Published in final edited form as:

Proc Int Conf Data Sci Adv Anal. 2019 October ; 2019: 392–402. doi:10.1109/dsaa.2019.00054.

Learning Personalized Treatment Rules from Electronic Health Records Using Topic Modeling Feature Extraction

Peng Wu, Tianchen Xu, Yuanjia Wang

Department of Biostatistics Columbia University

Abstract

To address substantial heterogeneity in patient response to treatment of chronic disorders and achieve the promise of precision medicine, individualized treatment rules (ITRs) are estimated to tailor treatments according to patient-specific characteristics. Randomized controlled trials (RCTs) provide gold standard data for learning ITRs not subject to confounding bias. However, RCTs are often conducted under stringent inclusion/exclusion criteria, and participants in RCTs may not reflect the general patient population. Thus, ITRs learned from RCTs lack generalizability to the broader real world patient population. Real world databases such as electronic health records (EHRs) provide new resources as complements to RCTs to facilitate evidence-based research for personalized medicine. However, to ensure the validity of ITRs learned from EHRs, a number of challenges including confounding bias and selection bias must be addressed. In this work, we propose a matching-based machine learning method to estimate optimal individualized treatment rules from EHRs using interpretable features extracted from EHR documentation of medications and ICD diagnoses codes. We use a latent Dirichlet allocation (LDA) model to extract latent topics and weights as features for learning ITRs. Our method achieves confounding reduction in observational studies through matching treated and untreated individuals and improves treatment optimization by augmenting feature space with clinically meaningful LDA-based features. We apply the method to EHR data collected at New York Presbyterian Hospital clinical data warehouse in studying optimal second-line treatment for type 2 diabetes (T2D) patients. We use cross validation to show that ITRs outperforms uniform treatment strategies (i.e., assigning same treatment to all individuals), and including topic modeling features leads to more reduction of post-treatment complications.

I. Introduction

Personalized treatment decision making has been proposed as a paradigm shift from the universal, “one-size-fits-all” strategy [1] to address substantial heterogeneity between patients affected by chronic disorders such as type 2 diabetes (T2D) and mental disorders [2]. Recently, personalized medicine research is fast growing and greatly facilitated by large-scale data collection and technological advances in data storage and processing [1].

pw2394@cumc.columbia.edu.

Appendix

This research is approved by the Institutional Review Board (IRB). Due to confidentiality restrictions from the IRB, the individual-level EHR data presented here cannot be shared. However, we share the codes for analyzing these data at https://github.com/PWuJuveOKC/T2D_EHR_LDA.

Personalized medicine allows healthcare providers to tailor treatment or treatment sequences to individual patient accounting for personalized information and patient heterogeneity [3].

One way to personalize medicine is to prescribe treatment using individualized treatment rules (ITRs) which are mappings from a patient's feature space (e.g., biomarkers of patient's health status) to the space of potential treatment decisions. Randomized controlled trials (RCTs) are considered to be of high internal validity due to the virtue of randomizing treatments so that RCTs are not subject to unmeasured confounding. However, due to their stringent inclusion and exclusion criteria, RCTs lack generalizability or external validity to a broader and diversified population [4]. Studies comparing T2D patient population in EHRs and participants of RCTs for T2D found a large difference in their distributions [5]. In contrast, EHRs provide a complementary resource for learning ITRs in a large real world patient population. Recently, large scale EHR data has been leveraged to characterize treatment pathways in medical decision making [6]. Due to the nonexperimental characteristics in retrospectively documented observational data, it is difficult to infer causality or valid ITRs without employing confounding and selection bias reduction techniques [7], [8]. It is beneficial to integrate medical domain knowledge in feature engineering with valid machine learning and statistical methods in alleviating biases to better estimate optimal decision rule in observational studies.

Machine learning methods are commonly used to optimize treatment decisions by estimating ITRs according to patient level feature variables. Indirect methods such as Q-learning [9], [10] and A-learning [11] fit a regression model to predict clinical outcome under alternative treatments and derive optimal ITRs by comparing predicted outcomes. Such methods are subject to model misspecification. Alternatively, outcome weighted learning (O-learning) and augmented O-learning [12], [13] are proposed to directly maximize clinical outcome after treatment. O-learning converts treatment optimization to a classification problem without requiring a model to predict the treatment responses. Tree-based methods use criteria related to testing for treatment by feature variable interaction effect to partition patients [14], [15]. Most of these methods are suitable for estimating ITRs from RCTs and cannot be directly applied to observational studies without adjusting for various sources of biases. Recently, a matching-based learning method referred as M-learning [16] is proposed to generalize O-learning methods under a unified framework and can be applicable to EHRs. However, M-learning method does not optimally engineer feature variables from EHRs.

EHR data contains rich patient level features to build effective ITRs. Some information in EHRs is documented in the form of text notes or similar structure that enable researchers to use natural language processing (NLP) techniques to better extract the information. NLP techniques including latent semantic analysis (LSA) [17], probabilistic latent semantic analysis (pLSA) [18] have been widely used to represent features in a lower dimensional space. LSA is a well-known approach that maps count type data of documents to a reduced latent semantic space based on Singular Value Decomposition (SVD) while pLSA improves LSA in a statistical way that relies on a mixture decomposition [18]. Furthermore, latent Dirichlet allocation (LDA) is proposed as another generative probabilistic model for NLP. It overcomes some disadvantages of pLSA such as generalizability to new documents and overfitting problems [19].

In this paper, we present analyses methods to estimate optimal individualized treatments from EHR data with two main goals. The first goal is to achieve more effective feature extraction for identifying optimal ITRs with EHRs. The extracted features should be predictive in selecting optimal treatment and interpretable to clinical practitioners. The second goal aims at improving algorithms for treatment optimization and reducing confounding bias in statistical estimation especially in real world settings such as EHRs. We will use LDA as a type of topic models to analyze medication and condition domain of EHR data to extract features and apply three learning methods (Q-learning, O-learning and M-learning) to estimate the optimal ITRs depending on a patient's personalized characteristics such as demographics and predictive markers augmented by LDA-based features. We compare the performance of the three learning algorithms with or without augmenting feature space by LDA-based topic features. Additionally, we investigate important LDA-based features to provide the interpretation of ITRs using engineered topics in co-medication and diagnosis domains. The proposed method has the promise of discovering medical knowledge from observational data to facilitate individualized medical decision making.

II. Methods

A. Individualized Treatment Rule Framework

An individualized treatment rule (ITR) is a decision function that maps patients' features into the space of treatment options. In this paper, we consider single stage decision rules and binary treatment options. Extension to multiple treatment options can be achieved using methods in [20]. Let H_i denote the feature variables measured prior to treatment and R_i denote the clinical outcome or reward post treatment and we can assume a larger R_i corresponds to more desirable treatment effect (e.g., symptom relief). Let T denotes the treatment assignment taking values of -1 or 1 , and let $\mathcal{D}(H_i)$ denote an ITR. The value function used to evaluate an ITR associated with the decision rule $\mathcal{D}$ is defined as the expected post-treatment reward:

$$V(\mathcal{D}) = E^{\mathcal{D}}(R_i).$$

In particular, when the treatment assignment probability is known in an RCT, the value function in this expectation can be written as $V(\mathcal{D}) = E\left[\frac{R}{\pi(T, H)} I\{T = \mathcal{D}(H)\}\right]$, where $\pi(t, h)$ is the randomization probability for $T = t$ given $H = h$. In contrast, in an observational study (e.g., EHR data) treatment propensities $\pi(t, h)$ are usually unknown and need to be estimated from data using classification models such as logistic regression or random forest [21]. The value function can be estimated empirically from data as

$$V_n(\mathcal{D}) = \frac{1}{n} \sum_{i=1}^n \frac{I(T_i = \mathcal{D}(H_i)) R_i}{\pi(T_i, H_i)}. \quad (1)$$

Optimal ITRs are estimated from data by maximizing the above empirical value function (1).

B. Q-learning and O-learning Methods

There are many machine learning tools to estimate ITRs including Q-learning and O-learning methods. Q-learning is a reinforcement learning algorithm that plays an important role in estimating ITRs for multi-stage studies [10]. In Q-learning methods, regression-based approach is implemented to approximate R in $V(\mathcal{D})$ by the estimated treatment benefit through a model $R = \hat{f}(H, T)$ using H and interactions between H and treatment assignment T as predictors. The optimal ITR selected by Q-learning is expressed as $t^* = \max_{t \in \{1, -1\}} \hat{f}(H, t)$ [8]. Q-learning indirectly maximizes the value function and may be sensitive to incorrect model assumptions when feature space is high-dimensional [12], [8].

O-learning is a class of machine learning methods that directly maximize the expected clinical reward under a treatment assignment strategy. In the original O-learning, ITR estimation is transformed into a weighted classification problem by minimizing the empirical misclassification rate, $E\left[\frac{I(T_i \neq \mathcal{D}(H_i))R_i}{\pi(T_i, H_i)}\right]$ [12]. Augmented O-learning (AOL) is an improved version of O-learning by replacing R_i with a doubly-robust residual of $R_i - E[R_i|H_i]$ [13]. Efficiency gain can be guaranteed in the AOL by integrating O-learning with regression-based Q-learning in a doubly-robust manner [13]. However, O-learning methods are based on inverse probability weighting (IPW) of estimated propensities $\pi(T_i, H_i)$ and may suffer from several issues such as instability of weights estimation and imbalance of covariates distribution across treatment groups [16] when applied to observational data.

C. M-learning Method

In this paper, we focus on M-learning method. M-learning is a matching-based learning method that also directly maximizes the value function. It generalizes O-learning methods and integrates them with matching techniques. Specifically, M-learning performs matching to accurately measure individuals' treatment responses to alternative treatments, which requires less model specification and provides more flexibility in controlling confounders in the subgroups [16]. In observational studies, $\pi(T_i, H_i)$ is unknown and confounding bias commonly complicates analyses. It is well known that IPW-based methods for adjusting for confounding suffer from high variability especially when weights $\pi(T_i, H_i)$ are small in subgroups of individuals and lack control over subgroup propensity score balance [22]. Thus M-learning that avoids inverse weighting of $\pi(T_i, H_i)$ may be more advantageous over IPW-based approaches in the observational studies [16].

The key idea behind M-learning is that when two subjects are matched in terms of confounders or propensity scores of treatments but receive different treatments, the observed treatment leading to a larger clinical response should be more likely to be the optimal option. We describe M-learning method as follows: for any subject i , we first identify a matched set $\mathcal{M}_i$, which consists of subjects with opposite treatments but similar feature variables as subject i , where similarity is defined under a suitable distance metric. That is, we let

$$\mathcal{M}_i = \{j : T_j = -T_i, d(H_j, H_i) \leq \delta_i\}.$$

For any given ITR, we define the matching-based preference value function to maximize as

$$V_n(\mathcal{D}; u) = n^{-1} \sum_{i=1}^n |\mathcal{M}_i|^{-1} \sum_{j \in \mathcal{M}_i} u(|R_j - R_i|) \times \{I(R_j \geq R_i, \mathcal{D}(H_i) = -T_i) + I(R_j \leq R_i, \mathcal{D}(H_i) = T_i)\}. \quad (2)$$

where $|\mathcal{M}_i|$ is the size of the matched set $\mathcal{M}_i$ and $u(\cdot)$ is proposed as a monotonically increasing matching function to weight individual subjects. Common choice of $u(\cdot)$ can be $u(x) = 1$ or $u(x) = x$ according to users' objective. Note that when $u(x) = x$ and we choose one-nearest-neighbor matching, the value function is weighted by the absolute difference between the clinical reward in the matched pair, i.e., $|R_j - R_i|$. Equivalently, value function (2) is further written as

$$V_n(f; u) = n^{-1} \sum_{i=1}^n |\mathcal{M}_i|^{-1} \sum_{j \in \mathcal{M}_i} u(|R_j - R_i|) \times I(f(H_i)T_i \text{sign}(R_j - R_i) \leq 0), \quad (3)$$

where $\mathcal{D}(H) = \text{sign}(f(H))$ for some decision function. To estimate the optimal ITR in (3), it is equivalent to minimize the loss function

$$L_n(f; u) = n^{-1} \sum_{i=1}^n |\mathcal{M}_i|^{-1} \sum_{j \in \mathcal{M}_i} u(|R_j - R_i|) \times I(f(H_i)T_i \text{sign}(R_j - R_i) \geq 0). \quad (4)$$

Due to the discontinuity of the indicator function in (4), we replace the zero-one loss by alternative loss function for optimization. In this paper, we use hinge-loss and the optimization becomes a weighed support vector machine (SVM) problem for matched pairs:

$$V_\phi(f; u) = n^{-1} \sum_{i=1}^n |\mathcal{M}_i|^{-1} \sum_{j \in \mathcal{M}_i} u(R_j - R_i) \times \phi(-f(H_i)T_i \text{sign}(R_j - R_i)) + \lambda_n \|f\|_{\mathcal{H}_K}. \quad (5)$$

In $V_\phi(f, u)$, $\phi(x) = \max(1 - x, 0)$, λ_n is a penalty parameter and $\mathcal{H}_K$ is a reproducing kernel Hilbert space (RKHS) with kernel function $K(\cdot, \cdot)$. The solution to the dual problem of (5) using any off-the-shelf quadratic programming packages leads to the minimization of M-learning.

Additionally, the use of idea in doubly robust matching estimator (DRME) [23] can improve M-learning. The DRME incorporates both propensity score and prognostic score in matching set $\mathcal{M}_i$. The propensity score model $\pi(H)$ and prognostic score model $m(H)$ are defined as

$$\pi(H) = P(T = 1 | H) = g_1(H'\gamma_1)$$

and

$$m(H) = E(R | A = -1, H) = g_2(H'\gamma_2)$$

where γ_1, γ_2 denotes parameters for the two models. This extension in M-learning is doubly robust and is consistent even if one of the two models is misspecified [16].

It has been discussed in [16] that original O-learning and single-stage AOL are special cases of M-learning. Unlike O-learning methods which compare observed reward R with a constant or predicted outcome, M-learning directly compares observed individual rewards from the two subjects in the matched set $\mathcal{M}_i$. This demonstrates that M-learning methods have advantage over O-learning methods in ITR estimation leveraging individual information instead of predicted outcomes averaged across treatments [16].

D. EHR Data

The EHR data was collected and stored in a large clinical data warehouse (CDW) at New York Presbyterian Hospital (NYPH) which contains over 20 years of health information for about 4.5 million patients [24]. There are over 115,000 patients who had been diagnosed with type 2 diabetes (T2D) [5]. The quality of the EHR data including completeness, correctness, and plausibility was investigated to ensure suitability for studying treatment sequence of common diseases [25], [26]. Current T2D guidelines suggest metformin (MET) as the preferred first-line medication [27], but in the real world setting, there is no recommendation of optimal sequence of treatments on the long-term outcomes in the literature [28].

In this paper, we propose machine learning methods to estimate the optimal second-line treatment of T2D, specifically, we compare Met + insulin versus Met + sulfonylureas (SFU). To match this goal, we used a new-user design to select our cohort for analyses. Since our goal is to optimize second-line treatments, we included patients who had a first-line treatment and switched to a second-line treatment of SFU or insulin during 2008 and 2012. Each subject's observations are re-aligned at the time of switching (index time 0). A median baseline period of 12-months prior to the index time was established to extract baseline feature variables. The median follow-up period post index time to evaluate treatment response is 18-months. The primary outcome is the number of major complications of T2D after second-line treatment including essential hypertension, hyperlipidemia and hypercholesterolemia.

Unlike randomized experiment, EHR data are collected in an uncontrolled setting and pose challenges for inferring causal relationships in research studies. In our analyses, we address three major issues commonly confronting observational study research including confounding bias, selection bias and missing data. To alleviate confounding bias, we implement matching step in M-learning on potential confounders (e.g., lab test features),

two propensity scores constructed by two logistic regression models for lab measurement pattern features and demographics predictors and a doubly-robust prognostic score estimated from a linear regression model. To reduce selection bias due to missing post second-line treatment outcomes, IPW adjustment was implemented and a logistic regression model was used to predict the presence of at least one post-treatment measure during the follow up period. Lastly, missing features were imputed with chained equations.

E. LDA-Based Features for Co-medication and Diagnosis

In this paper, we extracted clinically meaningful features from two EHR data domains (i.e., medications and ICD diagnosis codes) using latent Dirichlet allocation model (LDA) as a topic model based approach. LDA is a three-level hierarchical generative model suitable for analyzing collections of discrete data and has broad applications from information retrieval to medical informatics [19]. When applied to EHRs, it can organize a large number of ICD diagnoses codes, procedure codes and medication prescription documentations into meaningful topics based on the correlation among variables. As a generative model, LDA is parameterized by distributions over topics, where each topic is characterized by a distribution over words in that topic. Thus, LDA models are highly interpretable.

In our EHR analyses, each patient can be viewed as a “document” in NLP applications. LDA model is adapted here to learn the hidden topics in the domain information for the collection of patients. In the medication prescription domain, the medication items in the generic name field except the T2D treatments are the “words” in a “patient’s document”. The latent topics learned in this domain are referred to as co-medication topics. Similarly, the ICD9 codes in the diagnosis domain are useful terms to learn topics of patient co-morbidities. The number of topics is fixed for the two domains and the latent topics were learned separately.

Specifically, the key idea of using LDA to analyze medication prescriptions and ICD codes in EHRs is that the patient population (collection of “documents”) can be represented as mixture of multiple latent subgroups or topics and each topic is defined to be a multinomial distribution over a collection of “vocabulary” of medications or conditions (i.e., terms) [19], [29]. LDA assumes a probabilistic generative process by which each patient is generated with K latent topics and each topic is assumed to follow a Dirichlet distribution over V medications or conditions in the “vocabulary”. We describe the process as follows [30]. Let $\mathcal{C}$ denote a corpus composed of M patients represented as $\mathcal{C} = \{c_1, c_2, \dots, c_M\}$.

1. For a patient of N medications or conditions, first a distribution over topics is drawn from

$$\boldsymbol{\theta} \sim \text{Dir}(\boldsymbol{\alpha}),$$

where the probability density of a k -dimensional Dirichlet variable $\boldsymbol{\theta}$ is

$$p(\boldsymbol{\theta} | \boldsymbol{\alpha}) = \frac{\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)} \theta_1^{\alpha_1 - 1} \dots \theta_k^{\alpha_k - 1},$$

$\Gamma(x)$ is the Gamma function, $0 < \alpha_i < 1$, $\forall i$ and $\sum_{i=1}^K \theta_i = 1$.

2. For each medication or condition of the patient, a topic $x \in \{1, \dots, K\}$ is drawn from

$$x \sim \text{Mult}(\theta),$$

where $\text{Mult}(\theta)$ is a multinomial distribution characterized by θ and $p(x = i | \theta) = \theta_i$

3. Given the topic, medications or conditions of the patient are drawn from a conditional multinomial distribution $p(w|x; \phi)$, where ϕ is the parameter of the per-topic distribution over medications or conditions.

Several Bayesian inference techniques can be used to approximate the posterior distribution

$$p(\theta, x | w, \alpha, \phi) = \frac{p(\theta, x, w | \alpha, \phi)}{p(w | \alpha, \phi)}$$

in LDA including variational inference, expectation propagation and Gibbs sampling [19], [31], [32]. We can obtain the marginal distribution of a patient as

$$p(w | \alpha, \phi) = \int_{\theta} \left(\prod_{n=1}^N \sum_{x_n=1}^K p(w_n | x_n; \phi) p(x_n | \theta) \right) p(\theta; \alpha) d\theta$$

and take the product of the marginal probabilities of patients to derive the probability of the corpus [19]. In LDA, the Dirichlet is drawn for each patient, and within this patient the multinomial topic distribution over medication or conditions is drawn repeatedly [30].

The weights of posterior probabilities or proportions for different topics given patients' observed medications and conditions obtained from Bayesian inference are the feature representation in a reduced dimension (dimension equals to the number of topics). In our application, both clinical interpretability and topic coherence measure were taken into consideration and 5 topics were used. The LDA-based features were constructed based on these weights and they were augmented to the feature space in our analyses in learning the personalized treatments and confounding reduction. We compared using LDA-based features with using a simple summary feature of a total number of medications or ICD diagnoses.

III. Results

A. Cohort Identification

Our cohort includes patients who were older than 18 years and had at least one T2D diagnosis between 1/1/2008 and 12/31/2012. The analysis was further restricted to those patients who were prescribed with Metformin as first-line medication [16]. There were over 1,200 patients who were augmented with at least one second-line treatment. The sunburst plot in Figure 2 displays treatment sequence with 2-3 stages for T2D patients in the

database. The most inner circle corresponds to the first-line treatment and the most commonly used first stage treatment is Metformin, which is consistent with the T2D treatment guideline. Our goal is to select optimal ITRs to reduce major complications. The ITRs were estimated in terms of lowering major complications of T2D measured by three ICD diagnosis counts as ordinal outcomes (0, 1, 2, 3).

The EHR data preprocessing flowchart is shown in Figure 1. Most features were extracted from four domains: demographics, medication prescriptions, lab test and diagnosis. Several important lab tests (e.g., HbA1c, glucose, HDL, LDL, BMI) were summarized by estimated initial test values and estimated rate of change using linear mixed effect models. Initial test value is a measure of patient's illness condition at the baseline of study and rate of change of test value characterizes patient's progression rate of disease. Lab measurement pattern features were constructed using the information in the lab test domain [16]. Lab measurement patterns in EHRs were shown to reflect patient's underlying health status as well as EHR documentation biases and thus informative of adjusting for confounding and matching patients [33]. In Figure 3, the measurement patterns of glucose test and HbA1c in terms of gap days in logarithm scale is different for patients on different second-line treatments, suggesting it can distinguish between patients who started insulin (MET+insulin) and those who received glyburide or glipizide (MET+SFU). Therefore the lab measurement pattern features were used as important predictors for computing a balancing propensity score [16] in M-learning. The features capturing temporal measurement patterns were created by discretizing a two-dimensional space of lab test values versus gap time between consecutive measurements according to high, medium, low quartiles [8]. Additionally, age, race and gender were also included as demographic variables for matching patients.

B. LDA Feature Representation

Topic model with LDA using bag of words is applied to the generic names in co-medications and condition ICD9 codes prior to second-line treatment initiation (during baseline period) with 5 latent topics respectively. There were over 700 generic medication names and over 2,800 unique ICD9 codes in this T2D EHR data. In the co-medication domain, bigram and trigram were created for compound words in generic names. The number of topics is fixed in our application for interpretability purposes. LDA-based features were constructed as weights for each of the 5 topics per patient. We use heatmaps with hierarchical clustering to show topic feature weights in all patients. The similarity metric we used to cluster patients is Hellinger distance [29] defined as

$$\sum_{k=1}^K (\sqrt{\hat{\theta}_{d_1,k}} - \sqrt{\hat{\theta}_{d_2,k}})^2,$$

where K is the fixed number of topics and $\theta_{d,k}$ denotes weight for topic k in each patient. In Figure 4a, the heatmap for co-medication topic features suggests that most patients have higher weights in only one or two topics. Thus, co-medications for most patients comprise of one or two latent topics. Figure 4b shows patterns of ICD diagnosis condition, and a relatively high proportion of patients have larger weights in the first condition topic, which

mainly consists of conditions such as hypertension, pure hypercholesterolemia, depressive disorder, hyperlipidemia and lumbago.

To understand the learned topics, we present visualization of some of the medication and condition topics in Figure 5 - 7. These topics were found to be informative of predicting optimal treatments in our subsequent analyses (next section). On the right panel of each figure in Figure 5 - 7, the top 30 most relevant terms (co-medications or conditions) corresponding to the topics are presented. For example, topic #3 in the co-medication domain is a cluster of medications used to treat heart disease, high blood pressure and high cholesterol. Figure 5 shows that this topic contains 5 most relevant prescriptions as aspirin, metoprolol, atorvastatin, enalapril and clopidogrel. In condition topic #2, the most relevant conditions include hypertension, hypercholesterolemia, pain in limb, atrial fibrillation and hyperlipidemia, which are common co-morbidities of T2D. In condition topic #5, the top most relevant conditions are benign hypertension, hyperlipidemia, asthma and obesity.

Based on topics learned from LDA models, we constructed association networks for medications and conditions separately in Figure 8 and 9. The medications or conditions (nodes in the network) in the same topic with a posterior probability of co-occurring higher than a threshold are linked by an edge and shown in the figures. The threshold is 1.5% for co-medications and 0.9% for conditions for visualization. The edge widths are proportional to posterior probabilities. In the condition network, mental disorders including anxiety, depressive disorder and schizoaffective as well as some side effects such as cough, headache and backache are clustered in the rightmost group. Cardiovascular system related disorders conditions such as old myocardial infarction, mixed hyperlipidemia, atrial fibrillation and coronary atherosclerosis are connected together in the bottom cluster. Several topic clusters are connected through “bridge conditions/procedures” such as hypertension, unspecified hyperlipidemia, and flu vaccination which are associated with conditions in other clusters. In the co-medications network, several anti-diabetic or heart disease medications are clustered together including Rosuvastatin, Pioglitazone, Warfarin and Sitagliptin. Drugs used to treat co-morbidities are connected in a group such as Levothyroxine (treat thyroid disease), Esomeprazole (treat gastroesophageal reflux disease) and Atenolol (treat high blood pressure and chest pain). In this co-medications network for T2D, aspirin, as a blood thinner and anti-inflammatory drug, plays an important role in bridging medications from several clusters since aspirin is often co-prescribed with other medications.

C. Learning Optimal ITR

In our final analysis cohort, there were 740 subjects in total. The cohort was further divided into high baseline HbA1c level group and low baseline HbA1c level group according to median initial HbA1c level of 8.5% to account for patient heterogeneity. There were 380 patients in the low baseline group among whom 240 (63%) was prescribed SFU as second-line medication and 140 (37%) was prescribed insulin. In the high baseline group, the number of patients received SFU and insulin were 208 (58%) and 152 (42%) respectively. The outcome R_i used to minimize the value function (4) is count of post-treatment major complications (essential hypertension, hyperlipidemia and hypercholesterolemia) during the follow up period. We performed 2-fold cross-validation 100 times, and computed the cross-

validated empirical value function under three learning methods (Q-learning, O-learning and M-learning).

In Table I, we present the previous results in [16] with the raw counts from co-medication and diagnosis domain during the baseline period included as feature variables instead of using LDA-based features. In Table II, we replaced the raw counts of medications and ICD9 condition codes for each patient with our extracted LDA-based topic features. Comparing this with excluding LDA-based features, all three learning methods improve with a lower value function of post-treatment major complications counts and a smaller variability in most cases. In the high baseline group (see Figure 10a), the worst performing learning method using raw counts is Q-learning, and it benefited most from including FDA-based features. In the low baseline group (see Figure 10b), M-learning using LDA-based features achieves the lowest value function with a mean of 0.71 and a standard deviation of 0.028. In Figure 10, we also note that all three ITR learning methods outperform the “one-size-fits-all” strategies to assign insulin to all patients or SFU to all. This result demonstrates that prescribing treatment based on each patient’s characteristics reduces the number of post-treatment major complications compared to universal treatment prescription.

We examine the important topic features for selecting optimal treatment by fitting M-learning with a linear kernel SVM including LDA-based features in the cohort of 740 patients. There were 427 (58%) patients predicted to have “MET + SFU” as the optimal second-line medication and 313 (42%) to have “MET + insulin”. We show the absolute value of all the standardized coefficients in Figure 11. Besides race and cluster membership variable (based on hierarchical clustering analysis for patient heterogeneity using a subset of standardized features), co-medication topic #3, condition topic #5 and #2 rank as the most important topic features. The most relevant terms in these three topics are presented in Figure 5 - 7. Medication topic #3 involves past prescriptions of drugs used to treat heart diseases and its risk factors, while condition topic #5 and #2 includes major T2D co-morbidities. These results suggest that to recommend optimal second-line treatments, physicians may focus on past heart disease medication prescriptions and past diagnostic history of hypertension, hypercholesterolem, and atrial fibrillation.

IV. Discussion and Conclusion

In this paper, we investigate topic modeling with LDA as a feature extraction tool for EHR data and the use of a matching-based machine learning method for individualized treatment optimization. The learned features from LDA can be used in the ITR learning step to improve interpretability of the ITRs. Compared to treatment rules estimated excluding LDA-based features, our method achieves better post treatment clinical outcome and a smaller variability. In addition, LDA-based features summarize a large number of ICD codes and medications into a few understandable topics. By examining important topics predictive of optimal treatment, our method provides an interpretable summary of main useful characteristics a physician needs to focus on when making personalized medical decisions.

Several extensions can be considered. It is desirable to consider a data-driven method to select the best number of topics in LDA-based feature extraction. A hierarchical,

comprehensive LDA-based feature extraction combining different domains including medications, conditions and procedures altogether may be useful. Furthermore, correlated topic models can be considered to model the correlation between topics within each medical domain and dynamic topic models is useful in capturing the temporal characteristics of the evolution of medical concepts. Lastly, we can extend nearest-neighbor matching in M-learning to other optimal matching techniques such as a continuous matching method using local kernel weighting. These extensions will provide more flexibility to M-learning, especially for observational studies.

References

- [1]. Collins FS and Varmus H, "A new initiative on precision medicine," *New England Journal of Medicine*, vol. 372, no. 9, pp. 793–795, 2015. [PubMed: 25635347]
- [2]. Fried EI and Nesse RM, "Depression is not a consistent syndrome: an investigation of unique symptom patterns in the star* d study," *Journal of affective disorders*, vol. 172, pp. 96–102, 2015. [PubMed: 25451401]
- [3]. Chakraborty B and Moodie EE, *Statistical methods for dynamic treatment regimes*. New York: Springer-Verlag, 2013.
- [4]. Cole SR and Stuart EA, "Generalizing evidence from randomized clinical trials to target populations: The actg 320 trial," *American journal of epidemiology*, vol. 172, no. 1, pp. 107–115, 2010. [PubMed: 20547574]
- [5]. Weng C, Li Y, Ryan R, Zhang Y, Liu F, Gao J, Bigger J, and Hripcsak G, "A distribution-based method for assessing the differences between clinical trial target populations and patient populations in electronic health records," *Applied clinical informatics*, vol. 5, no. 2, p. 463, 2014. [PubMed: 25024761]
- [6]. Hripcsak G, Ryan PB, Duke JD, Shah NH, Park RW, Huser V, Suchard MA, Schuemie MJ, DeFalco FJ, Perotte A, and etc, "Characterizing treatment pathways at scale using the OHDSI network," *Proceedings of the National Academy of Sciences*, vol. 113, no. 27, pp. 7329–7336, 2016.
- [7]. Rosenbaum PR, *Design of observational studies*. Springer, 2010, vol. 10.
- [8]. Wang Y, Wu P, Liu Y, Weng C, and Zeng D, "Learning optimal individualized treatment rules from electronic health record data," 2016 IEEE International Conference on Healthcare Informatics (ICHI), pp. 65–71, 2016.
- [9]. Watkins CJ and Dayan P, "Q-learning," *Machine learning*, vol. 8, no. 3-4, pp. 279–292, 1992.
- [10]. Qian M and Murphy SA, "Performance guarantees for individualized treatment rules," *The Annals of Statistics*, vol. 39, no. 2, p. 1180, 2011. [PubMed: 21666835]
- [11]. Murphy SA, "Optimal dynamic treatment regimes," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 65, no. 2, pp. 331–355, 2003.
- [12]. Zhao Y, Zeng D, Rush AJ, and Kosorok MR, "Estimating individualized treatment rules using outcome weighted learning," *Journal of the American Statistical Association*, vol. 107, no. 499, pp. 1106–1118, 2012. [PubMed: 23630406]
- [13]. Liu Y, Wang Y, Kosorok MR, Zhao Y, and Zeng D, "Augmented outcome-weighted learning for estimating optimal dynamic treatment regimens," *Statistics in Medicine*, vol. 37, no. 26, pp. 3776–3788, 2018 [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.7844> [PubMed: 29873099]
- [14]. Dusseldorp E and Van Mechelen I, "Qualitative interaction trees: a tool to identify qualitative treatment–subgroup interactions," *Statistics in medicine*, vol. 33, no. 2, pp. 219–237, 2014. [PubMed: 23922224]
- [15]. Laber E and Zhao Y, "Tree-based methods for individualized treatment regimes," *Biometrika*, vol. 102, no. 3, pp. 501–514, 2015. [PubMed: 26893526]

- [16]. Wu P, Zeng D, and Wang Y, “Matched learning for optimizing individualized treatment strategies using electronic health records,” *Journal of the American Statistical Association*, vol. 0, no. 0, pp. 1–23, 2019 [Online]. Available: 10.1080/01621459.2018.1549050
- [17]. Landauer TK, Foltz PW, and Laham D, “An introduction to latent semantic analysis,” *Discourse Processes*, vol. 25, no. 2-3, pp. 259–284, 1998 [Online]. Available: 10.1080/01638539809545028
- [18]. Hofmann T, “Probabilistic latent semantic analysis,” in *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*, ser. UAI’99. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1999, pp. 289–296. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2073796.2073829>
- [19]. Blei DM, Ng AY, and Jordan MI, “Latent dirichlet allocation,” *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, 3 2003 [Online]. Available: <http://dl.acm.org/citation.cfm?id=944919.944937>
- [20]. Zhou X, Wang Y, and Zeng D, “Outcome-weighted learning for personalized medicine with multiple treatment options,” in *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA) IEEE*, 2018, pp. 565–574.
- [21]. Lee BK, Lessler J, and Stuart EA, “Improving propensity score weighting using machine learning,” *Statistics in Medicine*, vol. 29, no. 3, pp. 337–346, 2010. [PubMed: 19960510]
- [22]. —, “Weight trimming and propensity score weighting,” *PloS One*, vol. 6, no. 3, p. e18174, 2011. [PubMed: 21483818]
- [23]. Antonelli J, Cefalu M, Palmer N, and Agniel D, “Doubly robust matching estimators for high dimensional confounding adjustment,” *Biometrics*, vol. 74, no. 4, pp. 1171–1179, 2018 [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1111/biom.12887> [PubMed: 29750844]
- [24]. Johnson SB, “Generic data modeling for clinical repositories,” *Journal of the American Medical Informatics Association*, vol. 3, no. 5, pp. 328–339, 1996. [PubMed: 8880680]
- [25]. Weiskopf NG and Weng C, “Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research,” *Journal of the American Medical Informatics Association*, vol. 20, no. 1, pp. 144–151, 2013. [PubMed: 22733976]
- [26]. Fort D, Weng C, Bakken S, and Wilcox AB, “Considerations for using research data to verify clinical data accuracy,” *AMIA Summits on Translational Science Proceedings*, pp. 2014: 211–217, 2014.
- [27]. Diabetes Control and Complications Trial Research Group, “The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus,” *New England Journal of Medicine*, vol. 329, no. 14, pp. 977–986, 1993. [PubMed: 8366922]
- [28]. Bennett WL, Maruthur NM, Singh S, Segal JB, Wilson LM, Chatterjee R, Marinopoulos SS, Puhon MA, Ranasinghe P, Block L et al., “Comparative effectiveness and safety of medications for type 2 diabetes: an update including new drugs and 2-drug combinations,” *Annals of internal medicine*, vol. 154, no. 9, pp. 602–613, 2011. [PubMed: 21403054]
- [29]. Blei DM and Lafferty JD, *Topic Models*. CRC Press, 2009.
- [30]. Blei DM, Ng AY, and Jordan MI, “Latent dirichlet allocation,” in *Advances in Neural Information Processing Systems 14*, Dietterich TG, Becker S, and Ghahramani Z, Eds. MIT Press, 2002, pp. 601–608. [Online]. Available: <http://papers.nips.cc/paper/2070-latent-dirichlet-allocation.pdf>
- [31]. Geman S and Geman D, “Stochastic relaxation, gibbs distributions, and the bayesian restoration of images,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, 11 1984 [Online]. Available: 10.1109/TPAMI.1984.4767596 [PubMed: 22499653]
- [32]. Minka T and Lafferty J, “Expectation-propagation for the generative aspect model,” in *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence*, ser. UAI’02. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2002, pp. 352–359. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2073876.2073918>
- [33]. Pivovarov R, Albers DJ, Sepulveda JL, and Elhadad N, “Identifying and mitigating biases in ehr laboratory tests,” *Journal of biomedical informatics*, vol. 51, pp. 24–34, 2014.

Fig. 1.
EHR Data Preprocessing Chart for Different Domains

Fig. 2. Sunburst plot of treatment sequence for T2D patients*

*: Layer corresponds to stage of treatment, patients with 2-3 stages were shown

Fig. 3.
Longitudinal Measurement Pattern (Gap Days) of HbA1c vs. Glucose

(a) Co-medication Topic Features

(b) Condition Topic Features

Fig. 4.
Heatmap of LDA-Features Clustered by Patients

Fig. 5.
Co-medication Topic #3

Fig. 6.
Condition Topic #2

Fig. 7.
Condition Topic #5

Fig. 8.

Association Network of Medication Prescriptions Based on Topics in LDA Model

(a) High Baseline Group

(b) Low Baseline Group

Fig. 10.

Empirical value function of ICD diagnosis count in EHR data with 100 2-fold cross-validations (a low value is desirable)

Fig. 11.
Feature Importance in M-learning with a Linear Kernel

TABLE I

Cross-validated Empirical Value Function for the Number of Major Complications Using Raw Count Data

High Baseline Group		
ITR Method	Mean (Std)	Median (Q1, Q3)
Q-Learning	0.88 (0.078)	0.88 (0.83, 0.93)
O-Learning	0.86 (0.050)	0.87 (0.83, 0.90)
M-Learning	0.84 (0.068)	0.83 (0.80, 0.88)
Low Baseline Group		
ITR Method	Mean (Std)	Median (Q1, Q3)
Q-Learning	0.81 (0.063)	0.81 (0.76, 0.85)
O-Learning	0.72 (0.033)	0.72 (0.70, 0.74)
M-Learning	0.73 (0.032)	0.73 (0.71, 0.75)

TABLE II

Cross-validated Empirical Value Function for the Number of Major Complications Using LDA Features

High Baseline Group		
ITR Method	Mean (Std)	Median (Q1, Q3)
Q-Learning	0.77 (0.065)	0.77 (0.73, 0.81)
O-Learning	0.81 (0.045)	0.81 (0.79, 0.84)
M-Learning	0.83 (0.088)	0.81 (0.78, 0.86)
Low Baseline Group		
ITR Method	Mean (Std)	Median (Q1, Q3)
Q-Learning	0.77 (0.048)	0.77 (0.73, 0.80)
O-Learning	0.72 (0.029)	0.72 (0.70, 0.74)
M-Learning	0.71 (0.028)	0.72 (0.70, 0.73)