Towards Systematic Evaluation of Interventions in Major Depressive Disorder

By

Barrett Jones

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

Master of Science

In

Biomedical Informatics

May 13, 2022

Nashville, Tennessee

Approved:

Colin G. Walsh, MD, MA, FAMIA

Warren Taylor, MD, MHSc

Daniel Fabbri, PhD, FAMIA

Acknowledgements and Funding

This work was funded by the National Library of Medicine Training Grant T15 LM007450-17. The work included in chapter 1 has been published in the proceedings of the 2021 American Medical Informatics Association Annual Symposium.

Table of Contents

List of Tables

List of Figures

Introduction

A learning health system "learns as quickly as possible about the best treatment for each patient—and delivers[1]." A vital function of the learning health system is the ability to systematically evaluate interventions[2]. The gold standard for intervention evaluation is the randomized control trial (RCT). Sequential Multiple Assignment Randomized Trial (SMART) trail designs are a type of RCT developed to imitate real world practices and a step towards systematic evaluation of interventions[3]. The Sequenced Treatment Alternatives to Relieve Depression (STAR*D) is an example of a SMART trial design to evaluate interventions in Major Depressive Disorder (MDD). This was a highly impactful trial, but limited in ability to evaluate intervention effectiveness at the various levels of the trial[4]. Electronic Health Records have improved the ability to conduct RCTs, by easing patient recruitment and tracking. However, RCTs can be high cost[5], and results may not replicate[6]. Advances in techniques for data collection and analysis, including the electronic health record (EHR), allow for the ability to evaluate interventions in an observational real-world setting.

Dynamic treatment regimes (DTR) are statistical methods able to model SMART trial designs like the STAR*D study[7–10]. DTRs model sequences of decision rules to make intervention recommendations accounting for dynamic nature of the patient environment[11]. DTR models require accurate and complete representation of the patient environment including the patient state, clinician actions and clinically relevant outcomes. This requirement poses challenges in the study of observational EHR data. In observational data, interventions are not randomly assigned and subject to confounding. Additionally, outcomes in MDD trials are often dependent on surveys that are not regularly collected in the EHR. In order to model DTRs with observational EHR data, we must develop accurate representations of patient state, clinician actions, and outcomes.

In chapter 1 we characterize medication treatment pathways to provide insight into clinician action trends at Vanderbilt University Medical Center (VUMC) in the treatment of MDD. In this study we use a Long-Short Term Memory (LSTM) autoencoder model to systematically characterize treatment pathways in MDD. LSTM autoencoder models generate representations of medication treatment pathways that account for temporality and complex interactions. Patients with similar pathways are clustered by the K-means algorithm. Clusters are characterized by analysis of medication utilization sequences and trends, as well as clinical features, such as demographics, outcomes and comorbidities. Cluster characterization identifies differences in medication utilization, medication heterogeneity, temporal trends, and clinical features between clusters. We find that treatment trajectories are associated with MDD endotypes including severe acute, low utilization moderate, and high utilization, chronic, severe, but managed.

The severe acute MDD endotype is relatively unmanaged and has higher rates of admission and self-harm/suicide attempt. In chapter 2 we begin to study the effectiveness of clinician intervention on avoiding these outcomes. Additionally, we generate representations for patient states, clinician actions, and outcomes and assess their utility. MDD is a prevalent phenotype commonly treated by forms of cognitive therapy, medication, or some combination. In rare cases patients may be referred for electroconvulsive therapy (ECT), or partial hospitalization. We study this wider range of clinician actions in chapter 2 and measure associations between clinician actions and outcomes in MDD. We hypothesize that clinician actions studied will have protective effects against outcomes indicative of severe depression after adjustment for patient state. We study patients with an indication of MDD in the EHR by measuring the association between clinician actions taking place during the first 90 days post indication of MDD and outcomes taking place in the subsequent 180 days. We adjust for patient state features including demographics, severity proxies and comorbidities. We then perform sensitivity analyses of model results.

We find that clinician actions have an increased risk association with outcomes. One clinician action outcome pair had a statistically significant reduction in risk—prescribing an initial antidepressant with unplanned admission. However, five action variables exhibited statistically significant associations with increased risk of an outcome after adjustment for state variables. Patient state representation is limited to what is available in the EHR and potential confounders are likely. This finding is consistent through our sensitivity analysis. The results of this study suggest potential confounding of intervention effect estimation. Further study of causal inference methods is necessary in this population.

The study of medication treatment trajectories in chapter 1 provides insight into discrete treatment pathways that appear to correspond to endotypes of MDD. This study informs our understanding of how treatment of MDD progresses. In chapter 2 we begin to explore evaluation of clinician interventions, focusing on initial clinician actions after diagnosis. We identify shortcomings with the data available and methods utilized that prompts further exploration of causal inference methods. Also, MDD is a chronic illness, so future studies will evaluate sequences of interventions over time extending beyond the initial interventions—building on the findings in chapter 1 and leveraging advances in DTR research.

Chapter 1-Characterization of medication treatment trajectories in Major Depressive Disorder

Background and Significance

An important step towards systematic evaluation is characterization of in place treatment pathways[2]. In this study, we perform a systematic characterization of treatment pathways of MDD from observational EHR data at VUMC. Clinical trials have shown the effectiveness of antidepressants in treatment of MDD relative to placebo[12], but there is a need for better understanding and representation of long term treatment practices[13–15]. Efforts to standardize treatment pathways have resulted in decreased variance in treatment of MDD over time, but heterogeneity of treatment pathways remains between institutions[2]. Due to the prevalence of MDD and impact on daily life, characterization of treatment pathways in MDD may be beneficial to healthcare organizations in improving outcomes and quality of care[16,17].

Past studies have modeled treatment paths by visualizing medication sequences[2,16]. We built upon these studies by modeling medication treatment pathways in the VUMC Synthetic Derivative[18] accounting for temporality as well as treatment resistant depression (TRD). TRD is defined as failure to respond to two or more antidepressant trials[19]. TRD is of particular interest in this study for the potential use of treatment pathways in prediction of TRD in future studies. In this study, electroconvulsive therapy (ECT) was used as a surrogate for TRD[20].

LSTM neural networks are deep learning algorithms that have the ability to model temporal data with complex relationships like those characteristic of medication data. An LSTM is a form of Recurrent Neural Network that stores information over extended time intervals and employs a gating method to address the exploding gradient problem found in some Recurrent Neural Network applications[21]. Autoencoder models have the ability to generate simplified encodings of complex data structures. Autoencoders are composed of two sub-models: an encoder and decoder. The encoder receives an input data set and reduces the dimensionality to a hidden layer. The decoder takes as input the hidden layer from the encoder model and reconstructs the input data[22,23]. The hidden layer generated by the encoder model is a denoised continuous vector representation of the model input. LSTM autoencoders have been used to learn representations for video sequences[24] and for biomedical endotyping[25]. In this study, we hypothesize that LSTM autoencoders can effectively represent medication treatment pathways and that characterization of encoding clusters will allow for differentiation of clusters by medication treatment pathways and clinically relevant variables.

Methods

Data Description

The patient cohort for this study included patients aged 18 to 90, diagnosed with MDD seen at least two times, six months apart, at VUMC located in the United States Mid-South in Nashville, TN. Patients with MDD were identified that had at least one of the following ICD-9 codes: 311.x, 296.2x, 296.3x, 300.4x or ICD-10 codes: F32.xx, F33.xx, F34.1 and their index diagnosis took place prior to 12/31/2016 to ensure a minimum of three years of medication data. We excluded any patients with a diagnosis of Bipolar Disorder or Schizophrenia and those that were not prescribed an antidepressant (see Table 1.5) within three years of diagnosis.

Medications were extracted for three years following first MDD diagnosis and subset to include only prescriptions. Because the population with TRD was of particular interest for downstream prediction of TRD status before it occurred, the cohort with MDD who received ECT, a surrogate for TRD in some studies[20], had medication data censored one day prior to ECT treatment date. Medications were grouped using the World Health Organization Anatomical Therapeutic Chemical (ATC) classification level 5[17]. Medications included in the study are listed in the Appendix Table A.1.

Unsupervised Feature Learning

Medication data were converted to a categorical quarterly time series with indicators for each medication class if observed during a quarter. The resulting dataset had dimension N x M x T, with N the number of patients, M the number of medication classes, and T the number of quarters in our observation window.

An LSTM autoencoder was fit on the medication data to generate a continuous-multidimensional representation of each patient's medication pathway. The model was built using the Keras Python library[26]. We tested multiple model structures in order to tune the number of encoder layers as well as the dimension of the encoder layers. We performed a grid search on 15 candidate models with one to three encoder layers and final encoder layer dimensions of 8, 16, 24, 32 and 64. Models were evaluated based on reconstruction mean squared error (MSE) on a 25% validation test set. Model selection was based on visual inspection of plots with encoding dimension on the x axis and MSE on the y. The model with an MSE such that increasing encoding dimension does not meaningfully reduce MSE was selected for clustering.

The encodings of the selected model were clustered using the K-means algorithm. The Elbow Method was utilized for selection of K—the number of clusters. In this method the sum squared distance from the cluster centroid is plotted against K and the "elbow" is selected by visual inspection as the point at which increasing K does not meaningfully reduce sum squared distance from the centroid[27].

Cluster Characterization

We characterized clusters by performing an analysis of demographics, clinical outcomes, comorbidities, and utilization trends. Demographics include gender, age and race. The clinical outcomes measured in the study are all cause mortality, admission, ER Visit and a TRD surrogate—ECT –within the three-year period of analysis. Comorbidities are identified using Agency for Healthcare Research and Quality (AHRQ) Clinical Classification Software (CCS)[28] to map ICD 9 and 10 codes to clinically meaningful comorbidities. We perform a Chi-square test of association between clusters and categorical demographics—gender and race—and perform ANOVA for continuous variables—age and comorbidity count—with post-hoc Bonferroni corrected confidence intervals to measure differences between clusters. Mental health comorbidity association with clusters were also analyzed by Chi-square with post-hoc Bonferroni corrected significance tests to identify clusters with enriched comorbidity prevalence.

Medication and visit utilization analysis included: calculating the mean number of prescriptions per patient, mean unique number of prescription classes per patient and single medication treatment rates—the proportion of patients prescribed a single medication class for the entire period of analysis. Trend graphs visualized three-quarter rolling averages of medication utilization. By cluster medication sequences were visualized by sunburst plots. For visualization purposes medications were combined into clinically meaningful groupings as defined by physicians from the VUMC department of psychiatry. These groupings are available in Appendix Table A.1. Additionally, we account for instances in which patients are prescribed multiple medications in a short time period. We will refer to these cases as Multi-medication-therapy and define it as any instance in which a patient is prescribed two different mediations (ATC 5 level) within a five-day period.

Results

Model Selection

Autoencoder models were fit and validation MSE was plotted by encoding dimension for each of the model structures in Figure 1. The three-layer encoder model outperforms the one and two encoder layer models at each candidate encoding dimension. Validation MSE in the three-layer encoder model is 0.0034 with encoding dimension equal to 24. At dimension 32 and 64 the validation MSE is 0.0032 and 0.0020 respectively. As can be seen in Figure 1, the rate at which the three-layer encoder model MSE (green) is decreasing is reduced at dimension 24. The model with three encoding layers and encoding layer dimension of 24 is selected for our final model. Encodings from the selected model were cluster using the K-means algorithm. The elbow method was utilized for selection of K—the number of clusters. Visual inspection of the Figure 1 (right) indicates that the "elbow" occurs at K=5.

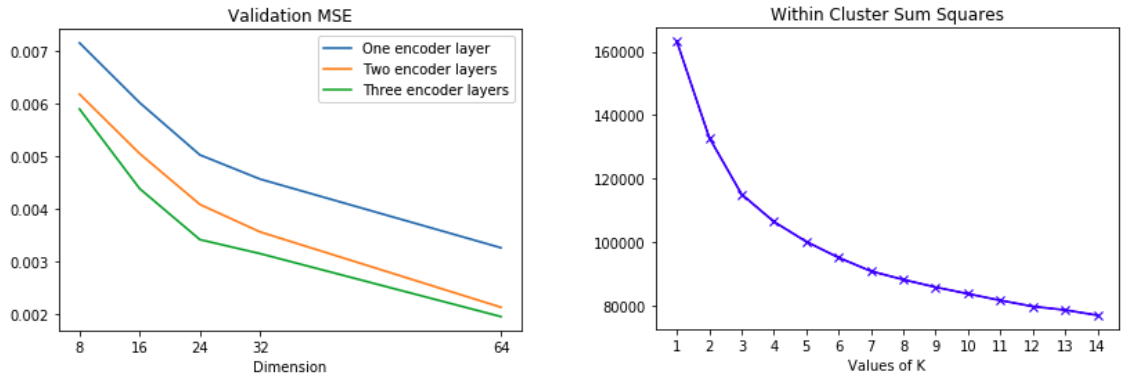

Figure 1.1. Autoencoder and K-means clustering model selection plots. The reconstructed MSE for each candidate model is plotted with lines colored by the number of encoding layers (left). The within cluster sum squares are plotted for the K-means algorithm with K=1 through K=14 (right).

Cluster Characterization

Summary statistics by cluster were calculated for demographics, outcomes and comorbidities. 46,454 patients met the inclusion criteria for this study. Five clusters were identified, cluster 2 being the largest (n=13,908, 29.1%). ANOVA and Chi-square tests were conducted for association between clusters and variables of interest. For each of the variables tested differences between clusters were found to be statistically significant. Post-hoc Bonferroni corrected confidence interval margin of errors are included in Table 1.1. The majority of patients in the study are female (66.4%) and Cluster 4 is higher proportion female (70.4%, 95% CI = 69.4%-71.4%). Cluster 5 has elevated mortality (3.6%, 95% CI = 3.3%-3.9%) and Cluster 2 has elevated admissions (66.6%, 95% CI = 66.0%-67.2%) and ECT (0.56%, 95% CI = 0.46%-0.66%). Cluster 3 has the highest CCS comorbidity count on average (17.6, 95% CI = 17.3-17.6).

Table 1.1. Clinical characteristics by cluster. For each metric in the table the margin of error for a Bonferroni corrected 95% confidence interval is included in parentheses.

		Cluster						
		1	**2**	**3**	**4**	**5**	**P-value**	**Total**
Demographics	**N**	6,238	13,908	10,616	5,112	10,580		46,454
	% Female	68.2 (0.9)%	64.5 (0.6)%	67.4 (0.7)%	70.4 (1.0)%	65.1 (0.7)%	<0.001	66.4%
	Mean Age	48.8 (0.6)	45.8 (0.4)	48.5 (0.5)	47.6 (0.6)	47.1 (0.5)	<0.001	47.3
Race/Ethnicity	**Black/Non-Hispanic**	9.6 (0.6)%	10.3 (0.5)%	8.6 (0.5)%	10.9 (0.8)%	8.7 (0.5)%		9.5%
	Black/Hispanic-Latino	0.0 (0.0)%	0.1 (0.1)%	0.1 (0.1)%	0.1 (0.1)%	0.1 (0.1)%		0.1%
	White/Non-Hispanic	85.6 (0.8)%	83.7 (0.5)%	86.9 (0.6)%	84.9 (0.9)%	85.6 (0.6)%	<0.001	85.3%
	White/Hispanic-Latino	1.2 (0.2)%	1.7 (0.2)%	1.4 (0.2)%	1.2 (0.3)%	1.5 (0.2)%		1.5%
	Other	3.7 (0.2)%	4.3 (0.2)%	3.0 (0.2)%	2.9 (0.3)%	4.0 (0.2)%		3.7%
Outcomes	**Mortality %**	2.4 (0.3)%	1.5 (0.2)%	2.2 (0.2)%	1.3 (0.3)%	3.6 (0.3)%	<0.001	2.2%
	Admission %	42.1 (1.0)%	66.6 (0.6)%	45.1 (0.8)%	42.6 (1.1)%	47.9 (0.8)%	<0.001	51.5%

		Cluster						
		1	**2**	**3**	**4**	**5**	**P-value**	**Total**
	ER Visit %	28.9 (0.9)%	27.9 (0.6)%	28.5 (0.7)%	29.5 (1.0)%	27.3 (0.7)%	0.024	28.2%
	ECT %	0.24 (0.1)%	0.56 (0.1)%	0.20 (0.07)%	0.17 (0.09)%	0.27 (0.08)%	<0.001	
Comorbidities	**Mean CCS Comorbidity Count**	16.0 (0.3)	13.3 (0.2)	17.6 (0.3)	15.6 (0.3)	16.7 (0.3)	<0.001	15.7

Medication utilization statistics by cluster are included in Table 1.2 and trend plots of medication and visit utilization in Figures 1.2 and 1.3. Clusters 2 and 4 have the lowest mean count of prescriptions (3.0, 3.6) during the period of analysis and the lowest number of unique prescriptions by medication class (1.5, 1.4) as well as the highest rates of single medication treatment (66.4%, 68.9%). Clusters 3 and 5 have the highest prescription counts (12.4, 9.8) and the lowest rates of single medication treatment (31.7%, 37.2%).

Table 1.2. Medication utilization summary statistics by cluster.

	Cluster				
	1	**2**	**3**	**4**	**5**
N	6,238	13,908	10,616	5,112	10,580
Mean prescription count	5.5	3.0	12.4	3.6	9.8
Mean unique prescription class count (ATC Level 5)	1.7	1.5	2.3	1.4	2.2
Single Medication Treatment Rates	58.5%	66.4%	31.7%	68.9%	37.2%

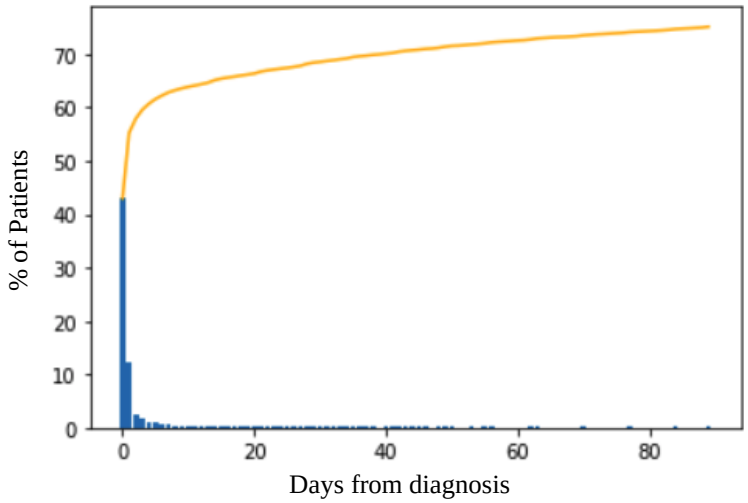

Figure 1.2. Pareto plot of time to antidepressant prescription from diagnosis. Histogram of density (blue) and cumulative density (orange).

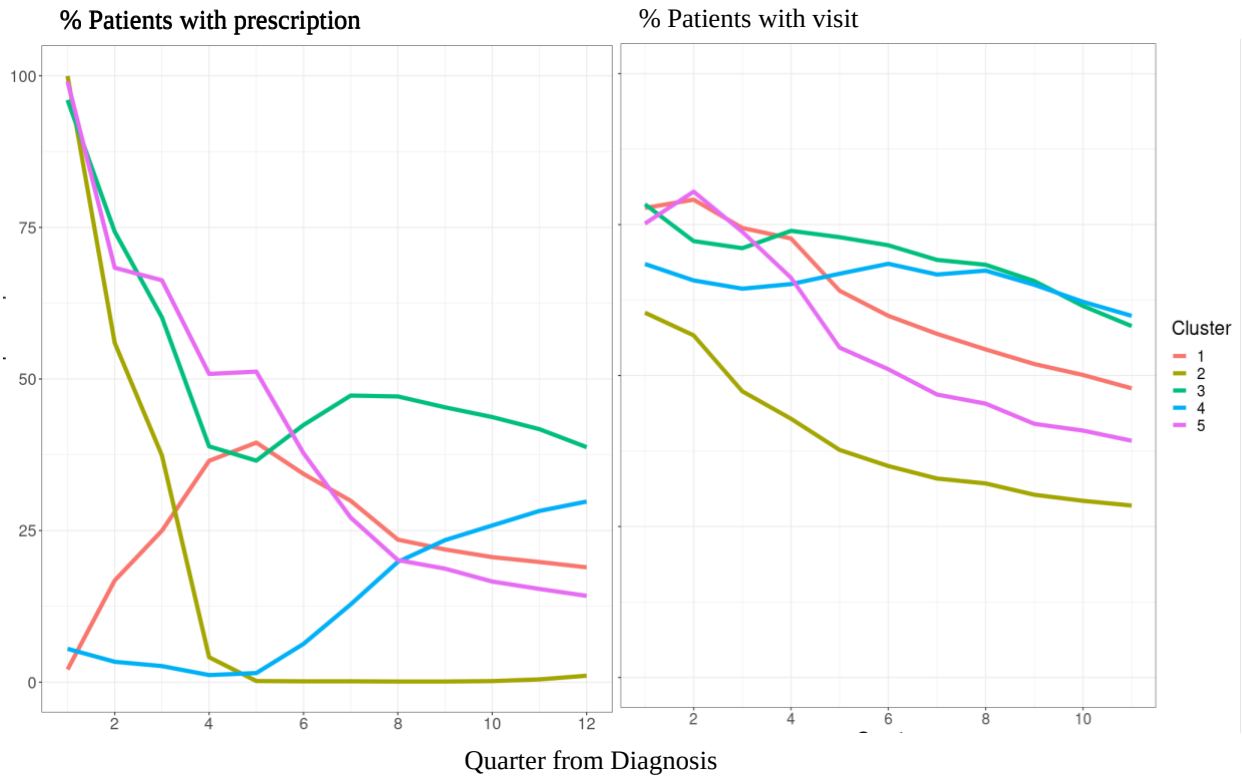

Figure 1.3. Medication and visit utilization trends. Three quarter moving average plots of the percentage of patients with a prescription (left) and visit (right) during each quarter of the period of analysis by cluster.

The most common initial medication type prescribed was selective serotonin reuptake inhibitors (SSRI) in the patients studied (50.1%). This ranged from 47.7% in cluster 4 up to 51.4% in cluster 3. For patients receiving an SSRI initially, the majority of these patients did not receive another prescription within the period of analysis (66.2%) and 9.1% were prescribed a SSRI as their second antidepressant. In our study population 18.4% of patients received multi-medication-therapy. Rates of multi-medication-therapy range between 11.5% in cluster 1 and 22.6% in cluster 2.

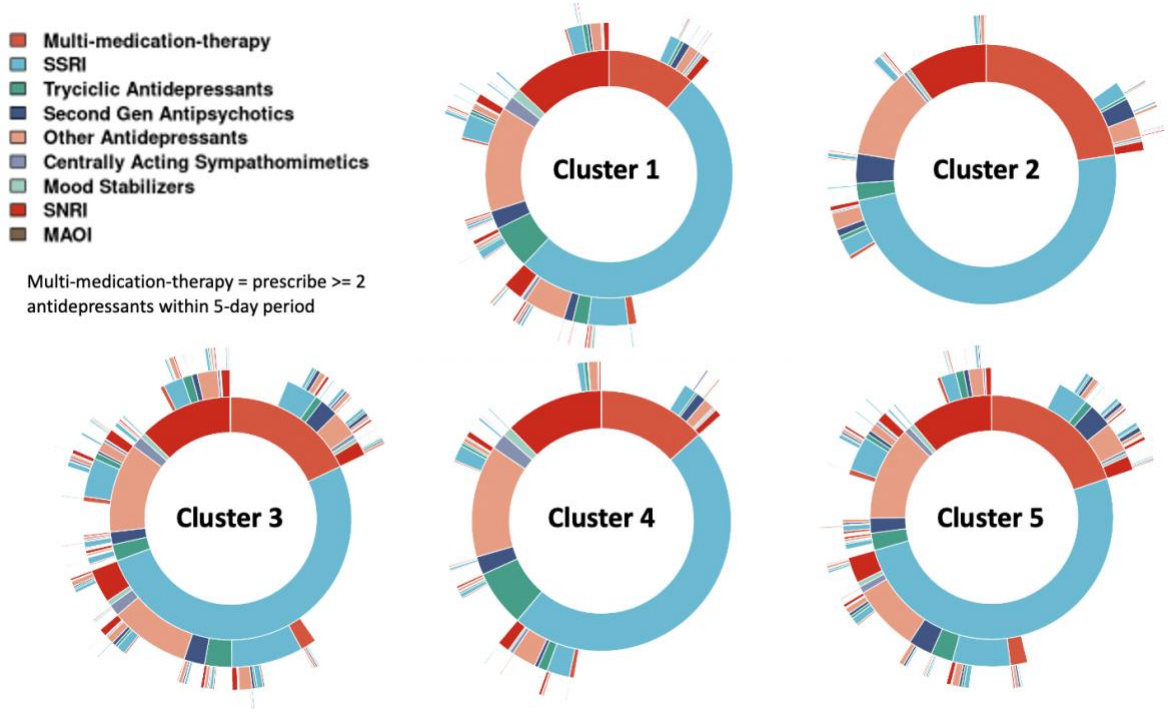

Figure 1.4. Sunburst plot of medication sequences by cluster. The first (inner) level of the plot represents the distribution of initial antidepressant types with the second level representing the second antidepressant type prescribed and so on.

The Chi-square analysis of association between mental health comorbidities and clusters was conducted to identify comorbidity-cluster pairs that deviated from expected frequency given the overall comorbidity frequency within the population and the cumulative comorbidity burden within the cluster. Results of the analysis are included in Table 1.3.

Table 1.3. Chi-square analysis of mental health comorbidities by cluster. Chi-square analysis to test association between cluster and mental health comorbidities. Post-hoc Bonferroni corrected significance tests of the Chi-square statistic are performed for each comorbidity-cluster pair at the 90%, 95%, and 99% significance levels (indicated by color).

	Cluster				
Comorbidity	**1**	**2**	**3**	**4**	**5**
mean CCS comorbidity count	16	13.3	17.6	15.6	16.7
anxiety disorders	41.6%	36.3%	50.3%	40.6%	47.8%
history of mental health and substance abuse	10.7%	13.3%	10.8%	11.3%	11.8%
substance related disorders	9.2%	12.4%	11.9%	10.1%	12.4%
adjustment disorders	10.6%	5.7%	12.1%	10.5%	11.4%
delirium dementia and amnestic/other cognitive disorders	8.6%	5.6%	8.8%	8.7%	9.1%
miscellaneous mental health disorders	8.0%	6.5%	7.8%	8.1%	7.7%
suicide and intentional self-inflicted injury	4.2%	10.7%	6.0%	4.6%	7.7%
attention deficit conduct and disruptive behavior disorders	5.7%	5.2%	7.7%	5.5%	6.3%
alcohol related disorders	4.2%	6.1%	5.2%	5.7%	5.7%
miscellaneous mental disorders	2.7%	1.6%	4.0%	1.9%	3.2%

Significance Level
99%
95%
90%
significant positive association
99%
95%
90%
significant negative association

Discussion

Overall, we find that the unsupervised learning methods are able to separate the study population into clusters using medication data alone. We find differences in medication utilization, medication heterogeneity, temporal trends, as well as statistically significant differences between clusters in clinical features. This work builds on the characterization of treatment pathways research[2,16,29] by leveraging machine learning methods that account for temporality and extending characterization to analyze relevant clinical features associated with treatment pathways.

Cluster characterization analysis examined three dimensions to describe the autoencoder and clustering outputs: utilization, medication heterogeneity, and temporal trends. Table 1.2 displays both utilization (mean prescription count) and medication heterogeneity (mean unique prescriptions, single medication treatment rates). The medication sequences in the sunburst plots (Figure 1.4) visualize medication heterogeneity between clusters and the quarterly moving average plots of percent patients with prescription and visit provide insight into temporal trends (Figure 1.3). In differentiating between clusters, medication utilization trends are a key factor. The profiles of clusters 3 and 5 are similar in aggregate medication utilization statistics, as well as medication sequences. However, medication utilization trends downward initially after diagnosis, but utilization sustains a consistent level in quarters 6 through 12 from MDD diagnosis. Conversely, cluster 5 medication utilization trends downward throughout the period of analysis.

In addition, we describe the clusters by clinical features such as demographics, comorbidities and outcomes. Differences between each of the demographics measured are found to be statistically significant between clusters. Outcomes measured include mortality, admission, ER visits, and ECT. Themes emerge from characterization analysis. For example, cluster 2 is characterized by higher rates of substance use disorders and self-harm as well as higher provision of ECT. These factors suggest this cluster captures high acuity psychiatric patients with depression at our medical center. This finding is also supported by medication treatment patterns, cluster 2 has high initial medication utilization and the highest rate of multi-medication treatment among the clusters. Clusters 1 and 4 are characterized by low utilization, outcomes and comorbidity profile. We suggest patients in these clusters have a moderate severity chronic MDD. Clusters 3 and 5 convey high utilization, and high comorbid burden with relatively low rates of substance abuse and self-harm. We assert clusters 3 and 5 are patients with severe, chronic, but managed MDD.

Our study is limited to analysis of data from a single institution. Our institution is an open system with finite specialty mental health access: patients included in our study might receive care at an institution other than VUMC or the patients in this study might be incident users of antidepressants[30]. We seek to limit this risk by requiring two visits to VUMC at least six months apart. Additionally, there are limitations in representation of the medication data. Our methods seek to address some of those limitations by representing medications as a time-series and leveraging an LSTM auto-encoder. In this work, we do not explicitly seek to capture concurrent medications. However, concurrent medications may be implicitly captured to some degree by the quarterly time series representation. In this study we limit to including only medications from VUMC at the time they are prescribed. Information about duration of prescription, adherence, or dose changes are not included in our model but remain germane to effective treatment of MDD.

Future work should seek to expand treatment pathway characterization to additional health systems and to analyze additional phenotypes. In addition, health information exchanges can provide a more complete picture of the patient health record[31].

Conclusion

We clustered the output of an LSTM autoencoder on time-series of antidepressant prescriptions in patients diagnosed with MDD. We identified common, clinically relevant patterns in prescribing practices. Cluster characterization established that prescribing practices differ on multiple dimensions: utilization, medication heterogeneity, temporal trends. Clusters are found to be associated with a clinically meaningful features including demographics, outcomes, and comorbidities. This method provides insight into endotypes of MDD.

Chapter 2-Modeling association between clinician interventions and outcomes in Major Depressive Disorder with observational Electronic Health Record data

Background and Significance

Prior EHR studies in MDD have characterized treatment trajectories[2] and developed predictive models of outcomes[13,14,32,33]. While these models often consider interpretability of model parameters, few are intended to make inferences about effects of clinical interventions. Such approaches have the potential to provide new intervention insights, as in a recent study that used observational EHR data to model COVID-19 intervention effects[34]. Observational studies are essential in medication safety research[35]. Examples include studies of the effects of Methylphenidate on suicide attempts and psychotic disorders[36,37], and the risk of bone fractures in hemodialysis patients treated with proton pump inhibitors[38]. Our current study builds on prior work by mining the EHR for features predictive of depression outcomes and fitting a statistical model to study associations between interventions and outcomes.

Regression analysis is a longstanding method for estimation of intervention effects[39,40]. It allows for estimation with adjustment for covariates that may confound the effects of the study. Selection of potential confounder variables is a non-trivial task—as the confounders are not always measured and adjustment for extraneous variables can reduce study power and introduce bias[41,42]. In observational studies the risk for confounding is always high and causal relationships between observed variables may be unknown. Sensitivity analyses provide insight into the validity of model performance and confounding risk[43]. Additionally, causal inference researchers have developed methods with the ability to adjust for unobserved confounding under certain circumstances[44–47].

Objective

In our study, we collaborated with clinical experts to understand and model interventions in MDD. We hypothesized that by integrating clinical expertise into data collection and model construction we could reduce confounding due to the observational nature of our study. The features in our model fall under the following categories: clinical interventions, states (patient history/status) and outcomes. Our objective was to model the effect of interventions on outcomes and assume that the relationships between interventions and outcomes are influenced by patient state. It follows that, if we can sufficiently represent the patient state using available EHR data we will be able to model the true effects of interventions on outcomes in a systematic manner[39]. Our secondary hypothesis is that after adjustment for state variables, the provider interventions studied will generally be protective and reduce the risk of negative outcomes[4,48,49]. If successful, we will have generated a representation of physician interventions from EHR data that influence clinical outcomes. These representations can be leveraged in systematic evaluation of clinician interventions when evaluating patients with MDD. A sensitivity analysis will be performed to assess model validity and inform future studies of causal inference methods in this population.

Methods

Our study included patients aged 18-90 years with an initial MDD event between 2013 and 2018. MDD events are defined based on ICD diagnosis codes, antidepressant prescription, or problem list mention of depression, and could be a first episode or recurrent episode. The ICD codes used to indicate depression are included in the appendix table A.4. Patients were required to have two visits at least 180 days apart at VUMC prior to the initial MDD event to reduce chances that patients were receiving a majority of care at an institution other than VUMC. Patients with a Schizophrenia or Bipolar Disorder ICD code were excluded.

The purpose of our model was to estimate effects of provider interventions on outcomes. To accomplish this, we structured our data temporally into intervention and outcome periods. In the intervention period, provider interventions were recorded indicating that a specific intervention took place during that period. In the outcome period, which follows the intervention period, we recorded outcome indicators. We defined the intervention period as the 90 days following the index MDD event (see figure 1.2) and the outcome period as the 180 days following the end of the intervention period, or days 91-270 following the index MDD event. This structure assumes that provider interventions taking place during the intervention period will influence future outcomes.

We selected the feature sets belonging to the categories: interventions, states, and outcomes based on the expertise of clinicians in the VUMC Department of Psychiatry and Behavioral Science. Intervention features include prescribing an antidepressant, prescribing a second antidepressant different from the index medication, increasing antidepressant dosage, reducing antidepressant dosage, referral or consult for behavioral health services, ECT consult and partial hospitalization.

For all interventions, orders data were queried. Interventions that were not related to medications were supplemented by notes data. Regular expressions were developed to identify mentions of a referral or consult for behavioral health related therapy, ECT consult and partial hospitalizations in both the notes and orders data. Antidepressant data were extracted from the RD by ATC code[17] according to a list of antidepressants created by clinical experts from the VUMC department of psychiatry. The list of medications has been included in Table 3 in the supplementary materials with ATC code.

Study outcomes were selected while considering clinician input as well as EHR availability. The resulting outcomes included utilization related outcomes—unplanned all-cause admissions and high utilization—as well as self-harm/suicide attempt. Healthcare utilization related outcomes have been shown in prior studies to be associated with severe depression[50–52] and are leveraged in this study to assess effectiveness of treatment interventions. Unplanned admissions were defined as patients admitted to the hospital excluding any admissions that may be considered part of planned treatment according to the Center for Medicare and Medicaid Services Unplanned Readmissions Algorithm[53]. High utilization was defined as any patient with two or more inpatient or emergency room visits with an MDD ICD code during the outcome period. The self-harm/suicide attempt outcome is ICD code based, codes were mapped to the AHRQ CCS[28] codes. The CCS provides a grouping of ICD codes intended to be clinically meaningful. Patients with a code belonging to the suicide and self-inflicted injury grouping were defined as having a self-harm/suicide attempt event.

The patient state includes demographics variables: age, gender and race/ethnicity; as well as comorbidities. Comorbidities are encoded using AHRQ CCS codes[28]. We select comorbidities to include in each outcome model using a bootstrapped elastic net[54–56]. The elastic net is a form of penalized regression often used for model variable selection. The elastic net is a combination of the L1 and L2 penalties with a mixing parameter α. The parameter α ranges from 0 to 1, when set to 1 the penalty is equivalent to L1 penalized regression, and when set to 0, the L2 penalty. The L1 penalty shrinks regression coefficients to zero more aggressively, so we set $\alpha = 0.5$. The elastic net method results in a more exhaustive variable selection by allowing for correlated comorbidities to be selected together[55]. To select the comorbidities to include for each outcome, we perform 20 bootstrap replications of the elastic net model for the full set of CCS codes in our population on each of the outcomes studied. Comorbidities included in each of the 20 bootstrap replication models were included as state variables in the corresponding outcome model. Additionally, the variables unplanned admission, high utilization, self-harm/suicide attempt, and ER visit that take place after the index MDD event and prior to the end of the intervention period were included as part of the patient state as a proxy for MDD severity[50–52].

We fit separate logistic regression models for each outcome—which regress the outcomes against intervention and state variables. We then perform a test for significance of interventions by the likelihood ratio test (LRT). The LRT is a commonly used statistical method to compare competing models. In our case—models with and without the interventions. Testing the hypothesis that addition of the intervention variables improves model fit. Due to sparse nature of the high utilization and self-harm/suicide attempt outcomes the interventions ECT referral and partial hospitalization were not included in these models. We observe both these interventions at low rates (Table 2.1) and were not able to reliably estimate coefficients and standard errors for these variables.

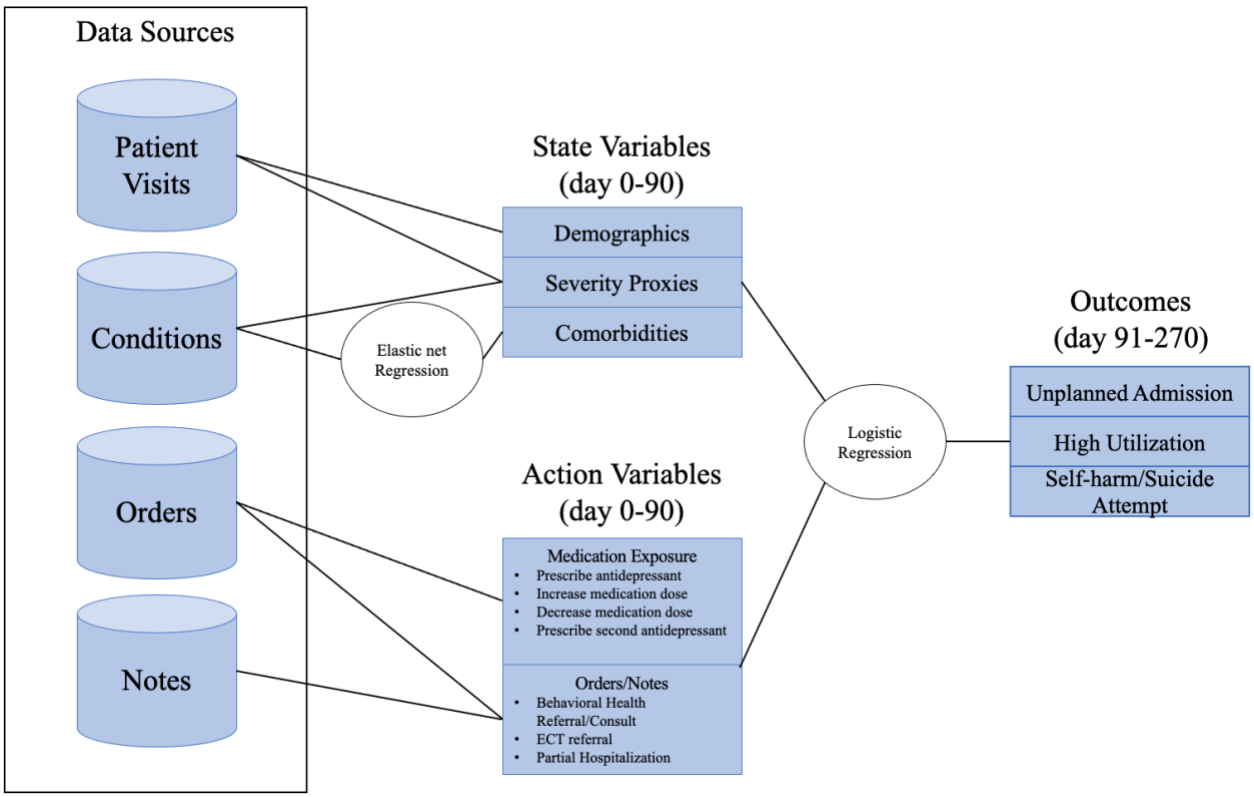

Figure 2.1-Overview of data extraction and processing

Our secondary analysis examines intervention feature associations with outcomes, and we perform a sensitivity analysis. This includes reporting model variable coefficients with corresponding confidence intervals for both intervention and state features. The sensitivity analysis reports the E-value, which is a measure of the odds ratio (OR) that a potential unobserved confounder would have to have with both the intervention and the outcome to produce a null effect in the intervention (OR = 1)[43]. In addition, we analyze changes in intervention feature coefficients as covariates are added to the model and perform a subgroup analysis. In the covariate analysis, we began with unadjusted odds ratios then add to the model in order: intervention features, demographics, severity proxies, and comorbidities and report intervention feature coefficients. The subgroup analysis measured changes to intervention feature coefficients as the data were subset by factors hypothesized to alter intervention effects. The subset groups include exclusion of patients with a cancer diagnosis, exclusion of patients with a pregnancy diagnosis code, those with a low comorbid burden, exclusion of patients with a substance abuse diagnosis code, and patients prescribed at least one antidepressant. We define a low comorbid burden as patients in the lowest quartile of the count of unique CCS codes. The threshold for inclusion in this group is less than or equal to 5 unique CCS codes.

All analyses were conducted in R Version 3.6.

Results

We identified 27,319 patients that met entry criteria. Of those, the most common index MDD event was diagnosis code (47.8%). All patients with an index MDD event identified in the problem list or antidepressant prescription had a later MDD diagnosis code. During the intervention period 59.7% of patients were prescribed an antidepressant. Of those prescribed an antidepressant, 29.6% were prescribed a second, different antidepressant during the intervention period. The most common outcome was unplanned admission (13.7%), while self-harm/suicide attempt and high utilization were relatively rare (0.5% and 1.2%). Overall, 64.2% of patients in our study received at least one intervention. Patients in each of the outcome categories were more likely to have received an intervention relative to the population total (unplanned admission = 67.4%, high utilization = 81.7%, self-harm/suicide attempt = 81.0%).

Table 2.1-Study Summary Statistics

	Unplanned Admission N=3,746		**High Utilization** N=333		**Self-harm/ suicide attempt** N=147		**Total** N=27,319	
	N	%	N	%	N	%	N	%
Index MDD Event								
Diagnosis Code	1,396	37.3%	102	30.6%	48	32.7%	13,065	47.8%
Antidepressant	1,605	42.8%	166	49.8%	82	55.8%	10,050	36.8%
Problem List	745	19.9%	65	19.5%	17	11.6%	4,204	15.4%
Gender								
Female	2,250	60.1%	189	56.8%	75	51.0%	17,621	64.5%
Race/Ethnicity								
Asian	37	1.0%	<11	<3.3%	<11	2.7%	379	1.4%
Black	477	12.7%	46	13.8%	23	15.6%	2,996	11.0%
White-hispanic/latino	52	1.4%	<11	<3.3%	<11	2.0%	387	1.4%
White-not hispanic/latino	3,080	82.2%	270	81.1%	113	76.9%	22,478	82.3%
Other	100	2.7%	<11	<3.3%	<11	2.7%	1,079	3.9%
Age								
Mean Age (SD)	50.1	(17.3)	49.1	(17.5)	37.7	(15.7)	48.1	(18.1)
Interventions								
Prescribe antidepressant	2,324	62.0%	264	79.3%	115	78.2%	16,308	59.7%
Second antidepressant	866	23.1%	137	41.1%	52	35.4%	4,840	17.7%
Increase dosage	63	1.7%	11	3.3%	<11	<7.5%	678	2.5%
Reduce dosage	32	0.9%	12	3.6%	<11	<7.5%	208	0.8%
Behavioral Health referral	825	22.0%	106	31.8%	39	26.5%	4,150	15.2%
ECT consult	17	0.5%	<11	<3.3%	<11	<7.5%	55	0.2%
Partial hospitalization	11	0.3%	<11	<3.3%	<11	<7.5%	116	0.4%
Total w/ Intervention	2,525	67.4%	272	81.7%	119	81.0%	17,535	64.2%

Comorbidity selection

The bootstrapped elastic net method for comorbidity selection resulted in 38, 5, and 14 comorbidities for the outcomes unplanned admission, high utilization and self-harm/suicide attempt respectively. Selected comorbidities are included in appendix table A.1. Mood disorders were common across all three outcome models and each included gastrointestinal related comorbidities (Other gastrointestinal disorders, Upper gastrointestinal disorders). Substance abuse and alcohol related disorders were included as comorbidities in the unplanned admissions and self-harm/suicide attempt models.

Model Results

We regressed each outcome against intervention and state variables and performed a LRT of association between interventions and outcomes after adjustment for state variables. Interventions were found to have statistically significant association with each outcome. We observed that prescribing an initial antidepressant had a protective effect—or a statistically significant reduction in the odds of an unplanned admission. Multiple variables exhibited

statistically significant associations with increased risk of an outcome, including prescribing a second antidepressant and behavioral health referral with unplanned admission, prescribe a second antidepressant and dose decrease with high utilization, as well as prescribing a second antidepressant with self-harm/suicide attempt.

Table 2.2-Model summary, LRT, and intervention coefficients

	Unplanned Admission			High Utilization			Self-harm/Suicide Attempt		
	LRT	$2.0x10^{-18}$		LRT	$6.1x10^{-7}$		LRT	$4.7x10^{-8}$	
	p-value			p-value			p-value		
		95% CI			95% CI			95% CI	
	OR	lower	upper	OR	lower	upper	OR	lower	upper
Prescribe Initial Med	0.67	0.58	0.78	1.14	0.71	1.87	1.82	0.94	3.73
Prescribe Second Med	1.19	1.00	1.42	1.58	1.04	2.41	2.09	1.12	3.86
Dose Increase	1.00	0.61	1.58	0.97	0.29	2.52	1.18	0.28	3.48
Dose Decrease	1.36	0.65	2.64	3.80	1.15	10.15	2.05	0.24	8.65
Behavioral Health Referral	1.23	1.04	1.46	1.31	0.85	1.99	1.49	0.77	2.77
ECT Referral	1.57	0.51	4.45						
Partial Hospitalization	1.15	0.34	3.02						

95% CI's are Bonferroni corrected for multiple analyses

Statistically significant increased risk of outcome

Statistically significant decreased risk of outcome

State Variable Effects

Figure 2.2 displays forest plots of state variable effects. For the unplanned admissions outcome, having a pregnancy complication during the intervention period had the largest effect (OR = 4.37, 95% CI = 3.30-5.77) and the second largest effect was normal pregnancy (OR = 3.46, 95% CI = 2.53-4.74). For high utilization the state variable with largest effect was having an unplanned admission during the intervention period (OR = 2.39, 95% CI = 1.83-3.13). Patients with a alcohol (OR = 2.66, 95% CI = 1.62-4.26) and substance disorders (OR = 2.37, 95% CI = 1.54-3.60) during the intervention period have statistically significant increased risk of self-harm/suicide attempt. Mood disorders were protective in each outcome model. Self-harm/suicide attempts during the intervention period were protective against self-harm/suicide attempt during the outcome period (OR = 0.033, 95% CI = 0.0019-0.15).

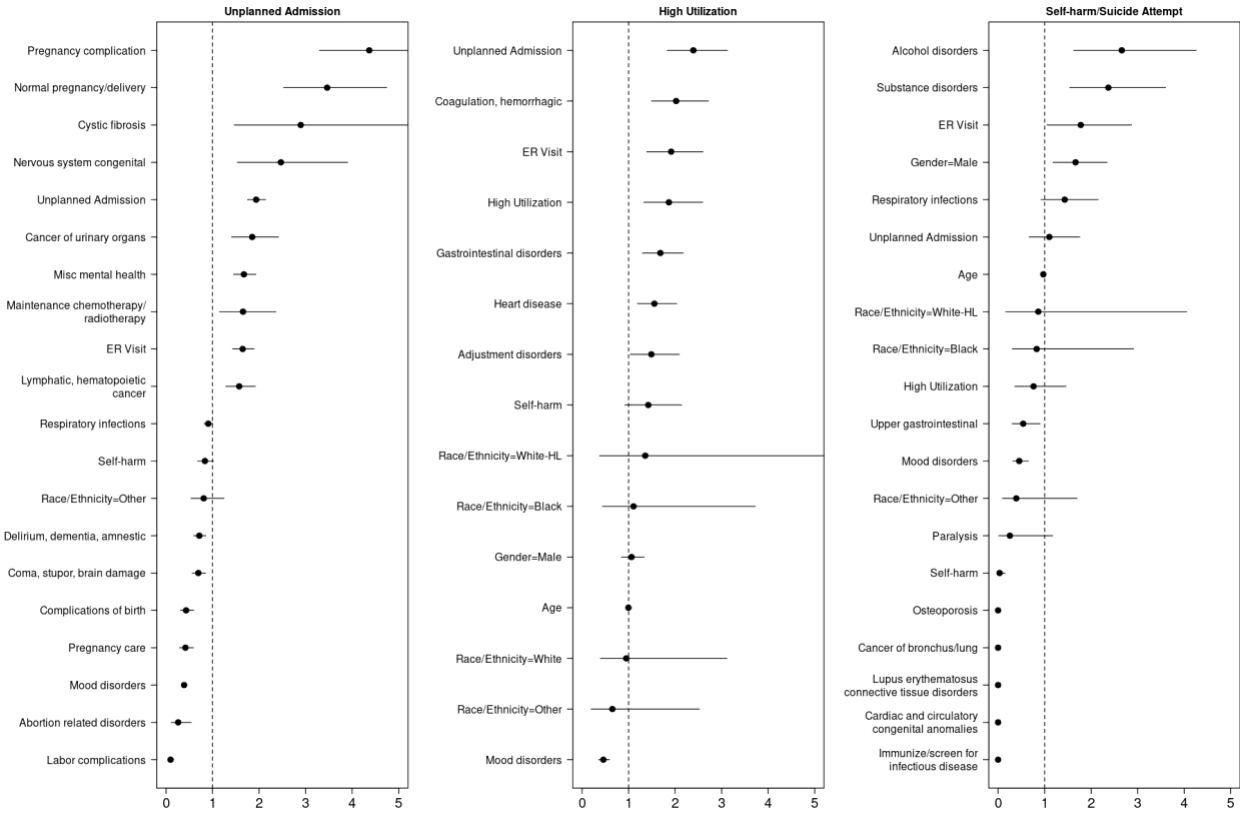

Figure 2.2-Patient state variable forest plots by outcome model. Top ten increased risk and protective effect ORs.

Sensitivity Analysis

The sensitivity analysis conducted in this study includes measurement of the required strength of hypothetical confounders, changes to intervention coefficients due to model covariates and subsets of the population. The E-value provides a measure of the strength of association potential confounders would need to have with both the intervention and the outcome to nullify the effect. E-values for interventions found to be statistically significant are reported in appendix table A.2. Dose decrease in the high utilization model has the highest E-value, 7.06, meaning a hypothetical confounder would have to have a relatively strong association—odds ratio magnitude greater than 7.06—with both self-harm/suicide attempt and prescribing an initial medication to nullify the observed effect.

In our covariate analysis we observed a majority of unadjusted intervention variables with a statistically significant increased risk of negative outcomes (65%). A single intervention outcome pair had an unadjusted protective effect—dose increase with unplanned admission. The association between prescribing an initial medication and unplanned admission flipped from unadjusted increased risk (OR = 1.12) to protective in the full model (OR = 0.67) and is significant at the 99.9% level.

Table 2.3-Covariate Analysis

Significance thresholds for Tables 3 and 4

99.9%	significant decreased risk
99%	
95%	
90%	
99.9%	significant increased risk
99%	
95%	
90%	

Significance thresholds are Bonferroni corrected

Unplanned Admission (N = 3,746, 13.7%)

	Unadjusted	Interventions	Interventions + Demographics	Interventions + Demographics + Severity Proxies	Full Model
Prescribe Initial Antidepressant	1.12	0.98	0.96	0.84	0.67
Prescribe Second Antidepressant	1.48	1.38	1.35	1.17	1.19
Dose Increase	0.64	0.58	0.64	0.76	1.00
Dose Decrease	1.14	1.16	1.20	1.24	1.36
Behavioral Health Referral	1.72	1.62	1.65	1.38	1.23
ECT Referral	2.82	2.11	2.05	1.76	1.57
Partial Hospitalization	0.66	0.49	0.55	0.67	1.15

High Utilization (N = 333, 1.2%)

Prescribe Initial Antidepressant	2.61	1.72	1.70	1.43	1.14
Prescribe Second Antidepressant	3.31	2.24	2.21	1.67	1.58
Dose Increase	1.35	0.75	0.79	0.85	0.97
Dose Decrease	5.10	3.41	3.50	3.28	3.80
Behavioral Health Referral	2.65	2.01	1.99	1.46	1.31

Self-harm and Suicide Attempt (N = 147, 0.5%)

Prescribe Initial Antidepressant	2.43	1.81	1.92	1.86	1.82
Prescribe Second Antidepressant	2.55	1.77	1.79	2.09	2.09
Dose Increase	2.28	1.50	1.05	1.19	1.18
Dose Decrease	3.70	2.10	2.00	2.04	2.05
Behavioral Health Referral	2.02	1.62	1.38	1.50	1.49

In our subgroup analysis we subset the study dataset to groups of interest and report intervention coefficients. The subgroups analyzed are: patients with no cancer diagnoses, patients with no pregnancy related diagnoses, low comorbid burden, no substance abuse codes, and those with at least one antidepressant prescription during the intervention period. In the unplanned admission model, prescription of an initial antidepressant is significant at the 99.9% level in each of the subgroups with the exception of the at least one med subset because it is not included in this model. Prescribing a second antidepressant is significant at the 99.9% level in high utilization and self-harm/suicide attempt models in the exclude substance abuse subgroup.

Table 2.4-Subgroup Analysis

Unplanned Admission

	Full Model	Exclude Cancer	Exclude Pregnancy	Low Comorbidities (<= 5 CCS)	Exclude Substance Abuse	Prescribed >= 1 Med
Model N	27,319	21,731	26,038	8,524	21,968	16,308
Outcome N (%)	3,746 (13.7%)	2,540 (11.7%)	3,375 (13.0%)	602 (7.1%)	2,780 (12.7%)	2,324 (14.2%)
Prescribe Initial Antidepressant	0.67	0.70	0.71	0.54	0.62	NA
Prescribe Second Antidepressant	1.19	1.16	1.16	1.25	1.25	1.13
Dose Increase	1.00	1.02	1.00	1.43	1.05	1.01
Dose Decrease	1.36	1.49	1.42	1.61	1.18	1.39
Behavioral Health Referral	1.23	1.20	1.21	1.14	1.29	1.26
ECT Referral	1.57	1.38	1.49	6.65	3.01	1.73
Partial Hospitalization	1.15	1.20	1.07	2.68	1.42	1.07

High Utilization

Outcome N (%)	333 (1.2%)	209 (1.0%)	319 (1.2%)	52 (0.6%)	240 (1.1%)	264 (1.6%)
Prescribe Initial Antidepressant	1.14	1.16	1.12	0.96	0.90	NA
Prescribe Second Antidepressant	1.58	1.57	1.58	1.55	1.99	1.51
Dose Increase	0.97	1.03	1.00	1.14	1.04	0.91
Dose Decrease	3.80	4.60	3.95	5.70	2.32	3.75
Behavioral Health Referral	1.31	1.24	1.32	0.97	1.38	1.30

Self-harm and Suicide Attempt

Outcome N (%)	147 (0.5%)	129 (0.6%)	141 (0.05%)	46 (0.5%)	86 (0.4%)	115 (0.7%)
Prescribe Initial Antidepressant	1.82	1.84	1.70	2.07	1.52	NA
Prescribe Second Antidepressant	2.09	2.02	2.10	2.83	3.55	2.15
Dose Increase	1.18	1.28	1.15	1.65	1.39	1.04
Dose Decrease	2.05	2.33	2.08	3.01	2.44	2.00
Behavioral Health Referral	1.49	1.49	1.76	0.65	1.37	1.71

Discussion

In this study we sought to effectively represent provider interventions in the treatment of MDD using data available in the EHR and measure the association between provider interventions and outcomes after adjustment for patient state. We observed significant associations between provider interventions and the outcomes included in our study. Our secondary analysis measured the effects of specific provider interventions on outcomes. We hypothesized that, after adjustment for the patient's state, provider interventions would be protective. However, this result was not observed. Provider interventions tend to be associated with an increased risk of the outcomes studied after adjustment for patient state—we observed five increased risk intervention outcome pairs to just one protective association.

The sensitivity analysis provides further insight, showing that unadjusted associations between provider interventions and outcomes are strong, positive, and statistically significant. However, adjustment for patient state variables does reduce the effect size in many cases. For example, the unadjusted odds of an unplanned admission for patients prescribed an antidepressant is 12% higher than those not prescribed an antidepressant holding all other factors constant. After adjustment for patient state, patients being prescribed an antidepressant have a 33% reduction in odds of unplanned admission relative to those not prescribed an antidepressant. This effect is significant at the 99.9% level and consistent through our subgroup analysis. Patients receiving a psych referral or consult are at 2.65 times the risk of high utilization relative to those with no referral or consult. After adjustment for patient state this

risk is reduced to 1.31 times and is not statistically significant. This suggests that the patient state representation is reducing the increased risk effect sizes, but may be insufficient to identify true effects.

Even after adjustment for state variables we find that most interventions are associated with an increased risk of the outcomes in our study. We hypothesized that interventions in this study would have protective associations with outcomes after adjustment for patient state variables. However, the patient state representation is limited to what is available in the EHR. There are likely important factors influencing the outcomes of patients with MDD that are not included in this study, such as depression severity, diet, exercise, medication adherence, and genetics. These factors potentially introduce confounder bias to this study. Potential confounding bias is supported by the E-value analysis. The E-values range from 1.76 to 7.06 and five of the six E-values are less than 4. There are eight state variable-outcome pairs in this study with adjusted OR magnitude greater than 4. It is reasonable to expect that there are unobserved confounders of the statistically significant intervention-outcome associations that could nullify the observed effect.

In addition to confounder bias, adjustment for state variables may introduce risk of a collider bias[41,57]. It is possible that state variables, such as comorbidities have a mediating effect between interventions and the outcomes studied. If an unobserved variable exists that is associated with the mediator and outcomes, adjustment for the mediator can cause a collider effect to mask the true effect of interventions on outcomes. An example case being, if an intervention were causing one of the state variables, such as substance abuse, and there was an unobserved variable causal of substance abuse and the outcomes studied. Adjusting for substance abuse would cause a collider bias. However, in our covariate analysis section we fit the models with and without adjustment for patient state variables. As we add variables we observe a decreasing trend in odds ratios, suggesting a confounder bias that is partially addressed by patient state adjustment rather than a collider bias.

The subgroup analysis finds consistent treatment effects within subgroups of the population that were informed by prior work, as well as the analysis of state variable effects in figure 2.2. The low comorbidities and exclude substance abuse subgroups were intended to identify lower severity subpopulations where treatments were more effective. This analysis did not provide evidence for heterogeneous treatment effects. However, it may be beneficial to apply more data driven selections of subpopulations for analysis of heterogeneous treatment effects in this population[58,59].

As stated previously, the primary limitation is the incompleteness of information available in the EHR. Lack of depression severity measures, objective diagnostic evaluation to confirm MDD and patient lifestyle data impacts our study population and patient state representations leading to potential confounding of our results. The study outcomes were also limited by EHR availability. Survey based measures of depression severity are common in MDD trials[4,60,61], but were not regularly available in the EHR for the study population. In addition, we perform multiple analyses throughout the study. To account for multiple analyses, we perform a Bonferroni correction, reducing the power of our study. Future work will focus on addressing the potential unobserved confounding identified in this study. Causal inference methods such as instrumental variables[62] and latent variable based models[44,46] may be applied to this dataset for more meaningful results. There are also potential opportunities to improve patient state representations by further mining the EHR.

Conclusion

The results of this study suggest that insufficient covariates are accessible in the EHR to identify the true effects of clinician interventions on outcomes in the MDD population. Further study of methods that have the potential to adjust for unobserved confounding are necessary, as well as improved documentation of factors effecting outcomes in MDD.

Summary

In chapter 1 we lay the groundwork for systematic evaluation of interventions by modeling and characterizing in-place treatment trajectories of MDD. This study provides insight into distinct treatment pathways that correspond to endotypes of MDD, while informing our understanding of how treatment of MDD progresses. In chapter 2 we begin to explore evaluation of clinician interventions, focusing on initial—post MDD diagnosis—interventions. We outline in chapter 2 some of the potential confounders of our study. These include depression severity, social support, lifestyle factors, medication adherence, and genetics. Representation of depression severity is of primary concern. There are potential tools for measuring MDD severity including the Patient Health Questionnaire (PHQ) surveys. PHQ surveys have been shown to be associated with the Hamilton Depression Rating Scale[63]. Regular collection of PHQ surveys have potential to improve the ability to perform systematic evaluation of intervention in the MDD population by providing a consistent measure of depression severity. For the time-period studied these data are not widely available in the VUMC EHR. However, as part of an initiative to improve assessment of patient reported outcomes at VUMC, PHQ-8 utilization is increasing—with over 20,000 PHQ-8 surveys completed from 1/1/2021 to 10/10/2021.

Future studies will use the addition of PHQ-8 data and will assess the utility of causal inference methods to reduce confounding. One such method is the deconfounder algorithm developed by Wang and Blei in 2018[44]. This model infers a latent structure of confounding that can be identified by observed intervention selection. This method employs latent factor models to capture the latent structure of clinician actions. For example, high severity patients are more likely to receive an ECT consultation. A correctly fit latent factor model would be able to extract and represent severity as a continuous variable. The latent variables generated by the latent factor model can be used as substitute confounders to make inference about clinician interventions.

We have shown that the chronic nature of MDD results in dynamic patient states and clinician interventions requiring longitudinal evaluation. We model and characterize treatments in chapter 1 and can expand on this study by employing DTRs. As discussed earlier, these methods allow for intervention individualization based on the dynamic nature of a patient's state. DTRs optimize over future outcomes and recommend an intervention policy given the patient state. The ability to model DTRs in MDD will be dependent on accurately modeling intervention effects. A dynamic treatment regime based on the models in chapter 2 would likely recommend no actions in many cases—as a majority of the clinician actions studied are associated with increased risk of negative outcomes. The logistic regression model utilized in chapter 2 is one of the simplest models that can be used for this type of problem. DTR models such as Q-Learning and Marginal Structural Models lend themselves better to longitudinal data, interactions, and heterogeneous treatment effects.

Treatment of MDD is complex, but through advances in data collection and analysis methods we can better understand this phenotype and evaluate interventions. This thesis has characterized in place treatment practices and identified some of the difficulties in evaluation of medical interventions in the MDD population. DTR models informed by the research presented here will allow for the ability to mine what can be considered best practices from the EHR and identify where those practices may deviate from what is commonly done.

Appendix

Table A.1. Medications included in the study by Anatomical Therapeutic Chemical classification (ATC) and clinical grouping

Clinical Grouping	Medication Name (ATC Level 5)
Selective serotonin reuptake inhibitors (SSRI)	citalopram fluoxetine paroxetine escitalopram fluvoxamine sertraline
Tricyclic antidepressants	doxepin amitriptyline desipramine nortriptyline protriptyline clomipramine maprotiline amoxapine imipramine
Monoamine oxidase inhibitors, non-selective (MAOI)	tranylcypromine phenelzine isocarboxazid
Second generation antipsychotic medications	aripiprazole quetiapine olanzapine ziprasidone
Other antidepressants	vilazodone bupropion vortioxetine trazodone mirtazapine nefazodone
Centrally acting sympathomimetics	methylphenidate dexamfetamine
Mood Stabilizers	lithium lamotrigine
Serotonin-norepinephrine reuptake inhibitors (SNRI)	desvenlafaxine duloxetine venlafaxine milnacipran

Table A.2. Comorbidities selected by elastic net

Unplanned Admissions		High Utilization	Self-harm/Suicide Attempt
-Bacterial Infection	-Cystic fibrosis	-Coagulation and hemorrhagic disorders	-Immunizations and screening for infectious disease
-Viral Infection	-Coagulation and hemorrhagic disorders	-Adjustment disorders	-Osteoporosis
-Abortion related disorders	-Diseases of white blood cells	-Mood disorders	-Systemic lupus erythematosus and connective tissue disorders
-Complications mainly related to pregnancy	-Alcohol related disorders	-Diseases of the heart	-Cardiac and circulatory congenital anomalies
-Indications for care in pregnancy	-Substance related disorders	-Other gastrointestinal disorders	-Cancer of bronchus/lung
-Complications during labor	-Miscellaneous mental health disorders		-Disorders of lipid metabolism
-Complications of birth	-Delirium, dementia, amnestic		-Alcohol related disorders
-Normal pregnancy and or delivery	-Mood disorders		-Substance related disorders
-Infective arthritis and osteomyelitis	-Hereditary/degenerative nervous system		-Mood disorders
-Osteoporosis	-Epilepsy convulsions		-Paralysis
-Genitourinary congenital anomalies	-Coma, stupor and brain damage		-Eye disorders
-Nervous system congenital anomalies	-Other nervous system disorders		-Respiratory infections
-Complications	-Hypertension		-Chronic obstructive pulmonary disease and bronchiectasis
-Factors influencing healthcare	-Diseases of the heart		-Upper gastrointestinal disorders
-Cancer of lymphatic and hematopoietic tissue	-Respiratory infections		
-Maintenance chemotherapy/radiotherapy	-Non-infectious gastroenteritis		
-Cancer of urinary organs	-Other gastrointestinal disorders		
-Nutritional endocrine and metabolic disorders	-Disorders of teeth and jaw		
-Thyroid disorders	-Upper gastrointestinal disorders		

Table A.3. Intervention variable E-values

	Unplanned Admission	High Utilization	Self-Harm/ Suicide Attempt
Prescribe Initial Med	2.34		
Prescribe Second Med	1.66	2.53	3.60
Dose Increase			
Dose Decrease		7.06	
Behavioral Health Referral	1.76		
ECT Referral			
Partial Hospitalization			

Table A.4. Antidepressant List with ATC code

Code	Name	Code	Name	Code	Name
N06BA04	methylphenidate	N06AA10	nortriptyline	N06AX16	venlafaxine
N06BA02	dexamfetamine	N06AA11	protriptyline	N06AX17	milnacipran
N05AH04	quetiapine	N06AA04	clomipramine	N06AX06	nefazodone
N05AH03	olanzapine	N06AA21	maprotiline	N06AX23	desvenlafaxine
N05AE04	ziprasidone	N06AA17	amoxapine	N03AX09	lamotrigine
N05AN01	lithium	N06AA02	imipramine	N05AX12	aripiprazole
N06AF04	tranylcypromine	N06AX24	vilazodone	N06AB04	citalopram
N06AF03	phenelzine	N06AX12	bupropion	N06AB03	fluoxetine
N06AF01	isocarboxazid	N06AX21	duloxetine	N06AB05	paroxetine
N06AA12	doxepin	N06AX26	vortioxetine	N06AB10	escitalopram
N06AA09	amitriptyline	N06AX05	trazodone	N06AB08	fluvoxamine
N06AA01	desipramine	N06AX11	mirtazapine	N06AB06	sertraline

Table A.5. Depression ICD codes list

	Code	Description
ICD-9	311.x	Depressive disorder, not elsewhere classified
	296.2x	Major depressive disorder single episode
	296.3x	Major depressive disorder recurrent episode
	300.4x	Dysthymic disorder
ICD-10	F32.xx	Major depressive disorder, single episode
	F33.xx	Major depressive disorder, recurrent
	F34.1	Dysthymic disorder

References

1. Etheredge LM. Rapid Learning: A Breakthrough Agenda. *Health Aff (Millwood)*. 2014;33(7):1155-1162. doi:10.1377/hlthaff.2014.0043

2. Hripcsak G, Ryan PB, Duke JD, et al. Characterizing treatment pathways at scale using the OHDSI network. *Proc Natl Acad Sci U S A*. 2016;113(27):7329-7336. doi:10.1073/pnas.1510502113

3. Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research | Translational Behavioral Medicine | Oxford Academic. Accessed November 15, 2021. https://academic.oup.com/tbm/article-abstract/4/3/260/4561511

4. Sinyor M, Schaffer A, Levitt A. The Sequenced Treatment Alternatives to Relieve Depression (STAR*D) Trial: A Review. *Can J Psychiatry*. 2010;55(3):126-135. doi:10.1177/070674371005500303

5. Martin L, Hutchens M, Hawkins C, Radnov A. How much do clinical trials cost? *Nat Rev Drug Discov*. 2017;16(6):381-383.

6. OPEN SCIENCE COLLABORATION. Estimating the reproducibility of psychological science. *Science*. 2015;349(6251):aac4716. doi:10.1126/science.aac4716

7. Chakraborty B, Laber EB, Zhao Y. Inference for Optimal Dynamic Treatment Regimes using an Adaptive m-out-of-n Bootstrap Scheme. *Biometrics*. 2013;69(3). doi:10.1111/biom.12052

8. Schulte PJ, Tsiatis AA, Laber EB, Davidian M. Q- and A-Learning Methods for Estimating Optimal Dynamic Treatment Regimes. *Stat Sci*. 2014;29(4):640-661. doi:10.1214/13-STS450

9. Gottesman O, Johansson F, Meier J, et al. Evaluating Reinforcement Learning Algorithms in Observational Health Settings. *ArXiv180512298 Cs Stat*. Published online May 30, 2018. Accessed May 25, 2020. http://arxiv.org/abs/1805.12298

10. Yu C, Liu J, Nemati S. Reinforcement Learning in Healthcare: A Survey. :32.

11. Chakraborty B, Murphy SA. Dynamic Treatment Regimes. *Annu Rev Stat Its Appl*. 2014;1:447-464. doi:10.1146/annurev-statistics-022513-115553

12. Joffe R, Sokolov S, Streiner D. Antidepressant Treatment of Depression: A Metaanalysis. *Can J Psychiatry*. 1996;41(10):613-616. doi:10.1177/070674379604101002

13. Barak-Corren Y, Castro VM, Javitt S, et al. Predicting Suicidal Behavior From Longitudinal Electronic Health Records. *Am J Psychiatry*. 2016;174(2):154-162. doi:10.1176/appi.ajp.2016.16010077

14. Walsh CG, Ribeiro JD, Franklin JC. Predicting suicide attempts in adolescents with longitudinal clinical data and machine learning. *J Child Psychol Psychiatry*. 2018;59(12):1261-1270. doi:10.1111/jcpp.12916

15. Zhao J, Papapetrou P, Asker L, Boström H. Learning from heterogeneous temporal data in electronic health records. *J Biomed Inform*. 2017;65:105-119. doi:10.1016/j.jbi.2016.11.006

16. Zhang X, Wang L, Miao S, et al. Analysis of treatment pathways for three chronic diseases using OMOP CDM. *J Med Syst*. 2018;42(12):260. doi:10.1007/s10916-018-1076-5

17. WHO | 2. Anatomical Therapeutic Chemical (ATC) Classification. WHO. Accessed April 30, 2020. http://www.who.int/medicines/regulation/medicines-safety/toolkit_atc/en/

18. Roden DM, Pulley JM, Basford MA, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin Pharmacol Ther*. 2008;84(3):362-369. doi:10.1038/clpt.2008.89

19. Pandarakalam JP. Challenges of Treatment-resistant Depression. *Psychiatr Danub*. 2018;30(3):273-284. doi:10.24869/psyd.2018.273

20. Clements CC, Karlsson R, Lu Y, et al. Genome-wide association study of patients with a severe major depressive episode treated with electroconvulsive therapy. *Mol Psychiatry*. Published online January 22, 2021:1-11. doi:10.1038/s41380-020-00984-0

21. Hochreiter S, Schmidhuber J. Long Short-Term Memory. *Neural Comput*. 1997;9(8):1735-1780. doi:10.1162/neco.1997.9.8.1735

22. Goodfellow I. NIPS 2016 Tutorial: Generative Adversarial Networks. *ArXiv170100160 Cs*. Published online April 3, 2017. Accessed April 30, 2020. http://arxiv.org/abs/1701.00160

23. Doersch C. Tutorial on Variational Autoencoders. *ArXiv160605908 Cs Stat*. Published online August 13, 2016. Accessed April 30, 2020. http://arxiv.org/abs/1606.05908

24. Srivastava N, Mansimov E, Salakhutdinov R. Unsupervised Learning of Video Representations using LSTMs. *ArXiv150204681 Cs*. Published online January 3, 2016. Accessed April 30, 2020. http://arxiv.org/abs/1502.04681

25. Baytas IM, Xiao C, Zhang X, Wang F, Jain AK, Zhou J. Patient Subtyping via Time-Aware LSTM Networks. In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '17. Association for Computing Machinery; 2017:65-74. doi:10.1145/3097983.3097997

26. Home - Keras Documentation. Accessed April 30, 2020. https://keras.io/

27. Yuan C, Yang H. Research on K-Value Selection Method of K-Means Clustering Algorithm. *J*. 2019;2(2):226-235. doi:10.3390/j2020016

28. Clinical Classifications Software (CCS) for ICD-10-PCS (beta version). Accessed March 5, 2021. https://www.hcup-us.ahrq.gov/toolssoftware/ccs10/ccs10.jsp

29. Lin F ren, Chou S chao, Pan S mei, Chen Y mei. Mining time dependency patterns in clinical pathways. *Int J Med Inf*. 2001;62(1):11-25. doi:10.1016/S1386-5056(01)00126-5

30. Bobo WV, Pathak J, Kremers HM, et al. An electronic health record driven algorithm to identify incident antidepressant medication users. *J Am Med Inform Assoc*. 2014;21(5):785-791. doi:10.1136/amiajnl-2014-002699

31. Lenert L, McSwain BY. Balancing health privacy, health information exchange, and research in the context of the COVID-19 pandemic. *J Am Med Inform Assoc JAMIA*. 2020;27(6):963-966. doi:10.1093/jamia/ocaa039

32. Belsher BE, Smolenski DJ, Pruitt LD, et al. Prediction Models for Suicide Attempts and Deaths: A Systematic Review and Simulation. *JAMA Psychiatry*. 2019;76(6):642-651. doi:10.1001/jamapsychiatry.2019.0174

33. Burke TA, Ammerman BA, Jacobucci R. The use of machine learning in the study of suicidal and non-suicidal self-injurious thoughts and behaviors: A systematic review. *J Affect Disord*. 2019;245:869-884. doi:10.1016/j.jad.2018.11.073

34. Geleris J, Sun Y, Platt J, et al. Observational Study of Hydroxychloroquine in Hospitalized Patients with Covid-19. *N Engl J Med*. 2020;382(25):2411-2418. doi:10.1056/NEJMoa2012410

35. Lao KSJ, Chui CSL, Man KKC, Lau WCY, Chan EW, Wong ICK. Medication safety research by observational study design. *Int J Clin Pharm*. 2016;38(3):676-684. doi:10.1007/s11096-016-0285-6

36. Man KKC, Coghill D, Chan EW, et al. Association of Risk of Suicide Attempts With Methylphenidate Treatment. *JAMA Psychiatry*. 2017;74(10):1048-1055. doi:10.1001/jamapsychiatry.2017.2183

37. Man KKC, Coghill D, Chan EW, et al. Methylphenidate and the risk of psychotic disorders and hallucinations in children and adolescents in a large health system. *Transl Psychiatry*. 2016;6(11):e956-e956. doi:10.1038/tp.2016.216

38. Fusaro M, D'Arrigo G, Pitino A, et al. Increased Risk of Bone Fractures in Hemodialysis Patients Treated with Proton Pump Inhibitors in Real World: Results from the Dialysis Outcomes and Practice Patterns Study (DOPPS). *J Bone Miner Res*. 2019;34(12):2238-2245. doi:10.1002/jbmr.3842

39. McNamee R. Regression modelling and other methods to control confounding. *Occup Environ Med*. 2005;62(7):500-506. doi:10.1136/oem.2002.001115

40. Brazauskas R, Logan BR. Observational Studies: Matching or Regression? *Biol Blood Marrow Transplant*. 2016;22(3):557-563. doi:10.1016/j.bbmt.2015.12.005

41. Jiang T, Smith ML, Street AE, et al. A comorbid mental disorder paradox: Using causal diagrams to understand associations between posttraumatic stress disorder and suicide. *Psychol Trauma Theory Res Pract Policy*. Published online 2021:No Pagination Specified-No Pagination Specified. doi:10.1037/tra0001033

42. Pearl J. An Introduction to Causal Inference. *Int J Biostat*. 2010;6(2). doi:10.2202/1557-4679.1203

43. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med*. 2017;167(4):268-274. doi:10.7326/M16-2607

44. Wang Y, Blei DM. The Blessings of Multiple Causes. *ArXiv180506826 Cs Stat*. Published online April 14, 2019. Accessed December 9, 2020. http://arxiv.org/abs/1805.06826

45. Zhang L, Wang Y, Ostropolets A, Mulgrave JJ, Blei DM, Hripcsak G. The Medical Deconfounder: Assessing Treatment Effects with Electronic Health Records. In: *Machine Learning for Healthcare Conference*. PMLR; 2019:490-512. Accessed September 1, 2021. https://proceedings.mlr.press/v106/zhang19a.html

46. Ranganath R, Perotte A. Multiple Causal Inference with Latent Confounding. *ArXiv180508273 Cs Stat*. Published online March 1, 2019. Accessed September 30, 2021. http://arxiv.org/abs/1805.08273

47. Brookhart MA, Schneeweiss S. Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. *Int J Biostat*. 2007;3(1):Article 14. doi:10.2202/1557-4679.1072

48. Pagnin D, de Queiroz V, Pini S, Cassano GB. Efficacy of ECT in depression: a meta-analytic review. *J ECT*. 2004;20(1):13-20. doi:10.1097/00124509-200403000-00004

49. Thatte S, Makinen JA, Nguyen HNT, Hill EM, Flament MF. Partial hospitalization for youth with psychiatric disorders: treatment outcomes and 3-month follow-up. *J Nerv Ment Dis*. 2013;201(5):429-434. doi:10.1097/NMD.0b013e31828e1141

50. Sun Y, Möller J, Lundin A, Wong SYS, Yip BHK, Forsell Y. Utilization of psychiatric care and antidepressants among people with different severity of depression: a population-based cohort study in Stockholm, Sweden. *Soc Psychiatry Psychiatr Epidemiol*. 2018;53(6):607-615. doi:10.1007/s00127-018-1515-0

51. Snell C, Fernandes S, Bujoreanu IS, Garcia G. Depression, illness severity, and healthcare utilization in cystic fibrosis. *Pediatr Pulmonol*. 2014;49(12):1177-1181. doi:10.1002/ppul.22990

52. Kornstein SG, Joseph AC, Graves WC, Wallenborn JT. Prenatal Depression Severity and Postpartum Care Utilization in a Medicaid Population. *Womens Health Rep New Rochelle N*. 2020;1(1):468-473. doi:10.1089/whr.2020.0079

53. Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation. 2014 Measure Updates and Specifications Report Hospital-Wide All-Cause Unplanned Readmission – Version 3.0. Published online July 2014.

54. Teipel SJ, Grothe MJ, Metzger CD, et al. Robust Detection of Impaired Resting State Functional Connectivity Networks in Alzheimer's Disease Using Elastic Net Regularized Regression. *Front Aging Neurosci*. 2017;8. doi:10.3389/fnagi.2016.00318

55. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc Ser B Stat Methodol*. 2005;67(2):301-320. doi:10.1111/j.1467-9868.2005.00503.x

56. Friedman J, Hastie T, Tibshirani R, et al. *Glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models*.; 2021. Accessed August 10, 2021. https://CRAN.R-project.org/package=glmnet

57. Collider bias undermines our understanding of COVID-19 disease risk and severity | Nature Communications. Accessed October 25, 2021. https://www.nature.com/articles/s41467-020-19478-2

58. Powers S, Qian J, Jung K, et al. Some methods for heterogeneous treatment effect estimation in high dimensions. *Stat Med*. 2018;37(11):1767-1787. doi:10.1002/sim.7623

59. Varadhan R, Seeger JD. *Estimation and Reporting of Heterogeneity of Treatment Effects*. Agency for Healthcare Research and Quality (US); 2013. Accessed September 14, 2021. https://www.ncbi.nlm.nih.gov/books/NBK126188/

60. Chevance A, Ravaud P, Tomlinson A, et al. Identifying outcomes for depression that matter to patients, informal caregivers, and health-care professionals: qualitative content analysis of a large international online survey. *Lancet Psychiatry*. 2020;7(8):692-702. doi:10.1016/S2215-0366(20)30191-7

61. Kroenke K, Shen J, Oxman TE, Williams JW, Dietrich AJ. Impact of pain on the outcomes of depression treatment: Results from the RESPECT trial. *Pain*. 2008;134(1):209-215. doi:10.1016/j.pain.2007.09.021

62. Lorch SA, Baiocchi M, Ahlberg CE, Small DS. The Differential Impact of Delivery Hospital on the Outcomes of Premature Infants. *Pediatrics*. 2012;130(2):270-278. doi:10.1542/peds.2011-2820

63. Shin C, Lee SH, Han KM, Yoon HK, Han C. Comparison of the Usefulness of the PHQ-8 and PHQ-9 for Screening for Major Depressive Disorder: Analysis of Psychiatric Outpatient Data. *Psychiatry Investig*. 2019;16(4):300-305. doi:10.30773/pi.2019.02.01