\mathcal{X}} |\widehat{\alpha}(\mathbf{X}) - \alpha(\mathbf{X})| = o_p(1)$. Afterwards, it holds that $\sup_{\mathbf{x}\in\mathcal{X}} |q(\mathbf{X})| = o_p(1)$ through similar reasoning given by *Theorem 8* of Hansen (2008). The same technique proves,

$$
n^{1/2}\Lambda E\left[\widetilde{\mathbf{X}} \left\{ \widehat{Q}^{(np)}(\mathbf{X}, 0) - Q^{(np)}(\mathbf{X}, 0) \right\} \right] = \widetilde{H}_{n,1}^{(0)} + o_p(1). \tag{18}
$$

This leaves us with,

$$n^{1/2}\Lambda^{-1}E\left[\widetilde{\mathbf{X}}\left\{\widehat{C}(\mathbf{x})-C(\mathbf{X})\right\}\right]=\widetilde{H}_{n,1}^{(1)}-\widetilde{H}_{n,1}^{(0)}+o_p(1)=n^{-1/2}\sum_{i=1}^{n}\Psi_{np}(\mathbf{O}_i)+o_p(1).$$

Since $\Psi_{np}(\mathbf{O})$ is the influence function for $\widehat{\boldsymbol{\beta}}_{np}$ with $E\left[\Psi_{np}(\mathbf{O})\right]=0$ and variance $\mathcal{V}_{\boldsymbol{\beta}_{np}}=E\left[\Psi_{np}(\mathbf{O})\Psi_{np}'(\mathbf{O})\right]$, the *CLT* states it will converge to $\mathcal{N}_{p+1}(0,\mathcal{V}_{\boldsymbol{\beta}_{np}})$.

B Proof of Theorem 2.

The proof begins similarly to the proof of Theorem 1. Let $\widehat{C}^{(SS)}(\mathbf{X})=\widehat{Q}^{(SS)}(\mathbf{X},1,\widehat{\boldsymbol{\theta}}_1)-\widehat{Q}^{(SS)}(\mathbf{X},0,\widehat{\boldsymbol{\theta}}_0)$.

$$\begin{aligned}
n^{1/2}\left(\widehat{\boldsymbol{\beta}}_{SS}-\boldsymbol{\beta}\right) &= n^{1/2}\Lambda_N^{-1}E\left[\widetilde{\mathbf{X}}\left\{\widehat{C}(\mathbf{x})-C(\mathbf{X})\right\}\right]+o_p\left(1\right) && (19)\\
&= n^{1/2}\Lambda_N^{-1}E\left[\widetilde{\mathbf{X}}\left\{\widehat{Q}^{(SS)}(\mathbf{X},1,\widehat{\boldsymbol{\theta}}_1)-Q^{(SS)}(\mathbf{X},1,\boldsymbol{\theta}_1)\right\}\right] && (20)\\
&- n^{1/2}\Lambda_N^{-1}E\left[\widetilde{\mathbf{X}}\left\{\widehat{Q}^{(SS)}(\mathbf{X},0,\widehat{\boldsymbol{\theta}}_0)-Q^{(SS)}(\mathbf{X},0,\boldsymbol{\theta}_0)\right\}\right]+o_p\left(1\right). && (21)
\end{aligned}$$

Next, recall that

$$\widehat{Q}^{(SS)}(\mathbf{x},a,\widehat{\boldsymbol{\theta}}_a)=\frac{1}{\mathcal{K}}\sum_{k=1}^{\mathcal{K}}\widehat{Q}_k^{(np)}\left(\mathbf{x},a\right)+\widehat{\boldsymbol{\theta}}_a'\widetilde{\mathbf{x}}$$

and,

$$Q^{(SS)}(\mathbf{x},a,\boldsymbol{\theta}_a)=Q^{(np)}\left(\mathbf{x},a\right)+\boldsymbol{\theta}_a'\widetilde{\mathbf{x}}.$$

This allows us to rewrite Equation (20) in the following way,

$$n^{1/2}\Lambda_N^{-1}\left\{\frac{1}{N}\sum_{j=n+1}^{N}\widetilde{\mathbf{X}}_j\left\{\frac{1}{\mathcal{K}}\sum_{k=1}^{\mathcal{K}}\widehat{Q}_k^{(np)}\left(\mathbf{X}_j,a\right)+\widehat{\boldsymbol{\theta}}_a'\widetilde{\mathbf{X}}_j\right\}-Q^{(np)}\left(\mathbf{X}_j,a\right)+\boldsymbol{\theta}_a'\widetilde{\mathbf{X}}_j\right\}+o_p(1). \quad (22)$$

Equation (21) can be rewritten analogously. Define,

$$\widehat{S}_a^{\mathcal{K}} = \frac{1}{\mathcal{K}} \sum_{k=1}^{\mathcal{K}} \left\{ \frac{1}{N} \sum_{j=n+1}^{n+N} \widetilde{\mathbf{X}}_j \left\{ \widehat{Q}_k^{(np)}(\mathbf{X}_j, a) - Q^{(np)}(\mathbf{X}_j, a) \right\} \right\}.$$

Rearranging the summands in Equation (22) allows us to formulate Equation (19),

$$n^{1/2} \left(\widehat{\boldsymbol{\beta}}_{SS} - \boldsymbol{\beta} \right) = n^{1/2} \Lambda_N^{-1} \left\{ \Lambda_N (\widehat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1) + \widehat{S}_1^{\mathcal{K}} \right\} - n^{1/2} \Lambda_N^{-1} \left\{ \Lambda_N (\widehat{\boldsymbol{\theta}}_0 - \boldsymbol{\theta}_0) + \widehat{S}_0^{\mathcal{K}} \right\} + o_p(1) \quad (23)$$

If we follow similar logic to the proof of *Theorem 3.2* in Chakrabortty and Cai (2018) from this point on, we complete the proof of Theorem 2.

C Additional Tables for EMR Analysis.

The tables provided below use all 945 subjects to get point estimates and ASE estimates using the TR-OLS method discussed in Section 4. The propensity score model uses the same variables described in Section 7, but utilize the 945 subjects with treatment and response information available.

Acknowledgements

References

Beaulieu-Jones, B. K., Greene, C. S. et al. (2016), "Semi-supervised learning of the electronic health record for phenotype stratification," *Journal of biomedical informatics*, 64, 168–178.

Beaulieu-Jones, B. K., Lavage, D. R., Snyder, J. W., Moore, J. H., Pendergrass, S. A., and Bauer, C. R. (2018), "Characterizing and managing missing structured data in electronic health records: data analysis," *JMIR medical informatics*, 6, e11.

Chakrabortty, A., and Cai, T. (2018), "Efficient and adaptive linear regression in semi-supervised settings," *The Annals of Statistics*, 46, 1541–1572.

Chapelle, O., Scholkopf, B., and Zien, A. (2006), *Semi-Supervised Learning* MIT Press.

Fan, J. (1992), "Design-adaptive nonparametric regression," *Journal of the American statistical Association*, 87, 998–1004.

Garla, V., Taylor, C., and Brandt, C. (2013), "Semi-supervised clinical text classification with Laplacian SVMs: an application to cancer case management," *Journal of biomedical informatics*, 46, 869–875.

Hansen, B. E. (2008), "Uniform convergence rates for kernel estimation with dependent data," *Econometric Theory*, 24, 726–748.

Hersh, W. R., Weiner, M. G., Embi, P. J., Logan, J. R., Payne, P. R., Bernstam, E. V., Lehmann, H. P., Hripcsak, G., Hartzog, T. H., Cimino, J. J. et al. (2013), "Caveats for the use of operational electronic health record data in comparative effectiveness research," *Medical care*, 51(8 0 3), S30.

Hong, C., Liao, K. P., and Cai, T. (2019), "Semi-supervised validation of multiple surrogate outcomes with application to electronic medical records phenotyping," *Biometrics*, 75, 78–89.

Johnson, A. E., Pollard, T. J., Shen, L., Li-wei, H. L., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Celi, L. A., and Mark, R. G. (2016), "MIMIC-III, a freely accessible critical care database," *Scientific data*, 3, 160035.

Kellum, J. A., Lameire, N., Aspelin, P., Barsoum, R. S., Burdmann, E. A., Goldstein, S. L., Herzog, C. A., Joannidis, M., Kribben, A., Levey, A. S. et al. (2012), "Kidney disease: improving global outcomes (KDIGO) acute kidney injury work group. KDIGO clinical practice guideline for acute kidney injury," *Kidney international supplements*, 2, 1–138.

Le Gall, J., Loirat, P., Alperovitch, A., Glaser, P., Granthil, C., Mathieu, D., Mercier, P., Thomas, R., and Villers, D. (1984), "A simplified acute physiology score for ICU patients.," *Critical care medicine*, 12, 975–977.

Lee, J., Kothari, R., Ladapo, J. A., Scott, D. J., and Celi, L. A. (2012), "Interrogating a clinical database to study treatment of hypotension in the critically ill," *BMJ open*, 2, e000916.

Lehman, L. w., Saeed, M., Moody, G., and Mark, R. (2010), "Hypotension as a risk factor for acute kidney injury in ICU patients," *Computing in cardiology*, 37, 1095.

Linn, K. A., Laber, E. B., and Stefanski, L. A. (2015), "Estimation of dynamic treatment regimes for complex outcomes: Balancing benefits and risks," in *Adaptive Treatment Strategies in Practice: Planning Trials and Analyzing Data for Personalized Medicine* SIAM, pp. 249–262.

Maslove, D. M., Lamontagne, F., Marshall, J. C., and Heyland, D. K. (2017), "A path to precision in the ICU," *Critical Care*, 21, 79.

Moodie, E. E., Richardson, T. S., and Stephens, D. A. (2007), "Demystifying optimal dynamic treatment regimes," *Biometrics*, 63, 447–455.

Murphy, S. A. (2003), "Optimal dynamic treatment regimes," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65, 331–355.

Nahum-Shani, I., Qian, M., Almirall, D., Pelham, W. E., Gnagy, B., Fabiano, G. A., Waxmonsky, J. G., Yu, J., and Murphy, S. A. (2012), "Q-learning: A data analysis method for constructing adaptive interventions.," *Psychological Methods*, 17, 478.

Newey, W. K. (1994), "Kernel estimation of partial means and a general variance estimator," *Econometric Theory*, 10, 1–21.

Raghu, A., Komorowski, M., Celi, L. A., Szolovits, P., and Ghassemi, M. (2017), Continuous State-Space Models for Optimal Sepsis Treatment: a Deep Reinforcement Learning Approach,, in *Machine Learning for Healthcare Conference*, pp. 147–163.

Raghupathi, W., and Raghupathi, V. (2014), "Big data analytics in healthcare: promise and potential," *Health information science and systems*, 2, 3.

Robins, J. M. (2004), "Optimal Structural Nested Models for Optimal Sequential Decisions," in *Proceedings of the Second Seattle Symposium in Biostatistics* Springer, pp. 189–326.

Rosenbaum, P. R., and Rubin, D. B. (1983), "The central role of the propensity score in observational studies for causal effects," *Biometrika*, 70, 41–55.

Rubin, D. B. (1978), "Bayesian Inference for Causal Effects: The Role of Randomization," *The Annals of Statistics*, 6, 34–58.

Savova, G. K., Masanz, J. J., Ogren, P. V., Zheng, J., Sohn, S., Kipper-Schuler, K. C., and Chute, C. G. (2010), "Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications," *Journal of the American Medical Informatics Association*, 17, 507–513.

Schulte, P., Tsiatis, A., Laber, E., and Davidian, M. (2014), "Q-and A-learning methods for estimating optimal dynamic treatment regimes," *Statistical Science: a Review Journal of the Institute of Mathematical Statistics*, 29, 640–661.

van Walraven, C., Austin, P. C., Jennings, A., Quan, H., and Forster, A. J. (2009), "A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data," *Medical care*, pp. 626–633.

Wang, Y., Fu, H., and Zeng, D. (2018), "Learning optimal personalized treatment rules in consideration of benefit and risk: with an application to treating type 2 diabetes patients with insulin therapies," *Journal of the American Statistical Association*, 113, 1–13.

Wang, Y., Wu, P., Liu, Y., Weng, C., and Zeng, D. (2016), "Learning Optimal Individualized Treatment Rules from Electronic Health Record Data," *2016 IEEE International Conference on Healthcare Informatics (ICHI)*, pp. 65–71.

Watkins, C. J., and Dayan, P. (1992), "Q-learning," *Machine learning*, 8, 279–292.

Weiskopf, N. G., Hripcsak, G., Swaminathan, S., and Weng, C. (2013), "Defining and measuring completeness of electronic health records for secondary use," *Journal of biomedical informatics*, 46, 830–836.

Wilke, R., Xu, H., Denny, J., Roden, D., Krauss, R., McCarty, C., Davis, R., Skaar, T., Lamba, J., and Savova, G. (2011), "The emerging role of electronic medical records in pharmacogenomics," *Clinical Pharmacology & Therapeutics*, 89, 379–386.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2012), "A robust method for estimating optimal treatment regimes," *Biometrics*, 68, 1010–1018.

Zhao, Y., Zeng, D., Rush, A. J., and Kosorok, M. R. (2012), "Estimating individualized treatment rules using outcome weighted learning," *Journal of the American Statistical Association*, 107, 1106–1118.

Zhu, X., and Goldberg, A. B. (2009), "Introduction to semi-supervised learning," *Synthesis lectures on artificial intelligence and machine learning*, 3, 1–130.

Table 1: Average PCD and Average Value Function over 500 replications for $\widehat{\delta}^{opt}_{\boldsymbol{\beta}_{TR}}$ and $\widehat{\delta}^{opt}_{\boldsymbol{\beta}_{SS}}$. Empirical SEs are provided in parentheses.

			TR-OLS		SQ-SNP	
$\mu(\mathbf{X})$	Model	V_0	V	PCD	V	PCD
$(\boldsymbol{\omega}'\mathbf{X})^3$	Linear	0.56	0.54 (0.04)	0.92 (0.06)	0.56 (0.01)	0.96 (0.02)
	Cubic	0.24	0.22 (0.06)	0.87 (0.12)	0.24 (0.01)	0.93 (0.05)
	Sine	0.32	0.21 (0.12)	0.80 (0.17)	0.26 (0.06)	0.88 (0.09)
$(\boldsymbol{\alpha}'\mathbf{X})\,(1+\boldsymbol{\omega}'\mathbf{X})$	Linear	1.31	1.29 (0.05)	0.91 (0.06)	1.31 (0.01)	0.96 (0.02)
	Cubic	0.99	0.98 (0.05)	0.86 (0.11)	0.99 (<0.01)	0.94 (0.04)
	Sine	1.06	0.96 (0.11)	0.80 (0.16)	1.02 (0.04)	0.89 (0.06)

Table 2: Component-wise Bias, Empirical SE, Asymptotic SE, Coverage Probability (CP) for 95% CI, and RE under both baseline effects.

(a) $\mu(\mathbf{X}) = (\boldsymbol{\omega}'\mathbf{X})^3$

		TR-OLS				SQ-SNP				
Model	$\boldsymbol{\beta}$	Bias	ESE	ASE	CP	Bias	ESE	ASE	CP	RE
Linear	0	−0.010	0.220	0.209	0.94	−0.005	0.123	0.122	0.95	3.21
	1	−0.021	0.329	0.347	0.98	−0.008	0.182	0.173	0.93	3.28
	1	−0.020	0.355	0.347	0.97	−0.004	0.198	0.175	0.92	3.22
Cubic	0	−0.008	0.206	0.205	0.94	0.001	0.123	0.132	0.95	2.80
	0.41	0.002	0.352	0.338	0.95	−0.002	0.212	0.193	0.95	2.74
	0.81	0.002	0.423	0.390	0.94	−0.010	0.239	0.212	0.91	3.14
Sine	0	−0.006	0.171	0.168	0.96	0.004	0.118	0.116	0.95	2.13
	0.37	−0.007	0.282	0.270	0.94	−0.011	0.170	0.158	0.91	2.74
	0.37	0.011	0.296	0.272	0.94	0.011	0.176	0.161	0.92	2.82

(b) $\mu(\mathbf{X}) = (\boldsymbol{\alpha}'\mathbf{X})\,(1 + \boldsymbol{\omega}'\mathbf{X})$

		TR-OLS				SQ-SNP				
Model	$\boldsymbol{\beta}$	Bias	ESE	ASE	CP	Bias	ESE	ASE	CP	RE
Linear	0	−0.017	0.242	0.230	0.94	−0.007	0.114	0.116	0.95	4.53
	1	−0.012	0.348	0.326	0.94	0.011	0.150	0.151	0.93	5.37
	1	−0.021	0.352	0.347	0.94	−0.011	0.163	0.154	0.91	4.66
Cubic	0	−0.005	0.236	0.223	0.93	−0.026	0.121	0.122	0.95	3.77
	0.41	−0.014	0.341	0.313	0.93	0.003	0.155	0.154	0.92	4.86
	0.81	−0.004	0.432	0.392	0.91	−0.008	0.193	0.186	0.92	4.99
Sine	0	0.003	0.210	0.202	0.94	0.009	0.117	0.113	0.95	3.14
	0.37	0.002	0.296	0.289	0.94	−0.004	0.146	0.146	0.94	4.12
	0.37	−0.005	0.300	0.290	0.93	0.001	0.136	0.140	0.95	4.90

Table 3: Mean and SD of Fully Observed ($\mathcal{O}$) and Unlabeled ($\mathcal{U}$) data before scaling. Wilcoxon Rank Sum Tests and Kolmogorov−Smirnov tests to compare both data sets.

	$\mathcal{O}$		$\mathcal{U}$		P-values from Diagnostic Tests	
Predictors	Mean	SD	Mean	SD	Wilcoxon Test	KS Test
Mean MAP	65.068	7.793	65.091	7.971	0.842	0.897
Mean heart rate	84.988	17.055	83.650	17.140	0.115	0.171

Table 4: Estimated regression coefficients, ASE, and p-values for testing $H_0 : \boldsymbol{\beta}_{TR} = 0$ and $\boldsymbol{\beta}_{SS} = 0$.

	TR-OLS			SQ-SNP		
Predictors	$\boldsymbol{\beta}_{TR}$	ASE	P-value	$\boldsymbol{\beta}_{SS}$	ASE	P-value
Intercept	0.020	0.136	0.883	−0.050	0.126	0.695
Mean MAP	0.043	0.094	0.640	0.037	0.066	0.579
Mean heart rate	−0.096	0.057	0.089	−0.104	0.050	0.040

Table 5: Treatment allocation given by SQ-SNP and TR-OLS.

		SQ-SNP	
	Treatment	IV Fluid	Vasopressors
TR-OLS	IV Fluid	1430	29
	Vasopressors	459	1454

Table 6: Estimated regression coefficients, ASE, and p-values for testing $H_0 : \boldsymbol{\beta}_{TR} = 0$.

Predictors	$\boldsymbol{\beta}_{TR}$	ASE	P-value
Intercept	−0.049	0.040	0.220
Mean MAP	0.049	0.035	0.160
Mean heart rate	−0.102	0.046	0.025

Table 7: Estimated regression coefficients, ASE, and p-values for testing $H_0 : \boldsymbol{\beta}_{TR} = 0$ using all variables in the propensity score model.

	$\boldsymbol{\beta}$	ASE	P-value
Intercept	−0.625	0.852	0.463
baseline creatinine	−0.072	0.133	0.586
total urine output	−0.008	0.049	0.878
mean MAP	0.043	0.035	0.221
age	−0.042	0.057	0.465
mean heart rate	−0.108	0.045	0.016
Elixhauser	0.166	0.175	0.342
SAPS I	−0.001	0.207	0.997
gender	0.068	0.093	0.464
surgery service	0.056	0.085	0.509