

A Multi-Level Analysis of Major Health Challenges in the United States Using Data Analytics Approaches

Negar Darabi

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of:

Doctor of Philosophy
in
Industrial and Systems Engineering

Niyousha Hosseinichimeh, Chair
Konstantinos P. Triantis
Zhenyu (James) Kong
Vida Abedi

August 11, 2020

Falls Church, VA

Keywords: Infant Mortality, Performance Measurement, Data Envelopment Analysis, Opioid Epidemic, Statistical Analysis, Hospital Readmissions, Stroke, Machine Learning, Risk Factors

A Multi-Level Analysis of Major Health Challenges in the United States Using Data Analytics Approaches

Negar Darabi

ABSTRACT

The U.S. healthcare system is facing many public health challenges that affect population health, societal well-being, and quality of healthcare. Infant mortality, opioid overdose death, and hospital readmission after stroke are some of these important public health concerns that can impact the effectiveness and outcomes of the healthcare system. We analyze these problems through the industrial engineering and data analytics lens. The major goal of this dissertation is to enhance understanding of these three challenges and related interventions using different levels of analysis to improve the health outcomes. To attain this objective, I introduced three stand-alone papers to answer the related research questions.

In essay 1, we focused on the performance of the state's healthcare systems in reducing unfavorable birth outcomes such as infant mortality, preterm birth, and low birthweight using Data Envelopment Approach. We constructed a unique state-level dataset to answer this main research question: what does make a healthcare system more successful in improving the birth outcomes? Our results indicated that socioeconomic and demographic factors may facilitate or obstruct health systems in improving their outcomes. We realized that states with a lower rate of poverty and African-American women were more successful in effectively reduce unfavorable birth outcomes. In the second essay, we looked into the trends of the opioid overdose mortalities in each state from 2008 to 2017. We investigated the effect of four state laws and programs that have been established to curb the epidemic (i.e., dose and duration limitations on the initial prescription, pain management clinic laws, mandated use of prescription drug monitoring programs, and medical cannabis laws) in short and long-term, while we controlled for several protentional risk factors. The results of fixed-effect regression and significant tests indicated that state policies and laws were unlikely to result in an immediate reduction in overdose mortalities and comprehensive interventions were needed to restrain the epidemic. The third essay investigated the risk factors of 30-day readmission in patients with ischemic stroke at an individual level. We aimed to identify the main risk factors of stroke readmissions and prioritized them using machine learning techniques and logistic regression. We also introduced the most effective

predictive model based on different performance metrics. We used the electronic health records of stroke patients extracted from two stroke centers within the Geisinger Health System from 2015 to 2018. This data set included a comprehensive list of clinical features, patients' comorbidities, demographical characteristics, discharge status, and type of health insurance. One of the major findings of this study was that stroke severity, insert an indwelling urinary catheter, and hypercoagulable state were more important than generally known diagnoses such as diabetes and hypertension in the prediction of stroke 30-day readmission. Furthermore, machine learning-based models can be designed to provide a better predictive model. Overall, this dissertation provided new insights to better understand the three major challenges of the U.S. healthcare system and improve its outcomes.

A Multi-Level Analysis of Major Health Challenges in the United States Using Data Analytics Approaches

Negar Darabi

ABSTRACT (GENERAL AUDIENCE)

The major goal of a healthcare system can be summarized in three main objectives: preventing preterm birth and premature mortality, advancing the quality of life, and preparing for a good death. Despite all the national efforts to achieve these goals, the U.S. healthcare system still faces many obstacles and crises and suffers from inefficiencies. The U.S. infant mortality rate is still higher than any other comparable advanced country. The opioid overdose death rate has been steadily increasing since 1999 and has risen exponentially in recent years. Hospital readmissions especially in stroke patients impose a substantial cost burden on the healthcare system in the U.S. Also, readmitted stroke patients are at higher risk of mortality compared to the first admission. I believe that industrial engineering and data analytics approaches can help in advancing the understanding of these health challenges, their important risk factors, and effective interventions. In this dissertation, the main focus was on the performance, trends, variations, and processes of the healthcare systems. We applied innovative methods to provide answers to the following questions in three essays: What does make a healthcare system more successful in improving the birth outcomes? What factors do explain mortality from opioid painkillers? What are the determinants of state variations in mortalities from an opioid overdose? What is the impact of states' laws and programs and opioid prescription rates and overdose mortality rates? What are the most important contributors to stroke readmissions? The results of the first essay showed that not all the state's healthcare systems perform the same in terms of reducing unfavorable birth outcomes. States with lower people in poverty and lower African American women were more successful in improving their birth outcomes. The second study revealed that states with a higher share of uninsured people and binge drinkers were suffering from higher opioid overdose deaths. Also, our results implied that in addition to upstream prevention policies, states need to implement downstream programs to curb the epidemic. Finally, the third study showed that the top predictors of stroke readmissions within 30 days were consist of the severity of the stroke, insert an indwelling urinary catheter, being overweight, and malnourished. The results of this dissertation can help to educate policymakers and practitioners at state and organizational level in a way to

better serve the society and ultimately enhance the population health, quality of healthcare, and societal well-being.

Acknowledgments

Now that I have come to the end of a challenging journey, I am glad to have the opportunity to express my deepest gratitude and appreciation to the amazing individuals who supported me throughout the years.

My beloved husband, *Dr. Alireza Ebrahimvandi*. I am lucky and blessed that I have had your endless support from the beginning of this journey. You have always believed in me and encouraged me to go above and beyond. You are my best friend, soulmate, a wonderful colleague, and an amazing husband. Your love and support empower me to overcome any obstacle. Thank you for giving me the hope and courage that I needed every time!

I would like to sincerely thank my amazing advisor, *Dr. Niyousha Hosseinichimeh*. Working with you was one of the greatest honors of my life. You are a wonderful mentor, researcher, and role model. I will always cherish our conversations and friendship and I will truly miss our weekly meetings. Getting a Ph.D. could be very tough. However, you made it a pleasant and memorable experience for me. I am extremely thankful for your endless support, patience, and amazing personality. Thank you for being the best advisor I could ever ask!

I would also like to thank my dissertation committee members *Dr. Konstantinos P. Triantis*, *Dr. James Kong*, and *Dr. Vida Abedi*. It has been a privilege and honor working with you all. Thank you for your valuable feedback on my research. This doctoral dissertation would not have been accomplished without your support.

I am also grateful for the opportunity to work with my wonderful coauthors *Dr. Konstantinos P. Triantis*, *Dr. Vida Abedi*, and *Dr. Ramin Zand*. Dr. Triantis taught me the foundations of efficiency and performance measurements which opened the research idea for my first paper. The third paper of this dissertation was enabled by valuable advice and inputs from Dr. Abedi and Dr. Zand. I am very fortunate to have crossed paths with you!

I would like to thank *Dr. Navid Ghaffarzadegan*. You are one of the best teachers that I have ever had. I learned a lot from your lectures in the applied systems engineering class. You made me passionate and interested in System Dynamics. I am sincerely grateful for your support and guidance throughout my Ph.D.

Last but not least, I would like to express my special thanks to *my parents* and *my family* for their unconditional love, eternal support, and endless sacrifices. I am blessed to have a family who

always encouraged me to pursue higher education, be persistent, and do my best to make my dreams come true.

Table of Contents

Chapter 1. Introduction	1
1.1 Problem context.....	1
1.2 Research contributions	2
1.2.1 Essay I.....	2
1.2.2 Essay II.....	3
1.2.3 Essay III.....	5
References.....	8
Chapter 2. A DEA evaluation of states' infant mortality rate in the U.S.	10
Abstract	10
2.1 Introduction	10
2.2 Method	13
2.2.1 Data Envelopment Analysis	13
2.2.2 Heterogeneity in DMUs and Bootstrapping in DEA.....	14
2.3 Model Framework.....	15
2.3.1 Input and Output Variables	15
2.3.2 Data.....	17
2.4 Results	18
2.4.1 Outlier Analysis	18
2.4.2 Model Selection and Results Interpretation (Meta-Frontier Approach)	21
2.4.3 Model Selection and Results Interpretation (Bootstrapping).....	25
2.4.4 Hypotheses Testing	25
2.4.5 Slack Analysis and Policy Insights	26
2.5 Discussion	27
2.6 Implication	29

2.7	Future Work	29
	References.....	31
Chapter 3. <i>Associations between State Policies on Opioid Prescriptions and Opioid Overdose Death Rates in the United States, 2008–2017</i>.....		
	<i>Abstract</i>	33
3.1	Background	33
3.2	Study Data and Method	35
3.2.1	Data.....	35
3.2.2	Key Variables	35
3.2.3	Analyses.....	36
3.2.4	Limitations	37
3.3	Study Results	37
3.3.1	Variation in Opioid Overdose Mortality.....	37
3.3.2	Opioid Prescribing Trends.....	38
3.3.3	Opioid Overdose Death Rate Trends	39
3.3.4	Descriptive Statistics	39
3.3.5	Impact of State Policies and Other Factors on Opioid Prescribing Rates.....	40
3.3.6	Impact of Policies and Other Factors on Opioid Overdose Death Rates	41
3.4	Discussion	43
3.5	Conclusion.....	45
	References.....	46
Chapter 4. <i>Machine Learning-Enabled 30-day Readmission Model for Stroke Patients</i>		
	<i>Abstract</i>	48
4.1	Introduction	49
4.2	Methods	50

4.2.1	Database Description	50
4.2.2	Study Variables	50
4.2.3	Data Processing, Feature Selection, and Sampling.....	51
4.2.4	Model Development	52
4.2.5	Comparison Measures	52
4.3	Results	53
4.3.1	Study Design and Population Characteristics	53
4.3.2	Results of the Logistic Regression	55
4.3.3	Results of the Machine Learning Methods	56
4.3.4	Model Comparison and Performance	58
4.4	Discussion	60
4.5	Strengths, Limitations, and Future Directions	61
4.6	Conclusion.....	62
	References.....	64
Chapter 5.	Conclusion.....	66
5.1	Contributions.....	67
5.2	Broader Impacts.....	69
5.3	Implication	70
	References.....	70
Appendices		71
Appendix A	Bootstrapping, Robust PCA, and Slack Analysis	71
Appendix B	Pairwise Regression	80
Appendix C	Correlation Matrix.....	82

List of Figures

Figure 1. Robust PCA analysis for outlier's identification including inputs and outputs.....	20
Figure 2. Robust PCA analysis for outlier's identification including input, output, and environmental variables.....	20
Figure 3. Robust PCA analysis for outlier's identification including only environmental variables	21
Figure 4. Changes in opioid mortality rates in all 50 states and DC from 2008 to 2017.....	38
Figure 5. The number of states with different laws and programs to combat opioid epidemic along with the mean rates of opioid overdose mortality and opioid prescriptions in the U.S. from 2008 to 2017.....	39
Figure 6. ROC curves for ML models.	59
Figure 7. ROC curves for ML models with ROSE-sampling.	59

List of Tables

Table 1. Summary of the three essays.	7
Table 2. Preliminary list of input and output variables.	17
Table 3. Datasets used.	18
Table 4. Descriptive statistics of input, output, and environmental variables, 2017.	18
Table 5. Included variables in different DEA models.	22
Table 6. Descriptive statistics of efficiency scores in different DEA models.	22
Table 7. DEA results of BCC model for 50 states of the U.S. for Model 1.	24
Table 8. Bootstrapped regression results on the environmental factor	26
Table 9. Summary data on opioid overdose death rates and state-level variables from 2008 to 2017.	40
Table 10. Fixed-effect regression models of opioid prescription rates, for 50 states and DC from 2008 to 2017.	41
Table 11. Fixed-effect regression models of opioid overdose mortality rates, for 50 states and DC from 2008 to 2017.	43
Table 12. Descriptive statistics of continuous variables.	54
Table 13. Logistic regression analysis results for predictors of 30-day readmission in stroke patients.	56
Table 14. Performance metrics for machine learning models.	57
Table 15. Performance metrics, with ROSE sampling, for machine learning models.	57
Table 16. Performance metrics, with feature selection and ROSE sampling, for machine learning models.	58
Table 17. Variable importance scores of XGBoost model with ROSE sampling.	58

Chapter 1. Introduction

1.1 Problem context

The modern organized medicine began in the U.S. in the early 1900s. Physicians formed powerful associations like the American Medical Association (AMA), and the government passed legislation to overcome national and organizational health issues (1, 2). Despite all the national advances in the health and healthcare arena, the U.S. healthcare system still faces many unresolved and emerging crises that need to be addressed. In this doctoral dissertation, I plan to study some of the most critical healthcare challenges in different domains of the U.S. healthcare system using diverse levels of analysis.

Three specific aims guide this dissertation. First, I aim to study disparities in states' birth outcomes and provide insights on how well healthcare systems are performing with respect to major health indicators. Second, I investigate one of the current national crises—opioid overdose deaths—and study the impact of laws and programs implemented by state governments on the opioid epidemic. Third, I explore risk factors of 30-day readmission among stroke patients which improves the quality of care and assists healthcare providers to avoid penalties introduced by new regulations. To address these aims, I use industrial engineering and data science techniques. The proposed dissertation advances our understanding of how healthcare systems are performing relative to their benchmarks and explains variation in opioid overdose mortalities, as well as adding to the body of literature by exploring the risk factors of readmission for an acute condition using advanced techniques not used in the past studies of stroke patients.

The first essay focuses on the birth outcomes of healthcare systems at the state level and attempts to discern why some states are more successful than others in terms of improving their birth outcomes such as infant mortality rates (IMR), preterm birth rates, and low birthweight rates. In the second essay, I exploit state variations in opioid overdose mortalities in order to examine factors affecting mortalities from these painkillers and the effectiveness of policies that might mitigate the crisis. In essay three, I study risk factors that are associated with hospital readmissions of stroke patients within 30 days of discharge. To understand the urgency of these topics, it is important to look at statistical evidence and trends. First, in 2017, more than 22,000 infants died in the U.S. (3). The IMR in the U.S. is 71% higher than the mean IMR of comparable developed countries (4). Second, the number of deaths from opioid overdose in the U.S. has been

steadily increasing since 1999 and has risen exponentially in recent years. The mortality rate from overdose deaths in 2017 was 9.9 times the death from opioid overdoses in 1999 (5). Data from 2016 and 2017 shows that every day in the U.S. more than 130 people died from opioid drug overdoses (6). Third, hospital readmissions impose a substantial cost burden on the healthcare system in the U.S. It annually costs Medicare about \$26 billion (7). Centers for Medicare and Medicaid Services set 30-day readmissions as an indicator for poor hospital care and penalizes hospitals to reduce readmission rates as a national healthcare reform goal (8). On the other hand, more than 795,000 people experience one type of stroke every year in the U.S. which results in a \$34 billion cost for the healthcare system annually (9). Therefore, addressing these three challenges is valuable in terms of reducing the financial burden on the U.S. healthcare system and it can also help save many lives.

1.2 Research contributions

1.2.1 Essay I

In the first essay, I focus on evaluating the efficiency of healthcare systems in terms of their birth outcomes across states. Infant mortality rate (IMR) has been a major public health concern in the U.S. over the past decades. According to U.S. healthy people guidelines, reducing IMR is a top priority public goal in 2020 (10). In this work, I use mixed-method approaches such as data envelopment analysis (DEA), robust principal component analysis (ROBPCA), bootstrapping, and statistical (i.e. truncated regression) method to quantify the efficiency of states' health systems in reducing the unfavorable birth outcomes such as IMR. To our knowledge, there was no study that compared the performance of healthcare systems of the United States in terms of their birth outcomes at the state level. Prior studies that used DEA, compare the efficiency of healthcare systems in terms of life expectancy, adult mortality, or infant mortality across different countries (11-13) or consider one country and evaluate its healthcare system's performance over several years (14). However, there is no study in the literature that evaluates the healthcare performance of the U.S. in terms of birth outcomes (i.e., IMR, preterm birth, and low birthweight rate) at the state level. There are some articles that examine infant mortality rates of neighboring states using statistical methods. But they compare a few numbers of states just based on their geographical location, rather than using a mathematical model of choosing benchmarks. Therefore, this is the first study that uses DEA—which is a well-established method of comparison—to find the best practices for states that suffer from poor birth outcomes (i.e., infant mortality, preterm birth, and

low birthweight). To do so, I define the following research question: *what characteristics do make a healthcare system more successful in improving the birth outcomes at the state level?*

To conduct this study, we first construct a unique state-level database using multiple data sources (i.e. Centers for Disease Control and Prevention, Kaiser Family Foundation, and Area Health Resources File (AHRF)). Then we apply the DEA method to systematically benchmark, compare, and contrast healthcare systems' performance with respect to the birth outcomes. Using this method, we calculate efficiency scores for states' health systems based on the defined set of input and output variables. Then, we test the effect of contextual factors on the effectiveness of states' health systems. In a nutshell, the DEA method helps us to quantify the efficiency of each healthcare system, identify best practices for inefficient systems, and introduce opportunities for improvement.

The insights of our work provide policymakers a better understanding of how states' healthcare systems are performing and how efficient they are in achieving desirable birth outcomes (i.e., reducing IMR, preterm birth rate, and low birthweight rate). We introduce efficient and inefficient health systems and we identify the best practices for each inefficient state that suffers from poor birth outcomes. Additionally, we shed light on the impact of contextual variables on the outcomes. We clarify how disregarding these environmental factors leads to an overestimation of the system's efficiency. Furthermore, our findings can be informative for practitioners in the health systems in the identification of those who are at higher risk of giving unfavorable birth outcomes and need special clinical attention. Last but not least, our work provides an application contribution to the body of literature by showing how a well-established operation research method like DEA can be applied to provide a better understanding of a multi-dimensional health-related problem at the state level.

1.2.2 Essay II

In the second essay, I concentrate on the opioid epidemic in the United States, which has been declared a national emergency in 2017 (15). I exploit state variations in opioid overdose deaths to examine factors affecting mortalities from these painkillers. The effects of state laws and programs (i.e., dose and duration limitations on the initial prescription, pain management clinic laws, mandated use of prescription drug monitoring programs (PDMPs), and medical cannabis laws) on the opioid epidemic are arguably unknown and even contradictory. Therefore, I aim to assess the impact of these four laws and programs that have been enacted to curb the opioid

epidemic. I plan to address several major gaps in the literature. Past state-level studies have investigated the impact of policies such as PDMPs, medical cannabis laws, and pain management clinic laws on the number of mortalities from an opioid overdose (16-18). However, these studies have found conflicting results regarding the effect of mandated PDMPs on opioid overdose deaths. One study (17) found that mandating PDMPs along with pain clinic law implementations could reduce opioid overdose deaths by 12%, while others did not find a significant impact (16, 18). Also, state-level studies of opioid overdose deaths have included few or no explanatory factors, although individual-level studies indicate that factors such as type of health insurance, binge drinking, and tobacco use are significantly associated with opioid mortalities. Moreover, some of these studies did not consider the lag between policy implementations and change in outcomes. A delay may exist in the effect of opioid policies on overdose deaths. Also, as most opioid-related policies were implemented after 2013—and the timelines of prior studies range from 1999 to 2013—the investigation into long-term impacts of the policies has been limited. As such, considering the sharp growth in deaths after 2013, it is critical to analyze recent data (19). Therefore, we define specific questions in this study to address these gaps: What factors are associated with mortality from opioid painkillers? What are the determinants of state variations in mortalities from an opioid overdose? Is there any correlation between states' laws and programs, and opioid prescription rates and overdose mortality rates?

To do so, we selected several independent variables that were shown to be important in prior studies. These variables include the median income of household, insurance coverage rate, tobacco usage, and binge drinking. We also examined the effect of four state programs and laws that were enacted to combat the opioid epidemic with and without their lags using fixed-effect regression. We created a panel dataset for this purpose, which includes a 10-year time frame from 2008 to 2017 to be able to analyze the most recent U.S. data available.

Our finding shows that the implementation of pain management clinic laws, mandated PDMPs, and limitations on dose and duration of first prescriptions were significantly associated with fewer opioid prescribing rates from 2008 to 2017. However, we did not find a negative association between these laws and the rate of opioid overdose deaths, which may imply that these laws influence death rates with a delay. States with a higher share of uninsured and binge drinkers have higher overdose death rates. These findings suggest that mandating the use of PDMPs for prescribers and dispensers and reducing the number of uninsured and binge drinkers might lower overdose deaths rate in the long run.

1.2.3 Essay III

In the third essay, I focus on identifying predictors of 30-day readmissions after ischemic stroke at the individual-level. The high rates of readmissions are associated with significant cost burden on the U.S. healthcare system and readmitted patients are at greater mortality risks compared to the first admissions (20). I aim to identify and prioritize factors associated with 30-day readmissions after stroke which is the first time that such comparison has been reported for 30-day readmissions in stroke. Further, I want to find the most effective predictive model by comparing the results of different machine learning techniques and logistics regression. After performing a literature review, I identified several gaps that need to be addressed. First, most studies use multivariate logistics regression to develop predictive models. Second, these studies only consider the main effects of the risk factors. However, the relationship between independent factors can be more complicated which may involve interactions among variables. Therefore, one goal of this study is to improve prediction power by applying more advanced methods such as machine learning (ML) techniques. Third, many studies consider readmission after 90 days or one year as a dependent variable which is a long follow-up period, as CMS penalizes healthcare systems for readmission under 30 days. Fourth, a majority of past studies include a limited number of risk factors and few clinical diagnoses or procedures in their predictive model.

In this work, we address all of these gaps and improve the prediction performance of readmissions in stroke patients using a wide range of potential risk factors and the proper machine learning techniques. I use a novel dataset including the electronic health records (EHR) of stroke patients extracted from two stroke centers within Geisinger Health System from 2015 to 2018. This data set includes a comprehensive list of potential predictors including patient's comorbidities, demographical characteristics, clinical intervention, discharge status, and type of health insurance. Further, I compare the results of ML models and the traditional logistic regressions to find out the significant predictors of readmission risk as well as the best predictive model.

The result of this work will provide a better understanding of the main drivers of hospital readmissions after stroke. We will also prioritize and quantify a combination of individual-level risk factors for hospital readmissions. The proposed model may facilitate the identification of patients who are at higher readmission risks after stroke and need special clinical attention. This can help healthcare managers and practitioners to provide better care delivery and reduce stroke-related readmissions and financial costs.

Overall, this dissertation advances the understanding of three major healthcare challenges in the U.S. from a multi-level perspective. I provide health professionals and policymakers with theoretical and empirical foundations of risk assessment for reducing unfavorable birth outcomes, opioid-related mortalities, and unplanned 30-day readmissions in stroke patients. Table 1 shows a summary of the research objective(s), gaps, research questions, contributions, and methods that I use for each specific aim. This dissertation proceeds from three essays in chapters two, three, and four, in which each essay is a stand-alone paper. In the fifth chapter, I provide a conclusion including the major findings, contributions, and broader impacts.

Table 1. Summary of the three essays.

Gaps	Objectives	Research Question	Methods	Contributions
1) No study evaluates the performance of states' healthcare systems in terms of their birth outcomes in the U.S. 2) There is no study that compares the birth outcomes of states and provides benchmarks using data envelopment analysis (DEA).	We analyze the performance of states' healthcare systems regarding their birth outcome to identify the best practices for inefficient states. We also aim to explore the effect of contextual variables such as race on the efficiency scores.	1) What factors do make a healthcare system more successful in improving the birth outcomes? 2) Which states' health systems are performing better in terms of their birth outcomes? 3) What is the effect of contextual variables on the performance of healthcare systems?	Mixed methods: DEA, ROBPCA, Bootstrapping, and Truncated regression	1) Classifying which state's health system was efficient or inefficient with the DEA approach using a unique dataset that has not been studied before. 2) Identifying benchmarks and explore rooms for improvement for inefficient states. 3) Identifying the effect of contextual variables on the healthcare systems' outcome and efficiency scores.
1) Past state-level studies have included few or no explanatory factors. 2) They did not consider lags between policy implementations and their visible effects. 3) The timeframe of these studies ranges from 1999 to 2013, however, considering the sharp growth in deaths after 2013, it is critical to analyze recent data.	We exploit state variations in overdose deaths from opioids to examine factors affecting mortality from these painkillers. We also aim to investigate the effect of state programs on the opioid epidemic.	1) What factors do explain mortality from opioid painkillers? 2) What are the determinants of state variations in opioid mortalities? 3) Is there any correlation between state laws and programs and overdose mortalities?	Fixed-effect	1) Identifying factors associated with state variation in opioid mortality rates. 2) Exploring the effect of four state programs and laws on the opioid epidemic. 3) Considering the effects of these laws with delays. 4) Analyzing the most recent data that was available.
1) Most studies used logistics regression. 2) Only the main effects of the risk factors with no interactions were considered. 3) Most of these works considered a long follow up period for readmission as a dependent variable. 4) A few risk factors or predictors were considered in their model and most cases none or a limited number of included clinical diagnoses and procedures.	We explore the risk factors of 30-day readmissions in stroke patients and provide a better understanding of the determinants of readmissions in the Geisinger Health System. Also, we aim to identify the most effective predictive model.	1) What risk factors are the most important predictors of 30-day readmissions for ischemic stroke patients? 2) What is the best predictive model for estimation of the chance of 30-day readmission in ischemic stroke patients?	LR and ML	1) Improve prediction performance of readmission scores and identified a wide range of potential risk factors using different ML techniques and logistics regression. 2) Provide insights to help healthcare policymakers to enhance the quality of care and eliminate stroke-related readmissions and financial costs. 3) Facilitate the identification of patients who have higher chances of readmission after stroke and need special clinical attention.

References

1. MidlevelIU. Timeline: History of the U.S. Health Care System 2013 [Available from: <https://midlevelu.com/blog/timeline-history-us-health-care-system>.]
2. Allison Wasalaski AD, Alexandria Holladay, and Issabella Sanderson. United States Health Care Timeline 2018 [Available from: <https://www.sutori.com/story/united-states-health-care-timeline--2BBv9e4GsgGgsSDU6irbPW6K>.]
3. Centers for Disease Control and Prevention C. Infant Mortality 2019 [Available from: <https://www.cdc.gov/reproductivehealth/maternalinfanthealth/infantmortality.htm>.]
4. Sawyer B, McDermott D. How does infant mortality in the US compare to other countries? Peterson-Kaiser Health System Tracker. 2019.
5. Kaiser Family Foundation. Opioid Overdose Deaths and Opioid Overdose Deaths as a Percent of All Drug Overdose Deaths 2019 [Available from: <https://www.kff.org/other/state-indicator/opioid-overdose-deaths/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D>.]
6. U.S. Department of Health and Human Services. The opioid epidemic in numbers, 2017 [Available from: <https://www.hhs.gov/opioids>.]
7. LaPointe J. 3 Strategies to Reduce Hospital Readmission Rates, Costs 2018 [Available from: <https://revcycleintelligence.com/news/3-strategies-to-reduce-hospital-readmission-rates-costs>.]
8. Lichtman JH, Leifheit-Limson EC, Jones SB, Wang Y, Goldstein LB. Preventable Readmissions Within 30 Days of Ischemic Stroke Among Medicare Beneficiaries. *Stroke*. 2013;44(12):3429-35.
9. Center for Disease Control and Prevention. Stroke Facts 2017 [Available from: <https://www.cdc.gov/stroke/facts.htm>.]
10. Office of Diseases Prevention and Health Promotion. Maternal, Infant, and Child Health 2020 [Available from: <https://www.healthypeople.gov/2020/topics-objectives/topic/maternal-infant-and-child-health>.]
11. Mirmirani S, Li H, Ilacqua JA. Health care efficiency in transition economies: an application of data envelopment analysis. *International Business & Economics Research Journal (IBER)*. 2008;7(2).
12. Benneyan J, Ceyhan ME, Sunnetci A, editors. Data envelopment analysis of national healthcare systems and their relative efficiencies. the 37th international conference on computers and industrial engineering; 2007.
13. Cetin VR, Bahce S. Measuring the efficiency of health systems of OECD countries by data envelopment analysis. *Applied Economics*. 2016;48(37):3497-507.
14. Ibrahim MD, Daneshvar S. Efficiency analysis of healthcare system in Lebanon using modified data envelopment analysis. *Journal of healthcare engineering*. 2018;2018.
15. Christie C, Baker C, Cooper R, Kennedy PJ, Madras B, Bondi P. The president's commission on combating drug addiction and the opioid crisis. Washington, DC: United States Government Printing Office. 2017.
16. Bachhuber MA, Saloner B, Cunningham CO, Barry CL. Medical cannabis laws and opioid analgesic overdose mortality in the United States, 1999-2010. *JAMA internal medicine*. 2014;174(10):1668-73.
17. Dowell D, Zhang K, Noonan RK, Hockenberry JM. Mandatory provider review and pain clinic laws reduce the amounts of opioids prescribed and overdose death rates. *Health Affairs*. 2016;35(10):1876-83.
18. Paulozzi LJ, Kilbourne EM, Desai HA. Prescription Drug Monitoring Programs and Death Rates from Drug Overdose. 2011.
19. Centers for Disease Control and Prevention. Drug Overdose Deaths in the United States, 1999–2017 2018 [Available from: <https://www.cdc.gov/nchs/products/databriefs/db329.htm>.]

20. Noh AM, McCormick L, Modak J, Fortunato G, Staff I. High mortality among 30-day readmission after stroke: predictors and etiologies of readmission. *Frontiers in neurology*. 2017;8:632.

Chapter 2. A DEA evaluation of states' infant mortality rate in the U.S.

Abstract

States vary substantially in terms of their infant mortality and preterm birth rates. Despite all the attention and effort at national level in improving birth outcomes, still many states' healthcare systems are suffering from high rates of unfavorable birth outcome. Hence, the important question is that what makes a healthcare system more successful in improving the birth outcomes at state level. To answer this question, a state-level database was built using data from the Centers for Disease Control and Prevention, Kaiser Family Foundation, and Area Health Resources File. Then, we implement a mixed method approach that includes Data Envelopment Analysis (DEA), robust principal component analysis, bootstrapping, and statistical methods such as truncated regression to systematically compare and contrast the performance of 50 state's healthcare systems with respect to their birth outcomes (i.e., infant mortality, preterm birth, and low birthweight) and provide benchmarks. Our findings reveal that socio-economic and demographic characteristics such as poverty rate and number of African-American women per 1000 population significantly explain the major variation in healthcare systems' performance in terms of their birth outcomes. Therefore, disregarding these environmental factors leads to over estimation of the system's efficiency. Furthermore, our work contributes to the body of literature by showing how a well-established operations research method like DEA can be applied to better understand multi-dimensional health-related problem at the state level.

2.1 Introduction

Infant mortality is an important indicator of health status and quality of healthcare in any society. Infant mortality rate (IMR) is defined as "the number of infants who die before their first birthday per 1,000 live births". The IMR in the U.S. was 71% higher than developed countries in 2015 (1). Also, after 15 years of decline, the U.S. IMR started to increase in 2015 (1). More than 22,000 infants died in the U.S. in 2017 (2). This high rate of infant mortality, which is higher than other developed countries, has become a central focus of many researchers and policymakers. In addition to the high rate of IMR in the U.S., there are significant variations in IMR across states.

Although some states have been successful in reducing it, there are many states that are still suffering from high rates of infant mortality (3).

States healthcare systems have decentralized decision-making structures; they design the care delivery plans based on specific characteristics of their population. Additionally, heterogeneity exist among states in their demographics and healthcare infrastructure as they are different in their available resources. Some states might be able to better allocate their available resources and improve the birth outcomes, while others might not be that successful. Therefore, identifying the best practices of states' healthcare systems and understanding why they are performing better in reducing unfavorable birth outcomes is crucial for improving the overall U.S. healthcare system and for serving the community better.

Previous studies offered explanations on disparities in birth outcomes across states. Many compared IMRs between White and non-White, others investigated birthweights and infant mortalities of Whites versus Native Americans in the U.S. However, their comparison is limited to a few number of states and they use conventional statistical methods for their analysis (4, 5). Other articles analyzed the birth outcomes of neighboring states (6, 7) using such statistical methods. They relied only on the geographical location to compare the birth outcomes of a few numbers of states; however, there are many other factors like healthcare expenditures, socioeconomic factors, and personal factors that can affect birth outcomes and the efficiency of a healthcare system (8, 9). Therefore, there is a need to use a multi-dimensional quantitative benchmarking approach such as Data Envelopment Analysis (DEA) as a method of comparison (20, 21).

In the area of healthcare efficiency, majority of studies have focused on organizational level comparisons like hospitals (10, 11), or on the national level. Articles using national level data either used data envelopment analysis (DEA) to compare the efficiency of healthcare systems in terms of life expectancy, adult mortality, or infant mortality across different countries (9, 12, 13), or focused on one country and evaluate the healthcare system over several years (8). However, no study has compared the performance of healthcare systems of all 50 states in terms of their birth outcomes (i.e., IMR, preterm birth rate, and low birthweight rate). Thus, the major research question that we aim to answer is: *which factors make a healthcare system more successful in terms of achieving successful birth outcomes at the state level?*

One of the main points of departure from prior studies that focused on the performance measurement of healthcare systems lies in considering the impact of contextual variables on the

outcomes of healthcare systems. Prior studies examined the efficiency of healthcare systems only from an operational characteristics point of view and did not consider the effect of environmental variables (8, 9, 12, 13). Exploring the literature surrounding the most important risk factors associated with unfavorable outcomes guided us toward the impact of demographical and socio-economic factors on unfavorable birth outcomes. Scholars in this area introduced race as one of the most important predictors of preterm births (3). Additionally, studies focusing on race and ethnic disparities in infant mortality in the US showed that African-American women are at greater risk of giving unfavorable birth outcomes (14). These women were exposed to an interpersonal racial discrimination which was associated with being at higher risk of preterm delivery (15). Hence, considering the number of African-American women in each state when making decisions at the state level is crucial. By incorporating this demographical factor in the model and exploring its effect on the states' birth outcomes, one can provide more comprehensive insights for policymakers. Therefore, one of the important hypotheses that we aim to test in this paper is:

Hypothesis 1: States with a lower rate of African-American women are more successful in reducing the rate of unfavorable birth outcomes.

Scholars who studied the risk factors of unfavorable birth outcomes found that neighborhood poverty was one of the important factors associated with higher chances of giving preterm birth and low birthweight (16). Other studies showed an association between infant mortality rate and living in the disadvantaged area with higher poverty (17-19). Therefore, it is important to consider the impact of poverty on birth outcomes. Including this environmental factor in our analysis may shed light on the effect of living in poverty as a socio-economic factor on the performance of healthcare systems in each state. The second hypothesis that we would like to test is:

Hypothesis 2: State with higher poverty rate are less successful in reducing unfavorable birth outcomes.

In this paper, we analyze the effectiveness of states' healthcare systems with respect to their birth outcomes (i.e. IMR, preterm birth rate, and low birthweight rate) to advance our understanding of how states are performing relative to their best practices. We use mixed method approaches that includes DEA, robust principal component analysis (ROBPCA), bootstrapping, and statistical method such as truncated regression to quantify the effectiveness of each state's efforts in reducing IMR, preterm birth rate, and low birthweight rate considering multiple inputs and outputs. Additionally, we test proposed hypotheses to investigate whether environmental factors such as

number of African-American women and poverty rate obstruct the reduction of unfavorable birth outcomes by applying bootstrapping and truncated regression analysis.

2.2 Method

2.2.1 Data Envelopment Analysis

Data Envelopment Analysis (DEA) is a method in the operations research field for performance measurement and discovery of benchmarks. This methodology evaluates a production process, which consists of inputs and outputs that may be either fixed or variable. The DEA method was introduced by Charnes, Cooper, and Rhodes (20). This method is a multi-criteria performance assessment technique evaluating the relative efficiency of DMUs, which is based on Farrell (21). In the original DEA model known as CCR, the efficiency of a DMU is obtained through maximization of the weighted ratio such that the ratio for all DMUs must be less than or equal to one. The CCR model has a constant return to scale assumption, which was later enhanced to variable-returns-to-scale (VRS) by Banker, Charnes, and Cooper (22) and named BCC model. Like CCR, BCC is a radial-based DEA model. Each of these models may consider an input- or an output-oriented approach. Input-oriented models measure performance based on minimizing inputs while maintaining a constant level of output. Output-oriented models compute efficiency scores by maximizing output levels while at the same time keeping inputs constant.

In this study, we conduct an input-oriented BCC model with the variable returns-to-scale (VRS) assumption. We use VRS because intuitively states with higher health expenditure per capita deliver a better care relative to those that spends less. Moreover, studies in the literature, which applied DEA to compare the efficiency of healthcare systems across different countries considered variable returns-to-scale in their analysis (13). We picked the input-oriented model because we wanted to examine how performance of inefficient states can be improved by using the allocated resources more efficiently. In addition, we do not have any control over the results of the birth outcomes and it is not easy for states to improve the birth outcomes radically.

Linear programming formulation for BCC model with the VRS is provided in equation (2). In this formulation, it is assumed that by increasing the inputs, the outputs do not increase proportionally (i.e. $\sum_{j=1}^n \lambda_j = 1$). Suppose we have n DMUs and each DMU_j ($j = 1, \dots, n$) utilizes $X_j = \{x_{ij}\}$ inputs ($i = 1, \dots, m$) to produce $Y_j = \{y_{rj}\}$ set of outputs ($r = 1, \dots, s$). The efficiency score of each

DMU_j (θ_j) can be define as the weighted summation of outputs divided by weighted summation of input variables:

$$\theta_j = \frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \quad (1)$$

In equation (1), u_r ($r = 1, \dots, s$) denotes output weights and v_i ($i = 1, \dots, m$) represents input weights. The BCC model solves the following linear programming model to evaluates the efficiency of DMU_j ($j = 1, \dots, n$):

$$\min_{\theta, \lambda, s^+, s^-} z = \theta_j - \sum_{r=1}^s s_r^+ - \sum_{i=1}^m s_i^- \quad (2)$$

$$s.t. \quad \theta_j X_j - \sum_{j=1}^n \lambda_j x_{ij} - s^- = 0$$

$$\sum_{j=1}^n \lambda_j y_{rj} - s^+ = Y_j$$

$$\sum_{j=1}^n \lambda_j = 1$$

$$\lambda, s^+, s^- \geq 0$$

where $\lambda = \lambda_j$ ($i = 1, 2, \dots, n$) is a semipositive vector in $\mathbb{R}^+$, s^+ denotes output shortfalls ($s^+ \in \mathbb{R}^s$), and s^- input excesses ($s^- \in \mathbb{R}^m$). When performing DEA analysis, we need to solve n linear programming problems—like model (2)—for each DMU.

2.2.2 Heterogeneity in DMUs and Bootstrapping in DEA

In the DEA analysis, when there is a heterogeneity among DMUs, by using the traditional BCC model—which estimates one technology frontier—we might end up with unreasonable estimates for efficiency scores (23-26). This issue is relevant to healthcare systems because there are variations between the clinical factors (e.g., amount of expenditures per person and insurance coverage) and socio-economic factors, like the average household income and demographics, which result in different birth outcomes. For example, a state like Massachusetts, which has one of the highest health expenditures per capita cannot be pooled with a state like Utah with the

lowest health expenditure per capita. On the other hand, many of these clinical, personal, and socio-economic factors can be considered as inputs or outputs, however, some of them are uncontrollable/environmental. The environment that the transformation process occur in it, can be affected by some uncontrollable variables responsible for underperformance of the DMUs. Triantis, Sarayia (26) and Simar and Wilson (27) provided various methods of incorporating environmental variables into the model. Based on the assumptions that Simar and Wilson (27) proposed in their algorithm, the uncontrollable variables need to be separable from inputs and outputs. Therefore, after checking the significance of correlations between environmental variables and input and output variables, we confirmed that our proposed environmental variables are separable from the set of inputs and outputs. Further, we use a bootstrapping technique and then apply truncated regression to analyze the effect of the uncontrollable variables. These step address the heterogeneity among DMUs. More explanation on bootstrapping approach and its related formula is provided in appendix A.

2.3 Model Framework

2.3.1 Input and Output Variables

Every system has some inputs, which will be transformed to one or more outputs by performing a production process. In our study, the entities that were converting inputs into outputs was the healthcare system of each state. Therefore, we considered 50 states plus the District of Columbia as the Decision-Making Units (DMUs). Also, since different states were following relatively similar policies and regulations at the national-level, we labeled them as DMUs. Hence, we were interested in comparing the performance of healthcare system across states. To get the most accurate results in an efficiency analysis, we needed to include all descriptive variables. Every healthcare system has resources that could be characterized by a set of variables. These variables are medical staff, healthcare professions, health facilities, and hospital expenditures. Healthcare systems perform under these given resources and can be represented by these inputs. On the other hand, the major goals of every healthcare system can be categorized in three objectives: preventing preterm birth and premature mortality, advancing quality of life, and preparing for a good death. Considering these characteristics of healthcare systems and using similar studies in the literature (8, 9, 13, 28-30), we came up with a primary list of variables including five inputs, three outputs, and two environmental variable. The five primary input variables in this study are the number of obstetric and gynecologists, physicians, patient beds,

pregnant women who smoke (per 1,000 people), and amount of healthcare expenditures per capita. Also, we included three output variables in this step: infant mortality rate—which denotes the number of infants who die under 12 months, low birthweight rate (i.e. weight less than 5 lb.), and preterm birth rate—which is the number of babies who were born before 37 weeks of pregnancy per 1,000 live births. Also, two environmental variables were the number of African-American women per 1000 people and poverty rate.

All of the primary outputs are undesirable variables. There are several methods to deal with undesirable variables in the literature (31). We apply the suggested approach by Seiford and Zhu (32), which keeps all the data points positive and does not require any changes in the assumptions of the DEA model. In this approach, the undesirable output (y_r) will be transformed to a desirable variable (y'_r) by this function: $y'_r = -y_r + w > 0$, which is a linear monotone decreasing transformation and holds classification invariance property. We selected $w = 1000$ because the data for infant mortality, low birthweight, and preterm birth rates are reported per 1000 live births. The new desirable outputs can be interpreted as the infant survival rate—which represents the number of babies who survived for at least one year—, the number of infants with normal birthweight (i.e. weight more than 5 lb.), and the number of full-term¹ infants per 1,000 live births. For example, if the IMR is 6 per 1000, using the abovementioned transformation function, the infant survival rate will be $994 = -6 + 1000$ per 1000 live births.

Finally, we have two uncontrollable or environmental variables which are number of persons in poverty and number of African-American mothers per 1000 people. Literature showed that poverty is one of the risk factors of unfavorable birth outcomes (16-19). Also, historical health records show that African-American women have higher risk of unfavorable birth outcomes (14, 15). Therefore, these influential variables are considered as environmental variables. The list of all inputs, outputs, and environmental variables are provided in Table 2.

¹ A baby born after 37 weeks of pregnancy is considered full-term.

Table 2. Preliminary list of input and output variables.

Type of variable		Name of variable
Inputs	x_1	Number Obstetrics & Gynecologists per 1000
	x_2	Number of physicians per 1000 people
	x_3	Number of patients' beds per 1000 people
	x_4	Amount of healthcare expenditure per capita (\$)
	x_5	Number of pregnant women who smoke per 1000 people
Outputs	y_1	Infant survival rate (per 1000 live birth)
	y_2	Number of babies with normal birthweight (per 1000 live birth)
	y_3	Number of full-term infants (per 1000 live birth)
Environmental	z_1	Number of African-American mothers (per 1000)
	z_2	Number of persons in poverty (per 1000)

2.3.2 Data

In this study, we created a unique state-level dataset, a cross-sectional data of 50 states for the year 2017, that has not been studied before. We obtained the U.S. linked birth dataset from the CDC and combined it with two other state-level datasets, Area Health Resources File (AHRF) and Kaiser Family Foundation (KFF) (see Table 3). The uniqueness of this dataset lies in the reliability and the diversity of the included factors that were gathered from different sources. This dataset encompasses information from birth certificates, death certificates, healthcare resources, professions, and health expenditures from all 50 states.

The linked birth dataset provides information from birth and death certificates for every baby who died before 12 months old in the U.S. This linked dataset includes infants' information such as age, birth weight, gestational age, and some information regarding their parents such as race and maternal smoking status (33). The Area Health Resources File provides data on the healthcare resources of 50 states. It includes the number of general physicians, the number of obstetric and gynecologists, and hospital resources such as the number of hospital beds (34). After combining the three datasets, we performed data cleaning and preparation in *STATA 14.0* and coded the mathematical models in *R 3.6.0*. The most recent data for some of the inputs and outputs belonged to 2018; however, for some other variables, the data was available only until 2017. We provide the list of inputs, outputs, and the environmental variables along with their sources and descriptive statistics in Table 4.

Table 3. Datasets used.

Dataset	Data Provider	Year
Linked Period Birth/Infant Death Records	CDC	2017
Kaiser Family Foundation	KFF	2017
Area Health Resources File	U.S. Area Health Resources and Service Administration (AHRF)	2017

Table 4. Descriptive statistics of input, output, and environmental variables, 2017.

	Variables	Mean	Standard Deviation	Median	Min	Max	Data source
x_1	Number Obstetrics & Gynecologists per 1000	0.12	0.038	0.11	0.07	0.31	AHRF
x_2	Number of physicians per 1000 people	2.86	1.058	2.67	1.72	8.46	AHRF
x_3	Number of patients' beds per 1000 people	3.13	0.909	2.85	1.90	6.27	AHRF
x_4	Amount of healthcare expenditure per capita (\$)	8,332	1,256	8,108	5,982	11,944	KFF
x_5	Number of pregnant women who smoke per 1000 people	96.27	46.754	90.75	19.78	252.42	CDC
y_1	Infant survival rate (per 1000 live birth)	994.03	1.202	994.10	991.27	996.34	CDC
y_2	Number of infants with normal birthweight per 1000 live birth	986.13	2.876	985.98	978.45	990.87	CDC
y_3	Number of full-term infants per 1000 live birth	900.84	12.162	903.66	864.52	924.85	CDC
z_1	Number of African-American mothers per 1000	146.35	125.501	125.68	6.44	544.14	CDC
z_2	Number of persons in poverty per 1000 people	135.84	30.358	134.00	76.00	210.00	AHRF

2.4 Results

2.4.1 Outlier Analysis

To identify the influential observations, we needed to select the principal components (i.e., the best linear combination of the variables). These influential observations were known as outliers and identification of such factors provided a measure for the ratio of outliers compared to the regular data points and it was a measure for potential heterogeneity. We conducted the ROBPCA to distinguish the influential observations (outliers) from regular observations. This helped us calculate the vertical distance of observations from the principal hyperplane. This distance is known as “orthogonal distance”. We could also measure the horizontal distance of observations from the center of the principal hyperplane or “score distance”. These two values provided a robust measure of heterogeneity among observations. More explanation on the ROBPCA

approach and how to calculate orthogonal distance and score distance is provided in Appendix A. The analysis of ROBPCA was conducted using “robustbase” and “rrcov” packages in RStudio.

We performed the ROBPCA analysis for three sets of variables. First, we only included inputs and outputs (see Figure 1), then we considered all input, output, and environmental variables (Figure 2), and lastly, we included only environmental variables (Figure 3). As depicted in Figure 1, two good leverage points (green dots), Mississippi and West Virginia were identified. Also, in this figure, nine orthogonal outliers (blue points) were found (i.e., Connecticut, District of Columbia (DC), Maryland, Massachusetts, New York, North Dakota, Rhode Island, South Dakota, and Vermont). These DMUs showed a high output production without the corresponding input consumption. As it was shown in Figures 1, 2, and 3, the distribution of data points was different especially in Figure 3 where we only considered environmental variables.

Figure 1 and Figure 2 identified the principal components in terms of the operational characteristics of the transformation functions. These two figures show that the only orthogonal outlier that was located far from the other DMUs was DC. We also did not find any bad leverage points. However, once we considered only the environmental variables, four additional DMUs (Louisiana, Maryland, Mississippi, and New Mexico) (see Figure 3) switched to bad leverage points. Also, as shown in Figure 3, the distance between orthogonal outliers further increased and more DMUs were located close to DC. This highlights the existence of heterogeneity among DMUs. In another words, states varied substantially in the rate of poverty and African-American population. According to the results of ROBPCA, we kept all the data points except DC because its demographic was very different from the other 50 states and we continued the DEA analysis and bootstrapping with 50 states.

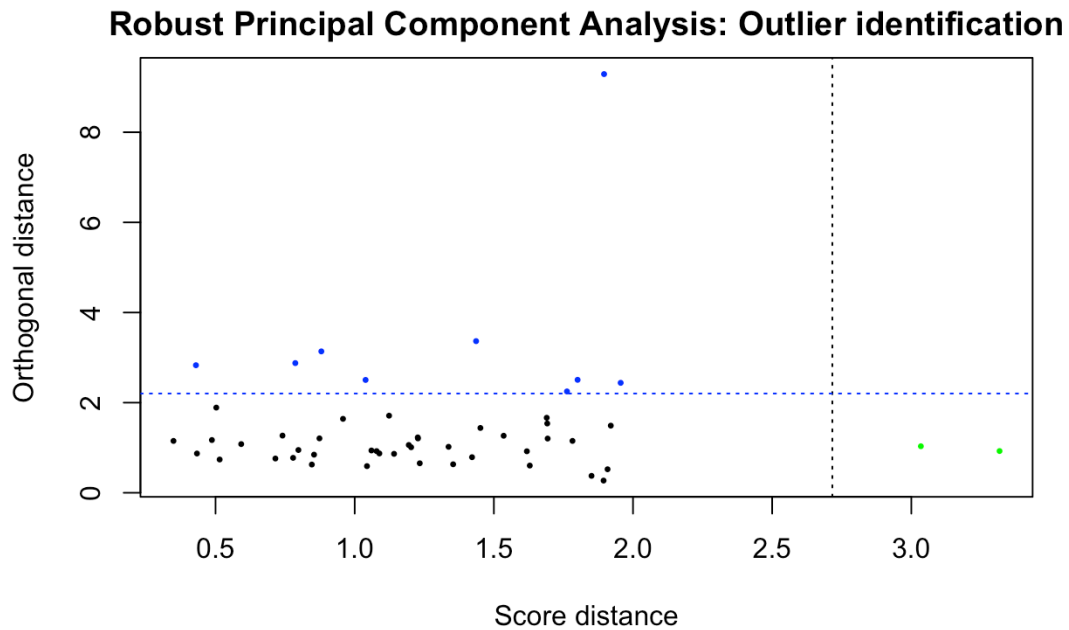


Figure 1. Robust PCA analysis for outlier's identification including inputs and outputs.

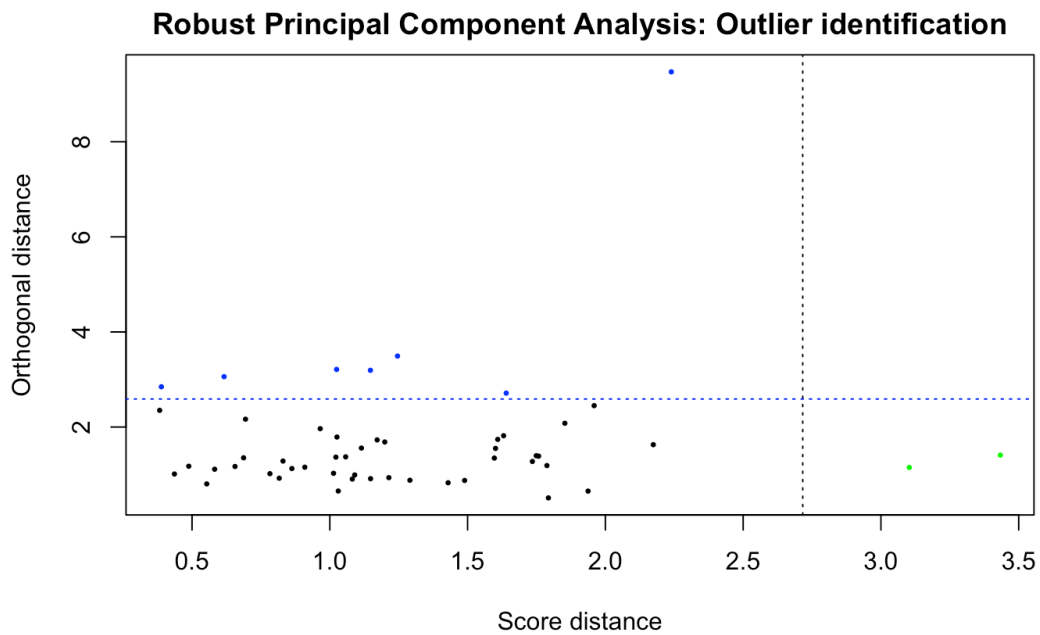


Figure 2. Robust PCA analysis for outlier's identification including input, output, and environmental variables.

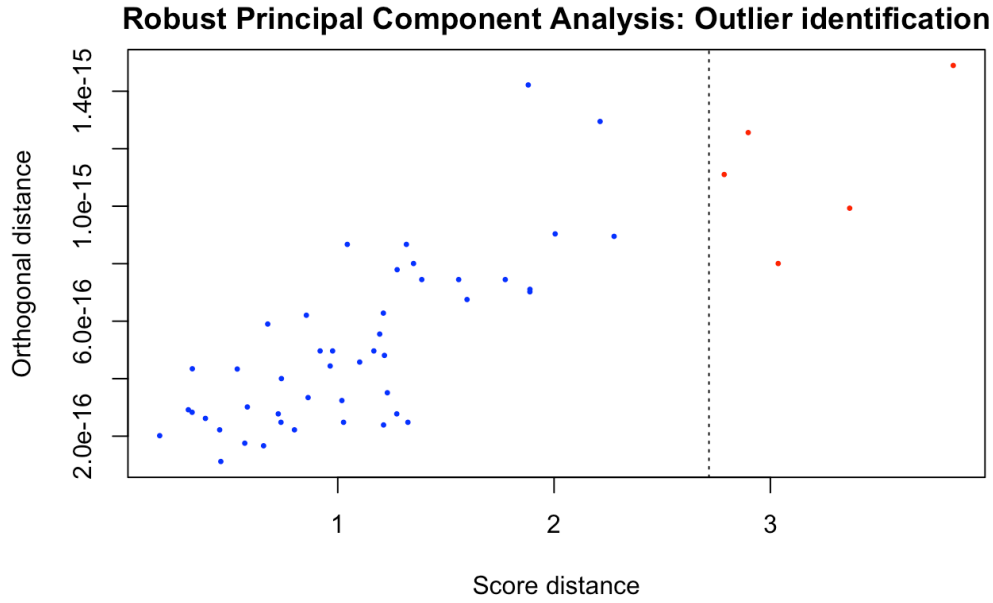


Figure 3. Robust PCA analysis for outlier's identification including only environmental variables

2.4.2 Model Selection and Results Interpretation (Meta-Frontier Approach)

The analysis of efficiency for 50 states is performed using the “Benchmarking” R-package in RStudio. We applied BCC input-oriented with variable return to scale (VRS)². As it is shown in

Table 5, we apply three different models. All the input variables are included in these three models, but each of them has different outputs. Model 1 estimates efficiency scores including three outputs. In Model 2, we included only full-term birth rates and normal birthweight rates, which the associated undesirable outputs (i.e., preterm birth rate and low birthweight rate) can be considered as reversible events. Finally, in Model 3, we only considered infant survival rate as the only output variable.

² We also applied BCC output-oriented model. The efficiency scores for BCC output-oriented were different from input-oriented. Therefore, the VRS is hold and one can say VRS is a reasonable assumption.

Table 5. Included variables in different DEA models.

Type of variable		Name of variable	Model 1	Model 2	Model 3
Inputs	x_1	Number of Obstetrics & Gynecologists per 1000 people	✓	✓	✓
	x_2	Number of physicians per 1000 people	✓	✓	✓
	x_3	Number of patients' beds per 1000 people	✓	✓	✓
	x_4	Amount of healthcare expenditure per capita	✓	✓	✓
	x_5	Number of pregnant women who smoke per 1000 people	✓	✓	✓
Outputs	y_1	Infant survival rate (per 1000)	✓		✓
	y_2	Number of babies with normal birthweight (per 1000)	✓	✓	
	y_3	Number of full-term babies (per 1000)	✓	✓	

The descriptive statistics under BCC DEA model are provided in Table 6. In Model 1, which is considered as our full model, 12 DMUs out of 50 were identified as efficient. Among these 12 DMUs, 10 (20% of the 50 states) were considered as BCC-efficient with zero slacks. However, two DMUs (4%) were identified as radial-efficient with non-zero slacks. Model 2 and Model 3 had only six (12%) and five (10%) DMUs as BCC-efficient, respectively. Both models had more radial-efficient DMUs (i.e., 8%) compared to Model 1. The range of efficiency scores for Model 1 was wider with minimum and maximum efficiency score of 0.700 and 1.000. Model 3 which only includes infant survival rate as output, had fewer number of efficient DMUs compared to Models 1 and 2, but with a range that was not as wide as Model 1 (Table 6).

Table 6. Descriptive statistics of efficiency scores in different DEA models.

DEA Models	Average	Standard Deviation	Minimum	Maximum	BCC-efficient DMUs	Radial-efficient DMUs	Number of inefficient DMUs
Model 1	0.877	0.093	0.700	1.000	10	2	38
Model 2	0.996	0.003	0.989	1.000	6	4	40
Model 3	0.998	0.001	0.995	1.000	5	4	41

We provided the results of Model 1 including the efficiency scores and list of peers for inefficient states in Table 7. For each peer, we reported the intensity variable (i.e., λ), in which the highest value for λ indicates the primary reference peer. Based on this table, California, Colorado, Idaho, Iowa, Massachusetts, Nevada, Oregon, Texas, Utah, Vermont, Washington, and Wyoming were identified as efficient states. These states were using their resources efficiently in comparison

with inefficient states. Comparing the efficiency scores found by Model 1 with the results of ROBPCA highlighted important points. All the good leverage points and orthogonal outliers found by considering only operational characteristics of the healthcare systems are identified as inefficient DMUs (Massachusetts and Vermont were exceptions). Also, all the bad leverage DMUs identified by including only environmental characteristics of states in the ROBPCA were found inefficient. Therefore, the results of the meta-frontier approach were in line with what we found in the robust PCA analysis. Additionally, comparing the results of these two approaches guided us toward the potential impact of environmental factors on the performance of DMUs. Hence, in the next section, we incorporated the impact of environmental variables in the model using a bootstrapping approach.

Table 7. DEA results of BCC model for 50 states of the U.S. for Model 1.

State			Eff. Score	Peers (λ)	State			Eff. Score	Peers (λ)
1	Alabama		0.873	ID (0.39), UT (0.61)	26	Montana		0.896	ID (0.95), WY (0.05)
2	Alaska		0.903	ID (0.44), WA (0.30), WY (0.17)	27	Nebraska		0.769	ID (0.51), UT (0.49)
3	Arizona		0.974	ID (0.27), UT (0.70), CA (0.03)	28	Nevada		1.000	—
4	Arkansas		0.920	ID (0.88), UT (0.12)	29	New Hampshire		0.893	ID (0.17), WA (0.71), VT (0.12)
5	California		1.000	—	30	New Jersey		0.819	ID (0.15), UT (0.11), CA (0.70), CO (0.04)
6	Colorado		1.000	—	31	New Mexico		0.921	ID (0.26), UT (0.52), WA (0.22)
7	Connecticut		0.824	CA (0.75), OR (0.25)	32	New York		0.724	UT (0.18), CA (0.57), CO (0.25)
8	Delaware		0.831	ID (0.56), UT (0.21), WA (0.23)	33	North Carolina		0.861	UT (0.87), WA (0.02), OR (0.11)
9	Florida		0.857	ID (0.05), UT (0.30), NV (0.65)	34	North Dakota		0.970	ID (0.75), WA (0.25)
10	Georgia		0.913	ID (0.03), UT (0.97)	35	Ohio		0.731	ID (0.41), UT (0.59)
11	Hawaii		0.941	UT (0.54), WA (0.04), CA (0.14), OR (0.29)	36	Oklahoma		0.956	ID (1.00)
12	Idaho		1.000	—	37	Oregon		1.000	—
13	Illinois		0.754	UT (0.88), OR (0.12)	38	Pennsylvania		0.700	ID (0.53), UT (0.47)
14	Indiana		0.807	ID (0.76), UT (0.24)	39	Rhode Island		0.879	WA (0.09), CA (0.60), VT (0.31)
15	Iowa		1.000	—	40	South Carolina		0.854	ID (0.28), UT (0.72)
16	Kansas		0.845	ID (0.52), UT (0.48)	41	South Dakota		0.778	ID (1.00)
17	Kentucky		0.819	ID (0.60), UT (0.40)	42	Tennessee		0.819	ID (0.06), UT (0.94)
18	Louisiana		0.765	UT (1.00)	43	Texas		1.000	—
19	Maine		0.759	ID (0.55), UT (0.07), WA (0.38)	44	Utah		1.000	—
20	Maryland		0.841	UT (0.31), CA (0.34), OR (0.035)	45	Vermont		1.000	—
21	Massachusetts		1.000	—	46	Virginia		0.792	UT (1.00)
22	Michigan		0.772	ID (0.25), UT (0.75)	47	Washington		1.000	—
23	Minnesota		0.775	ID (0.30), UT (0.33), WA (0.27), CO (0.10)	48	West Virginia		0.778	ID, (1.00)
24	Mississippi		0.918	ID (0.84), NV (0.16)	49	Wisconsin		0.860	ID (0.47), WA (0.53)
25	Missouri		0.789	ID (0.44), UT (0.56)	50	Wyoming		1.000	—

2.4.3 Model Selection and Results Interpretation (Bootstrapping)

The efficiency scores from meta-frontier approach, which is provided in Table 7, represent the non-parametric frontier estimators with biases. Uncontrollable variables may affect the results of the DEA and introduce bias to the efficiency scores. To find bias-corrected efficiency scores, we considered the effects of contextual variables in the model using the bootstrapping approach that Simar and Wilson (27) introduced. The analysis of bootstrapping for 50 states was performed using the “rDEA” package in RStudio and the results are shown in Table A1 of the appendix. In Table A1, we provided the efficiency scores from meta-frontier method along with the Bootstrapping results. The efficiency scores from bootstrapping approach showed that most of those DMUs which were identified as fully efficient in meta-frontier approach had a large bias. Existence of such biases for efficient DMUs might be because of heterogeneity among DMUs and their differences in the socio-demographic conditions such as poverty rates.

2.4.4 Hypotheses Testing

We used bootstrapped DEA analysis to incorporate the effect of environmental variables on the effectiveness of states' healthcare system. These variables were African-American women and poverty rates. We then applied truncated regression using the same rDEA package as we used for bootstrapping to test these hypotheses. We investigated whether the effects of these environmental variables were significant in determining the efficiency scores or not. The analysis was based on the proposed algorithm by (27) which includes two calculation loops. In the first loop, the bias-corrected distance function is calculated with replicates. In the second loop, the bootstrapped estimates for the regression were identified with replication. We ran the analysis at $\alpha = 0.05$ and $\alpha = 0.10$ significance levels. The results of truncated regression and bootstrapping confidence intervals (CI) for the three proposed models are provided in Table 8. The CI for rate of African-American women did not include zero in all three models. This means that the ratio of African-American women is positively associated with the response variable at 0.05 significance level. In these regressions, the response variable was the inverse of bias-corrected efficiency scores. Therefore, a smaller independent variable means a better efficiency score. Hence, the positive estimated coefficient for African-American rate means that the efficiency score is lower for those DMUs with higher rate of African-American women. However, the effect size of this contextual variable is larger in models with infant survival rate in their output set (Model 1 and Model 3).

The estimated parameter for poverty rate in Model 1 included zero, which means the effect of this variable might not be significant on the efficiency scores at 5% significance level. However, the poverty rate was significantly associated with the response variable in Model 2 and Model 3 (p -value < 0.05). Also, the coefficient of this variable was positive in all three models. Considering the dependent variable, which is the inverse of bias-corrected efficiency scores, DMUs with greater poverty rate received smaller efficiency score.

Table 8. Bootstrapped regression results on the environmental factor

Environmental Variables	Model 1			Model 2			Model 3		
	Coef.	90% Confidence Interval		Coef.	90% Confidence Interval		Coef.	90% Confidence Interval	
Intercept	1.293 **	1.148	1.475	0.996 **	0.995	0.999	0.999 **	0.998	1.001
African-American ratio	0.004 **	0.0001	0.007	0.002 **	0.001	0.003	0.003 **	0.001	0.005
Poverty	0.005 *	-0.020	0.006	0.003 **	0.002	0.004	0.002 **	0.001	0.003

** p -value < 0.05 ; * p -value < 0.10

2.4.5 Slack Analysis and Policy Insights

Data envelopment analysis can distinguish efficient and inefficient DMUs. More importantly, DEA is capable of identifying the best practices for inefficient units and provide insights on where they can focus to improve their performance. Also, it is crucial to provide a fair comparison by finding the best role models for each state given the heterogeneity among states. Table 7 provides the list of best practices for each inefficient state. Therefore, if state policymakers want to make changes and make the healthcare system more efficient, they need to look at their peers and learn from their behaviors. In addition, slack analysis will be helpful in exploring improvement options for each state's healthcare system which can help policymakers find where they can save or spend more to improve the performance.

As mentioned in the method section, the DEA is an operations research method which solves a linear programming (LP) model for each DMU. In the formulation of each LP model with an objective function and a set of constraints, there is a slack associated with each constraint. These slacks refer to input excesses or output shortfalls. Input slack can be identified by the difference between the current/actual input and the target input. Using the input slacks, one can identify potential improvement opportunities in the input side to improve the efficiency of the inefficient DMUs (35, 36).

We performed a slack analysis and identified target inputs using BCC input-oriented with the VRS model. Considering these input targets and by comparing with the actual input values, we were able to calculate input slacks and slack percentages. The slack analysis for 50 states was performed using the “Benchmarking” package in RStudio and the results are provided in Table A2 of appendix A. The “percentage of slacks” in Table A2 shows the amount of changes that inefficient states need to make in every input variable to improve their efficiency scores. For example, New York received an efficiency score of 0.724 and is identified as an inefficient DMU. According to Table A2, if the policymakers in this state want to improve the efficiency of the healthcare system, they need to increase the number of obstetric and gynecologists and the number of physicians by 17% and 8% respectively. Besides, the number of hospital beds, the amount of health expenditure, and the number of smoker mothers need to be decreased by around 7%, 2%, and 50%, respectively. The same insights can be found for other states from the results of the slack analysis (see Table A2 of appendix A).

2.5 Discussion

In this study, we evaluated the performance of states’ healthcare systems in improving the birth outcomes including infant mortalities, preterm births, and low birthweights. To perform these analyses, we used mixed method approaches that include DEA, ROBPCA, bootstrapping, and statistical methods such as truncated regression. In our analysis, we collected data from different data sources (i.e., CDC, KFF, and AHR) to build our dataset including five inputs, three outputs, and two environmental variables. The inputs were number of obstetric and gynecologists, physicians, patients’ beds, smoker pregnant women per 1,000 population, and amount of healthcare expenditure per capita. All the output variables were categorized as undesirable; therefore, we converted them to desirable variables (i.e. infant survival, full-term birth, and normal birthweight rate). We applied an input-oriented DEA model with variable return-to-scale including 50 DMUs. Further, we used bootstrapping DEA to incorporate the effect of environmental variables (i.e., number of African-American women and poverty rate) on the efficiency scores.

In the first part of our analysis, we ran three DEA models. In Model 1, we estimated the overall efficiency scores, including all three output variables. In Model 2, we included only preterm birth rates and low birthweight rates in the output side. Although these two unfavorable outcomes were associated with a high cost burden on the healthcare system, they can be reversible. In Model 3, we only included infant mortality because it is one of the most important indicators of overall health in a society and it is also an irreversible event. Our findings revealed that the performance of the

healthcare systems are similar in terms of improving only reversible or irreversible birth outcomes because their efficiency scores were in the same range. Another major finding was that less than one fourth of healthcare systems are efficient if all the birth outcomes are considered at the same time. However, if we only include irreversible or reversible birth outcomes, even fewer healthcare systems or DMUs are fully efficient. The lowest number of efficient DMUs belonged to Model 3. This means that most of the states need to focus on reducing infant mortality rates. Most of the states (i.e., 80%) also need to improve their performance and reduce their preterm birth and low birthweight rates.

The results of our analysis indicated that the major variations in healthcare systems' performance in terms of their birth outcomes can be significantly explained by contextual factors. Socio-economic and demographic characteristics of states affect the outcome and performance of their healthcare systems. Truncated regression results showed that the ratio of African-American population is significantly associated with the estimated efficiency scores. This factor is positively correlated with the efficiency scores at 5% significance level in all of the models. Our findings are consistent with the results of other studies in the literature that studied the risk factors of unfavorable birth outcomes (14, 15).

The second contextual variable was ratio of persons in poverty in each state. This factor was also significantly associated with estimated efficiency scores. States with greater poverty rate have lower efficiency scores in all of the models. Poverty rate had larger effect size in Model 2, which accounts for adverse birth outcomes that are reversible. This result is in line with the findings of other studies in the recent years (16-19). However, its effect size is smaller in Model 3 and it is not significant (at 0.05 level) in Model 1. This result pointed to the fact that living in poverty and disadvantaged areas is associated with giving premature birth because the underprivileged might have limited access to adequate care or having difficulty seeking care.

In the last part of the results, we performed the slack analysis to explore improvement opportunities for inefficient states. Slack analysis is an important domain in the operations research domain. This method highlights the crucial role of industrial and systems engineering and operations research in healthcare-related decisions. Industrial engineers can advance the understanding of policymakers on the performance of a healthcare system and can offer insights for improvement.

2.6 Implication

To our knowledge, this is the first study that assessed the performance of healthcare systems at the state level. It is also the first study that applied DEA to quantify the effectiveness of the state's healthcare systems in reducing adverse birth outcomes. We provided a systematic method for states' policy makers to choose benchmarks for improving the performance of their healthcare system. In our analysis, other than infant mortality rate, which is an irreversible and undesirable event, we considered preterm birth and low birthweight rates, which are considered as reversible events. In a comparison of the performance of states' healthcare systems, it is important to consider the effect of factors that are not part of the transformation process. Therefore, to have a comprehensive evaluation and provide a fair comparison, we needed to include potential demographical and socio-economic factors in the analysis. To the best of our knowledge, no DEA study in the literature considered the impact of contextual variables on the performance of states' healthcare systems. Therefore, another novelty of this work was incorporating the effect of two environmental variables in the model. We realized that disregarding the contextual factors leads to an overestimation of the system's efficiency.

The result of our work provides policymakers new insights on how states' healthcare systems are performing and how efficient they are in achieving desirable birth outcomes. Also, our results bring policymakers and health providers' attention to the importance of considering contextual factors in their interventions. Furthermore, the results of our work can be informative for practitioners at the health systems in identification of those who are at higher risk of giving unfavorable birth outcome and need special clinical attention. This study can be considered as a stepping stone to apply mathematical modeling in the context of performance measurement of healthcare systems at state level.

2.7 Future Work

There are several avenues for future research and expanding the results of this work. First, in this study, we applied bootstrapping to incorporate the effect of environmental variables in the model. However, DEA can be combined with machine learning (ML) techniques. Therefore, one can use ML models to study the impact of environmental variables. Second, the results of outlier identification with and without including contextual variables in the ROBPCA analysis revealed that the distribution of DMUs varied substantially. Results of ROBPCA, considering only the environmental variables, points to another potential method of comparison that can be a cluster-

frontier approach. To have homogenous groups for comparison, we can use a cluster-frontier approach to cluster similar DMUs and find the frontier for each group. The cluster frontiers can then be compared with the overall frontier from the meta-frontier approach. Third, this work was the first study that incorporated the effect of contextual variables such as poverty rate and the number of African-American women in the context of performance measurement of the healthcare systems. However, there were challenges to include more environmental variables considering the level of analysis and the availability of data. Therefore, future research can be done by including other potential environmental factors.

References

1. Sawyer B, McDermott D. How does infant mortality in the US compare to other countries? Peterson-Kaiser Health System Tracker. 2019.
2. Centers for Disease Control and Prevention C. Infant Mortality 2019 [Available from: <https://www.cdc.gov/reproductivehealth/maternalinfanthealth/infantmortality.htm>].
3. Ebrahimvandi A, Hosseinichimeh N, Iams J. Understanding State-Level Variations in the US Infant Mortality: 2000 to 2015. *American journal of perinatology*. 2018.
4. Kleinman JC. State trends in infant mortality, 1968-83. *American journal of public health*. 1986;76(6):681-8.
5. Vanlandingham MJ, Buehler JW, Hogue C, Strauss LT. Birthweight-specific infant mortality for native Americans compared with whites, six states, 1980. *American journal of public health*. 1988;78(5):499-503.
6. He X, Akil L, Aker W, Hwang H-M, Ahmad H. Trends in infant mortality in United States: A brief study of the southeastern states from 2005–2009. *International journal of environmental research and public health*. 2015;12(5):4908-20.
7. Muoto I, Luck J, Yoon J, Bernell S, Snowden JM. Oregon's coordinated care organizations increased timely prenatal care initiation and decreased disparities. *Health Affairs*. 2016;35(9):1625-32.
8. Ibrahim MD, Daneshvar S. Efficiency analysis of healthcare system in Lebanon using modified data envelopment analysis. *Journal of healthcare engineering*. 2018;2018.
9. Benneyan J, Ceyhan ME, Sunnetci A, editors. Data envelopment analysis of national healthcare systems and their relative efficiencies. the 37th international conference on computers and industrial engineering; 2007.
10. Sikka V, Luke RD, Ozcan YA. The efficiency of hospital-based clusters: Evaluating system performance using data envelopment analysis. *Health Care Management Review*. 2009;34(3):251-61.
11. Braithwaite J, Westbrook MT, Hindle D, Iedema RA, Black DA. Does restructuring hospitals result in greater efficiency-an empirical test using diachronic data. *Health Services Management Research*. 2006;19(1):1-12.
12. Mirmirani S, Li H, Ilacqua JA. Health care efficiency in transition economies: an application of data envelopment analysis. *International Business & Economics Research Journal (IBER)*. 2008;7(2).
13. Cetin VR, Bahce S. Measuring the efficiency of health systems of OECD countries by data envelopment analysis. *Applied Economics*. 2016;48(37):3497-507.
14. MacDorman MF, editor Race and ethnic disparities in fetal mortality, preterm birth, and infant mortality in the United States: an overview. *Seminars in perinatology*; 2011: Elsevier.
15. Collins Jr JW, David RJ, Handler A, Wall S, Andes S. Very low birthweight in African American infants: the role of maternal exposure to interpersonal racial discrimination. *American journal of public health*. 2004;94(12):2132-8.
16. Cubbin C, Kim Y, Vohra-Gupta S, Margerison C. Longitudinal measures of neighborhood poverty and income inequality are associated with adverse birth outcomes in Texas. *Social Science & Medicine*. 2020;245:112665.
17. Taylor-Robinson D, Lai ET, Wickham S, Rose T, Norman P, Bambra C, et al. Assessing the impact of rising child poverty on the unprecedented rise in infant mortality in England, 2000–2017: time trend analysis. *BMJ open*. 2019;9(10):e029424.
18. Romero-Sandoval N, Alcázar DD, Pastor J, Martín M. Ecuadorian infant mortality linked to socioeconomic factors during the last 30 years. *Revista Brasileira de Saúde Materno Infantil*. 2019;19(2):295-301.
19. Mohamoud YA, Kirby RS, Ehrenthal DB. Poverty, urban-rural classification and term infant mortality: a population-based multilevel analysis. *BMC pregnancy and childbirth*. 2019;19(1):40.
20. Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. *European journal of operational research*. 1978;2(6):429-44.

21. Farrell MJ. The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General)*. 1957;120(3):253-81.
22. Banker RD, Charnes A, Cooper WW. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*. 1984;30(9):1078-92.
23. O'Donnell CJ, Rao DP, Battese GE. Metafrontier frameworks for the study of firm-level efficiencies and technology ratios. *Empirical economics*. 2008;34(2):231-55.
24. Topcu TG, Triantis K, Roets B. Estimation of the workload boundary in socio-technical infrastructure management systems: The case of Belgian railroads. *European Journal of Operational Research*. 2019;278(1):314-29.
25. Herrera-Restrepo O, Triantis K, Seaver WL, Paradi JC, Zhu H. Bank branch operational performance: A robust multivariate and clustering approach. *Expert Systems with Applications*. 2016;50:107-19.
26. Triantis K, Sarayia D, Seaver B. Using multivariate methods to incorporate environmental variables for local and global efficiency performance analysis. *INFOR: Information Systems and Operational Research*. 2010;48(1):39-52.
27. Simar L, Wilson PW. Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics*. 2007;136(1):31-64.
28. Ramanathan R. Evaluating the comparative performance of countries of the Middle East and North Africa: A DEA application. *Socio-Economic Planning Sciences*. 2006;40(2):156-67.
29. Asandului L, Roman M, Fatulescu P. The efficiency of healthcare systems in Europe: A data envelopment analysis approach. *Procedia Economics and Finance*. 2014;10:261-8.
30. Hadad S, Hadad Y, Simon-Tuval T. Determinants of healthcare system's efficiency in OECD countries. *The European journal of health economics*. 2013;14(2):253-65.
31. Scheel H. Undesirable outputs in efficiency valuations. *European journal of operational research*. 2001;132(2):400-10.
32. Seiford LM, Zhu J. Modeling undesirable factors in efficiency evaluation. *European journal of operational research*. 2002;142(1):16-20.
33. Centers for Disease Control and Prevention (CDC). Linked Birth and Infant Death Data 2018 [Available from: <https://www.cdc.gov/nchs/nvss/linked-birth.htm>].
34. Health Resources and Services Administration. Area Health Resources Files 2017 [Available from: <https://data.hrsa.gov/topics/health-workforce/ahrf>].
35. Ramanathan R. An introduction to data envelopment analysis: a tool for performance measurement: Sage; 2003.
36. Cooper WW, Seiford LM, Tone K. Data envelopment analysis. *Handbook on data envelopment analysis*. 2000:1-40.

Chapter 3. Associations between State Policies on Opioid Prescriptions and Opioid Overdose Death Rates in the United States, 2008–2017

Abstract

Opioid overdose mortalities have increased dramatically in the U.S. over the past 16 years. In this study, we evaluated the effect of four state policies on opioid prescriptions and overdose deaths controlling for insurance type, income, tobacco use, and binge drinking using a state-level panel data from 2008 to 2017. Although we found significant negative associations between opioid prescription rates and mandated prescription drug monitoring programs (PDMPs), limitations on dose/duration of the initial prescription and pain clinic laws, they were not negatively associated with overdose mortalities; however, once we considered a delay for mandated PDMPs, a significant association appeared. We also found that a higher share of uninsured people and binge drinkers are associated with higher overdose deaths. State laws are unlikely to result in an immediate reduction in overdose deaths and comprehensive interventions are needed to curb the epidemic.

3.1 Background

Opioid overdose has become a major issue in the U.S., and in 2017 (1), a presidential commission declared it a national emergency. Although the number of deaths from opioid overdose in the country has increased since 1999, the rate of growth in opioid overdose has surged in recent years. Between 2013 and 2017, the number of deaths from synthetic opioids—with the exception of methadone—increased by 71% compared to an 8% growth between 1999 and 2010 (2). Every day in the U.S., 130 people die from opioid overdoses (3).

In response, many states have enacted laws and initiated programs (e.g., Prescription Drug Monitoring Programs (PDMPs) and medical cannabis laws) to reduce the opioid epidemic and mortality rate. The PDMPs is an electronic database used in some states to track the prescribing and dispensing of opioid pain relievers. Although this program became operational in 1939 in California, and other states subsequently implemented it, prescribers and dispensers were not required to use it (4). Mandated use of PDMPs for prescribers and dispensers began in 2006; by

2017, 28 states had mandated this program. Medical cannabis laws, which allow the medical use of marijuana for adults, were passed for the first time in California in 1996. By 2017, medical cannabis laws, introduced to combat drug/opioid misuse, passed in 47 states (5).

Despite the efforts and resources allocated for such programs, opioid overdose deaths are still increasing in the U.S. Past studies have investigated determinants of opioid overdose deaths using individual-level data and the effect of PDMPs, medical cannabis laws, and pain clinic laws on opioid mortalities at the state level. Literature on the potential predictors of opioid dependency and overdose deaths at the individual level has explored the association among these outcome variables and several predictors such as demographics, employment status, income level, tobacco use, insurance type, history of alcohol abuse, and dose and duration of prescribed opioids (6-12). The studies have showed that age, gender, prescribed opioid dosage, and tobacco and alcohol use are significantly associated with opioid overdose deaths.

State-level studies have investigated the impact of policies such as PDMPs, cannabis laws, and pain management clinic laws on the mortalities from opioid overdose (13-15). These studies have found conflicting results regarding the effect of mandated PDMPs on opioid overdose deaths. One study (14) found that mandating PDMPs along with pain clinic law implementations could reduce opioid overdose deaths by 12%, while others did not find a significant impact (13, 15). In addition, past state-level studies of opioid overdose deaths have included few or no explanatory factors, although individual-level studies indicate that factors such as type of health insurance, binge drinking, and tobacco use are significantly associated. Moreover, some of these studies did not consider the lag between policy implementations and change in outcomes. It is quite possible that a delay exists in the effect of opioid policies on overdose deaths. In addition, as most opioid-related policies were implemented after 2013—and the timelines of prior studies range from 1999 to 2013—investigation into long-term impacts of the policies have been limited. As such, considering the sharp growth in deaths after 2013, it is critical to analyze recent data (2).

In this study, we utilize state variations in overdose death rates to examine factors affecting mortality from opioid overdoses over a ten-year time horizon and accounting for the aforementioned four gaps. First, unlike previous studies of state policies' impact on opioid overdose deaths, we control for multiple factors including median household income (logged), insurance coverage type, binge drinking, and tobacco use. Second, we investigate the impact of multiple states' interventions consisting of dose and duration limits on the initial prescription, pain management clinic laws, mandated use of PDMPs, and medical cannabis laws. Third, we include

lags of these policies/laws in regression models and explore whether these policies impact overdose death rates with a delay. Finally, our panel data includes recent years in which overdose deaths rate have increased rapidly, which enables investigation of the association between overdose death data and time-lagged interventions.

3.2 Study Data and Method

3.2.1 Data

We obtained our data from different sources. The dependent variable in our study is opioid overdose deaths per 100,000 people (age-adjusted) acquired from the Centers for Disease Control and Prevention (CDC) (16). According to the “International Classification of Diseases, Tenth Revision (ICD-10)”, opioid overdose deaths are defined under several codes based on the categories of opioid analgesic.

Some of the independent variables, including different types of health insurance coverage and median household income, were acquired from the U.S. Census Bureau (17, 18). The data for binge drinking and tobacco use were collected from the CDC (19, 20). Data on dosage and duration of opioid prescription policies and whether states’ mandate prescribers and dispensers to use PDMPs were gathered from AffirmHealth (21) and the “Prescription Drug Monitoring Program Training and Technical Assistance Center” (4), respectively. For pain management clinic laws, we used data from the “Prescription Drug Abuse Policy System (PDAPS)” database (22). Finally, medical cannabis laws were obtained from the “National Conference of State Legislatures (NCSL)” (5).

3.2.2 Key Variables

In this study, we created a panel data set for a ten-year time frame—2008 to 2017. We included the independent variables shown to be important in prior related studies. These state-level variables are median household income, insurance coverage rates (percentage of state residents who are either uninsured, using private insurance, Medicare, or Medicaid), binge drinking (percentage of adults who have at least four or more alcoholic drinks on one occasion), tobacco use (percentage of adults who smoke every day), and some state policies. We included four types of government policies and laws. First, we added dose/duration policies and guidelines and recommendations for prescribing opioids and limits on the dosage/duration (per day) of opioid

prescriptions. If a state had such limitations on an initial prescription in a specific year, we coded it as 1 and zero otherwise. Second, we included pain management clinic laws that prevent inappropriate large-quantity prescriptions by requiring state supervision over such clinics and requiring them to follow certain prescribing guidelines (23). If a state has regulated this law, we coded it as 1 and zero otherwise. Third, we added Prescription Drug Monitoring Programs (PDMPs), an electronic database that some states use to track the prescribing and dispensing of opioid pain relievers. If a state has mandated the use of a PDMPs for prescribers and/or dispensers for a specific year, we coded that as 1 and zero otherwise. Fourth, we included medical cannabis laws, which was coded 1 if a state has approved the medical use of marijuana and zero otherwise.

Before performing regression on our data set, we conducted a pairwise simple linear regression between all variables (see Figure B 1). First, this can help us to explore correlation between variables and check for multicollinearity effect. Second, we can see how spread the variables are to fix any clumpy-ness in them. We noticed that the correlation between median household income and tobacco usage is relatively high, therefore, we need to be aware of this multicollinearity in the interpretation of the results. Also, we can see that the opioid overdose death rates and median household income are also rather clumpy. This suggests that logarithm transformation might be useful on these variables. The result of pairwise linear regression after applying log transformations on the opioid overdose death rates and median household income is shown in Figure B 2. Based on this figure, after applying logarithm transformations, all the variables are randomly distributed and there is no sign of clumpy-ness.

3.2.3 Analyses

We analyzed our data set using state and year fixed-effect regressions to investigate the association between opioid overdose death rates and four state policies, controlling for multiple independent variables for 50 states plus the District of Columbia. We run two sets of regressions with two different dependent variables: an opioid prescribing rate per 100 population and an opioid overdose deaths per 100,000 people (log-transformed). Independent variables include insurance coverage rates, logarithm transformation of median household income, binge drinking rate, tobacco use, and four state-level laws and programs (i.e., mandated PDMPs, medical cannabis laws, pain clinic laws, and first prescription's dose/duration limits). To investigate the role of time delay in the association between opioid overdose deaths and government policies, we include the lags of each law/program as independent variables in separate regressions. We

operationalized policy lags by shifting the year of implementation. For instance, Arizona passed legislation for medical marijuana use in 2010; thus, to create a one-year lag of the medical cannabis laws we changed the binary variable in 2010 to zero and made the variable value into one in 2011. We experimented with lags from one to nine years for all state policies and programs. All statistical models were conducted in Stata 15.

3.2.4 Limitations

Our analysis has several limitations. First, the data for opioid-related overdose death rates provided by the “CDC’s National Center for Health Statistics” captures all overdose mortalities in which an opioid analgesic was involved. This includes those mortalities involving illicit drug use such as heroin. In addition, as overdose mortalities from legally prescribed fentanyl could not be distinguished from illegally produced fentanyl, mortality rates in this data set include both. Our dependent variable includes all opioid deaths, and thus our results might underestimate the impact of state laws on the death rates. However, previous prescribed users might switch to illegal heroin and an overdose can be still indirectly related to such laws. Second, since our study is based on state-level data and the aggregation level in the analysis might impact the effect size of the coefficients, it is important to interpret the results only on the state level. Third, the dose/duration limitation policies of initial prescriptions vary by state and type of chronic pain. Therefore, we were unable to include the exact threshold for prescription limitations and instead used a dummy variable showing whether a state had such policies in place. Fourth, as long as a law/policy was implemented at any point during a year, we coded it as one. A more precise way of formulation is to consider a fraction for the first year of implementation. However, because the time horizon of our study is ten years and we did not see any immediate effect of such policies, the limitation is less likely to affect the results. Although the fixed-effect method controls for any omitted variables that are unchanging over time—but do differ across states or vice versa—there could be some omitted variables that change across states and time.

3.3 Study Results

3.3.1 Variation in Opioid Overdose Mortality

Age-adjusted opioid overdose death rates varied substantially across states. Figure 4 shows the change in opioid overdose mortality rates between 2008 and 2017 for all states including the District of Columbia. Although the majority of states experienced an increase in the rate of opioid

overdose deaths, there were some states that experienced a reduction. For instance, deaths in Ohio rose to 32 people per 100,000, while Montana had a decline of 4 people per 100,000.

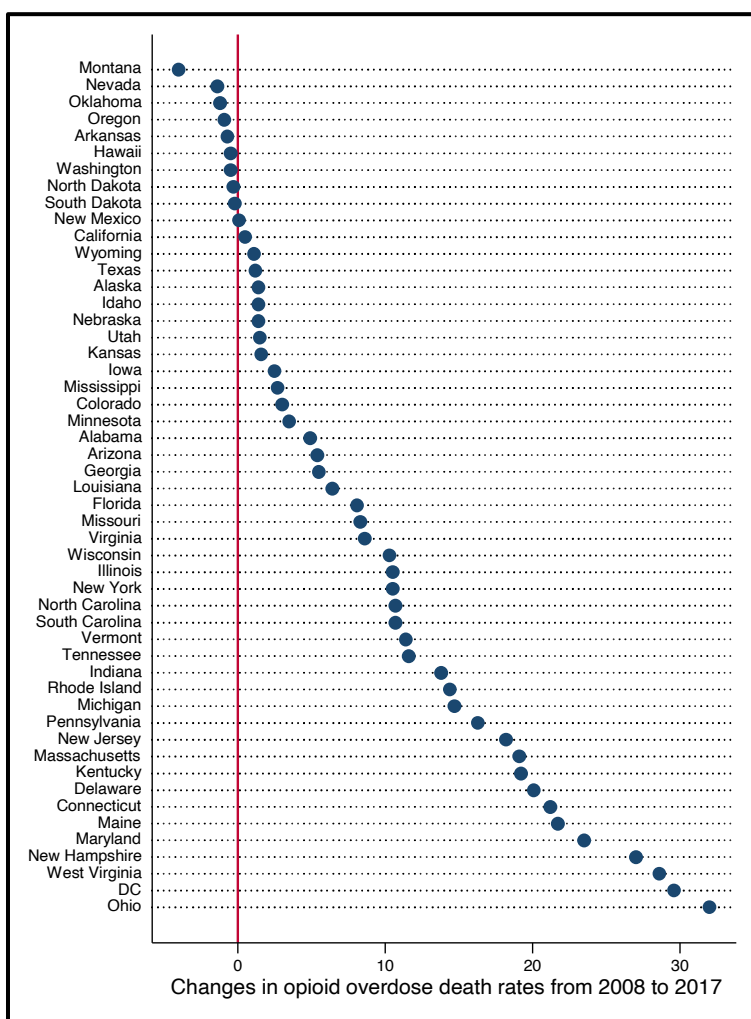


Figure 4. Changes in opioid mortality rates in all 50 states and DC from 2008 to 2017.

3.3.2 Opioid Prescribing Trends

The opioid prescription rate (i.e., number of prescriptions dispensed per 100 population) in the U.S. increased by 3.96% from 2008 to 2012 and then decreased after 2012. The opioid prescription rate was 58.7% in 2017, which is 27.80% lower than the rate in 2012 (see Figure 5). Concurrently, an increasing number of states adopted the four laws (i.e., dose/duration limitations, mandated PDMPs, pain management clinic laws, and the medical cannabis laws), which could explain the opioid prescription rate reduction after 2012 (Figure 5).

3.3.3 Opioid Overdose Death Rate Trends

Figure 5 shows prescribing rates and age-adjusted opioid overdose mortalities per 100,000 in the U.S., as well as the number of states that adopted dose/duration limitations, mandated PDMPs, pain clinic laws, and medical cannabis laws from 2008 to 2017. Although the prescribing rates in the U.S. declined after 2012, the mortality rates from opioid overdoses have increased substantially since 2013 (Figure 5). However, it should be noted that the increase in mortality slowed after 2016. Although the mortality rate increased by 27.88% between 2015 and 2016, it rose only by 12.03% from 2016 to 2017, or 56.86% less increase in opioid overdose mortalities per 100,000.

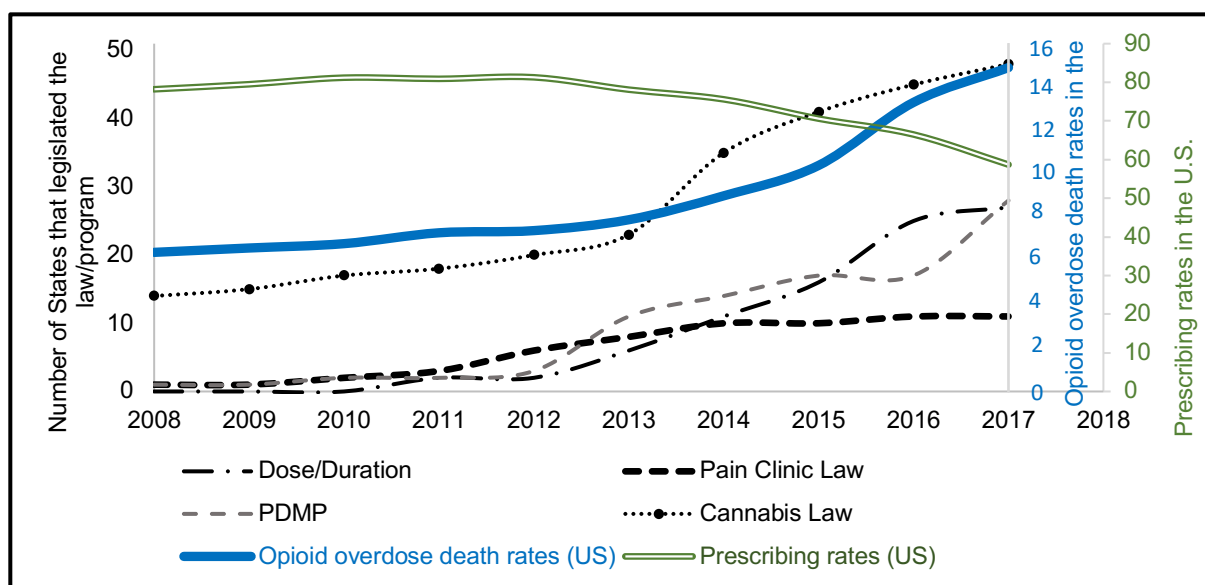


Figure 5. The number of states with different laws and programs to combat opioid epidemic along with the mean rates of opioid overdose mortality and opioid prescriptions in the U.S. from 2008 to 2017.

3.3.4 Descriptive Statistics

Table 9 represents the summary statistics for all variables from 2008 to 2017. In this table, we report the average, standard deviation, and range across the entire time frame. For instance, the average mortality rate from opioid overdoses is 10.15, which ranges between 0.70 and 49.80 people per 100,000. On average, 11.73% of states' residents were uninsured, 68.65% had private insurance, 17.71% were on Medicaid, and 16.11% were on Medicare. Since some people were enrolled in multiple types of health insurance programs, the percentage of people with different insurance coverage type does not add up to 100%.

In columns 4–7, we provided the average of each variable for every three years (i.e., 2008, 2011, 2014, and 2017) across all states to capture the variable variation over time. The death rate from opioid overdoses increased from 7.35 to 16.26 per 100,000 over 10 years during which time more residents were on Medicaid (from 14.84% to 19.96%) and Medicare (from 14.60% to 17.78%). The percentage of uninsured people increased from 2008 to 2011, but decreased from 2011 to 2017, while conversely, the share of the population with private insurance declined and then increased. Binge drinking did not exhibit any particular trend. The percentage of states that adopted policies related to opioid increased across all four categories.

Table 9. Summary data on opioid overdose death rates and state-level variables from 2008 to 2017.

Variable	Mean (SD)	Range	2008	2011	2014	2017
Opioid overdose mortalities per 100,000 people (age-adjusted)	10.15 (6.80)	0.70-49.80	7.35	8.34	10.44	16.26
Opioid prescribing rate (per 100 persons)	79.60 (23.40)	28.50-146.90	78.2	80.9	75.6	58.7
Insurance coverage						
Uninsured	11.73 (4.50)	2.50-23.80	13.40	13.79	10.68	8.07
Private	68.65 (5.90)	53.20-81.90	71.35	67.32	68.22	69.15
Medicaid	17.71 (4.40)	9.00-33.20	14.84	17.10	18.63	19.96
Medicare	16.11 (2.42)	8.70-23	14.60	15.42	16.67	17.78
Median household income	54048.78 (9281.10)	32338-83382	50747	50677	55160	61640
Binge drinking	16.47 (3.30)	6.60-27.20	15.32	18.34	16.40	17.13
Tobacco use (Smoke everyday)	13.37 (3.21)	5.60-24	13.95	15.49	13.04	12
Policy/Law						
Dosage/Duration	0.17 (0.38)	0-1	0	0.04	0.22	0.53
Pain clinic laws	0.12 (0.33)	0-1	0.02	0.06	0.20	0.22
Mandated PDMPs	0.20 (0.40)	0-1	0.02	0.04	0.29	0.55
Medical cannabis laws	0.54 (0.50)	0-1	0.27	0.35	0.69	0.94

3.3.5 Impact of State Policies and Other Factors on Opioid Prescribing Rates

Regression models of opioid prescribing rates are presented in Table 10. In all the regression models, the logarithm transformation of prescribing rates and three of the laws (i.e., dose/duration policies, mandated PDMPs and pain management clinic laws) are negatively and significantly associated with the prescribing rate. The implementation of pain clinic laws is associated with a 6% ($e^{-0.06} = 0.94 = 1 - 0.06$) reduction in the number of opioid prescriptions per 100 population.

Mandated PDMPs and dose/duration regulations are correlated with a 3% ($e^{-0.03} = 0.97 = 1 - 0.03$) and 2% ($e^{-0.02} = 0.98 = 1 - 0.02$) reduction in prescribing rate, respectively. However, medical cannabis laws are not significantly associated with the prescribing rates. These results hold after the addition of more explanatory variables into the model.

Table 10. Fixed-effect regression models of opioid prescription rates, for 50 states and DC from 2008 to 2017.

Variable	Model 1 (n=510)	Model 2 (n=510)	Model 3 (n=510)	Model 4 (n=510)	Model 5 (n=510)
Uninsured					-0.001 (-0.02, 0.01)
Medicaid					-0.005 (-0.02, 0.01)
Medicare					-0.001 (-0.04, 0.03)
Median income (logged)		0.04 (-0.08, 0.17)	0.04 (-0.11, 0.16)	0.03 (-0.11, 0.16)	0.03 (-0.11, 0.15)
Binge drinking			0.01 (-0.01, 0.01)	0.01 (-0.01, 0.01)	0.01 (-0.01, 0.01)
Tobacco use				0.01 (-0.01, 0.01)	0.01 (-0.01, 0.01)
Dose/Duration	-0.02* (-0.04, 0.01)	-0.02* (-0.04, 0.01)	-0.02* (-0.04, 0.01)	-0.02* (-0.04, 0.01)	-0.02* (-0.04, 0.01)
Pain management clinic laws	-0.06** (-0.11, -0.01)	-0.05** (-0.10, -0.01)	-0.06** (-0.10, -0.01)	-0.05** (-0.10, -0.01)	-0.05** (-0.10, -0.01)
Mandated PDMPs	-0.03* (-0.06, 0.01)	-0.03* (-0.05, 0.01)	-0.02* (-0.05, 0.01)	-0.03* (-0.05, 0.01)	-0.02* (-0.05, 0.01)
Medical cannabis laws	0.01 (-0.02, 0.03)	0.01 (-0.08, 0.17)	0.01 (-0.02, 0.03)	0.01 (-0.02, 0.03)	0.001 (-0.02, 0.03)
Overall	0.06	0.04	0.03	0.09	0.06
Within	0.829	0.829	0.830	0.831	0.834

*p-value <0.1, **p-value <0.05, ***p-value <0.01

3.3.6 Impact of Policies and Other Factors on Opioid Overdose Death Rates

Table 11 presents the regression results for overdose mortality rates. The sixth regression includes only the states' policies and laws. Based on this model, a positive association exists between opioid-overdose death rates and dose/duration at a 0.05 significance level and mandated PDMPs at a 0.01 significance level. In models 7, 8, and 9, more variables were added to the regression (i.e., logarithm of the median household income, binge drinking, and tobacco use). The coefficient size and level of significance remained the same across these regressions.

In the tenth regression, we included types of health insurance coverage. Private insurance, which has the highest prevalence, is the reference category. In this model, the mandated use of PDMPs is still positive and significant, but dose/duration variable is only significant at the 0.1 level. Additionally, the rate of persons who are either uninsured or on Medicaid compared to those with private insurance was significantly associated with opioid overdose death rates. In regression 10, a 1%-point increase in the mean value of the uninsured rate—from 11.73 to 12.73%—compared to private insurance, was associated with 13% () more opioid-overdose mortality rates (95%CI, 0.06 to 0.18) on average. Since the average opioid overdose mortality rate was 10.15 per 100,000 people, a 13% increase in mortality rate results in 1.32 more deaths (i.e., $10.15 \times 0.13 = 1.32$). Also, a 1%-point increase in the mean value of the percentage of state residents with Medicaid—from 17.71% to 18.71%—compared to private insurance, was associated with 4% (4) or, on average, 0.40 (i.e., 0.04×10.15) more opioid overdose deaths per 100,000 people. The magnitude effect of Medicaid is one third of the effect size of the uninsured rate. Moreover, in the tenth model, binge drinking is positively associated with opioid death rate ($p\text{-value} < 0.1$), but its effect size is relatively small.

Opioid overdose death rates and two opioid-related policies (dose/duration and mandated PDMPs) are positively associated in regressions 6–10, which could imply that states with increasing opioid overdoses should adopt such policies. In the eleventh regression, we examined whether these policies affect opioid overdose with a time delay. The lagged variables of state policies were added to the regression. Lagged dose/duration, pain management clinic, and cannabis laws were not significantly associated with opioid overdose death rates. However, a mandated PDMP with a five-year lag is *negatively correlated* with mortality rates from opioid overdose at the 0.1 level. We included different times lags of policy variables in the regression, and mandated PDMPs with a lag of less than five years were not significant. Based on the eleventh regression, after five years, the mandated use of PDMPs is associated with 13% ($e^{-0.14} = 0.87 = 1 - 0.13$) less annual mortality on average from opioid overdose compared to states without these programs. This is equivalent to 1.32 (10.15×0.13) fewer deaths per 100,000 people.

Adding lags in regression 11 only changes the sign of Medicaid variables, and the lags do not change the sign and level of significance for binge drinking and uninsured variables. A higher share of individuals with Medicaid compared to private insurance is associated with 5% higher mortality rates. A higher percentage of state residents who binge drink and higher rates of uninsured persons compared to those who have private insurance ($p\text{-value} < 0.01$) are positively

associated with 4.1% and 11.6% more opioid overdose mortalities, respectively. Based on Table 11, tobacco use and median household income are not significant in any of the regressions.

Table 11. Fixed-effect regression models of opioid overdose mortality rates, for 50 states and DC from 2008 to 2017.

Variable	Model 6 (n=510)	Model 7 (n=510)	Model 8 (n=510)	Model 9 (n=510)	Model 10 (n=510)	Model 11 (n=510)
Uninsured					0.12 *** (0.06, 0.18)	0.11*** (-0.10, 0.14)
Medicaid					0.04 * (0.003, 0.08)	0.05** (0.0004, 0.08)
Medicare					0.07 (-0.06, 0.2)	0.12 (-0.03, 0.28)
Median household income (logged)		-0.14 (-1.07, 0.80)	-0.21 (-1.07, 0.66)	-0.22 (-1.08, 0.64)	-0.065 (-0.93, 0.80)	-0.29 (-1.22, 0.65)
Binge drinking			0.02 (-0.02, 0.06)	0.02 (-0.02, 0.06)	0.03 * (-0.01, 0.07)	0.04** (0.001, 0.08)
Tobacco use				0.01 (-0.03, 0.06)	0.01 (-0.03, 0.05)	0.006 (-0.04, 0.05)
Dose/Duration	0.14 ** (0.01, 0.26)	0.13 ** (0.01, 0.25)	0.13 ** (0.01, 0.25)	0.13 ** (0.01, 0.23)	0.11* (-0.01, 0.23)	
Pain management clinic laws	0.03 (-0.12, 0.18)	0.02 (-0.14, 0.18)	0.01 (-0.15, 0.17)	0.02 (-0.14, 0.18)	0.07 (-0.07, 0.21)	
Mandated PDMPs	0.23 *** (0.09, 0.37)	0.23 *** (0.09, 0.37)	0.23 *** (0.1, 0.37)	0.23 *** (0.09, 0.37)	0.20 *** (0.07, 0.31)	
Medical cannabis laws	0.11 (-0.04, 0.27)	0.11 (-0.03, 0.27)	0.12 (-0.03, 0.26)	0.11 (-0.04, 0.27)	0.07 (-0.06, 0.21)	
Mandated PDMPs lag 5						-0.14 * (-0.30, 0.02)
Medical cannabis laws lag 5						-0.04 (-0.24, 0.15)
Overall	0.25	0.24	0.20	0.22	0.07	0.04
Within	0.51	0.52	0.52	0.52	0.57	0.54

*p-value <0.1, **p-value <0.05, ***p-value <0.01

3.4 Discussion

We used fixed-effect regression analyses to investigate the associations of four state laws and opioid-overdose death rates across states from 2008 to 2017. Our results show that implementation of three of the state programs and laws (i.e., dose/duration limitations of the first opioid prescription, pain management clinic law, and mandated PDMPs) are negatively associated with opioid prescribing rates. However, none of the state policies had a negative correlation with mortality rates. The significant positive correlation between the mandated use of

PDMPs and setting limitations on dose and duration of the first opioid prescription might imply that states with an increasing rate of opioid mortality rates have adopted such policies. This is consistent with the finding of a study that showed states with higher overdose death rates implemented both pain clinic laws and PDMPs (14).

Similar to our results, past state-level studies have showed that PDMP implementation had an immediate and sustained effect on the decline of prescribing of Schedule II opioid analgesics (24), and it was associated with approximately a 9% reduction in opioid prescription rates for Medicaid enrollees (25). We also examined the impact of delays on the relationship between the laws and overdose mortality rates and found that mandated PDMPs were associated with 13% lower overdose death rates after five years. These results were consistent with an article which showed the mandated use of PDMPs would not lead to a major reduction in opioid mortalities in the near future (26). The authors found that if the annual 7.5% reduction of prescription opioid misuse observed between 2008 and 2015 continues, prescription opioid deaths would decline by 10.7% in 2025.

Unlike Bachhuber and colleagues, we did not find a significant association between cannabis laws and overdose death rates (13). This could be explained by differences in our timeframe (their panel data includes years 1999 to 2010) and extra control variables in our regression. Additionally, we found that setting limitations on the dosage and duration of opioid painkillers for a patient's first prescription had no negative correlation with mortality rates.

Our findings revealed that mortality rates are higher in states with a greater share of uninsured residents compared to those with private insurance. A one percent increase in the uninsured rate was correlated with 1.32 more opioid overdose deaths per 100,000 people. Our result is consistent with findings reported in a paper that shows a 1%-point increase in insurance coverage results in 2.5 fewer opioid mortality per 100,000 among young adults (27). We also found a positive association between opioid mortalities and binge drinking, as has been found in prior studies (28). Although the effect of binge drinking seems negligible, it became significant in the presence of insurance coverage type. This result persisted after including mandated PDMPs and medical cannabis laws with a five-years lag.

Upstream prevention activities such as mandated PDMPs are effective in curbing the opioid crisis over the long term; however, to reduce overdose deaths, states need to focus on comprehensive approaches. Medication treatment is one of the few evidence-based treatments available for opioid use disorder. Yet, by 2016, only 36% of substance use treatment facilities offered any

medication treatment, and only 6% offered all three types of treatment approved by the Food and Drug Administration (i.e., methadone, buprenorphine, and extended-release naltrexone) (29). Future studies can examine the impact of combining upstream and downstream policies on overdose deaths rates.

3.5 Conclusion

Our analysis of state-level data shows that implementation of pain management clinic laws, mandated PDMPs, and limitations on dose and duration of first prescriptions were significantly associated with fewer opioid prescribing rates from 2008 to 2017. However, we did not find a negative association between these laws and the rate of opioid overdose deaths, which may imply that these laws influence deaths rate with a delay. States with a higher share of uninsured and binge drinkers have higher overdose deaths rates. These findings suggest that mandating the use of PDMPs for prescribers and dispensers and reducing the number of uninsured and binge drinkers might lower overdose deaths rate in the long run.

References

1. Christie C, Baker C, Cooper R, Kennedy PJ, Madras B, Bondi P. The president's commission on combating drug addiction and the opioid crisis. Washington, DC: United States Government Printing Office. 2017.
2. Centers for Disease Control and Prevention. Drug Overdose Deaths in the United States, 1999–2017 2018 [Available from: <https://www.cdc.gov/nchs/products/databriefs/db329.htm>].
3. U.S. Department of Health and Human Services. The opioid epidemic in numbers, 2017 [Available from: <https://www.hhs.gov/opioids>].
4. Prescription Drug Monitoring Program Training and Technical Assistance Center. State Profiles 2019 [Available from: <http://www.pdmpassist.org/content/state-profiles>].
5. National Conference State Legislatures. Prescribing Policies: States Confront Opioid Overdose Epidemic 2018 [Available from: <http://www.ncsl.org/research/health/prescribing-policies-states-confront-opioid-overdose-epidemic.aspx>].
6. Edlund MJ, Steffick D, Hudson T, Harris KM, Sullivan M. Risk factors for clinically recognized opioid abuse and dependence among veterans using opioids for chronic non-cancer pain. *Pain*. 2007;129(3):355-62.
7. Glanz JM, Narwaney KJ, Mueller SR, Gardner EM, Calcaterra SL, Xu S, et al. Prediction model for two-year risk of opioid overdose among patients prescribed chronic opioid therapy. *Journal of general internal medicine*. 2018;33(10):1646-53.
8. Turner JA, Saunders K, Shortreed SM, LeResche L, Riddell K, Rapp SE, et al. Chronic opioid therapy urine drug testing in primary care: prevalence and predictors of aberrant results. *Journal of general internal medicine*. 2014;29(12):1663-71.
9. Hasegawa K, Brown DF, Tsugawa Y, Camargo Jr CA, editors. Epidemiology of emergency department visits for opioid overdose: a population-based study. *Mayo Clinic Proceedings*; 2014: Elsevier.
10. Ives TJ, Chelminski PR, Hammett-Stabler CA, Malone RM, Perhac JS, Potisek NM, et al. Predictors of opioid misuse in patients with chronic pain: a prospective cohort study. *BMC health services research*. 2006;6(1):46.
11. Back SE, Payne RL, Simpson AN, Brady KT. Gender and prescription opioids: Findings from the National Survey on Drug Use and Health. *Addictive behaviors*. 2010;35(11):1001-7.
12. Boscarino JA, Rukstalis M, Hoffman SN, Han JJ, Erlich PM, Gerhard GS, et al. Risk factors for drug dependence among out-patients on opioid therapy in a large US health-care system. *Addiction*. 2010;105(10):1776-82.
13. Bachhuber MA, Saloner B, Cunningham CO, Barry CL. Medical cannabis laws and opioid analgesic overdose mortality in the United States, 1999-2010. *JAMA internal medicine*. 2014;174(10):1668-73.
14. Dowell D, Zhang K, Noonan RK, Hockenberry JM. Mandatory provider review and pain clinic laws reduce the amounts of opioids prescribed and overdose death rates. *Health Affairs*. 2016;35(10):1876-83.
15. Paulozzi LJ, Kilbourne EM, Desai HA. Prescription Drug Monitoring Programs and Death Rates from Drug Overdose. 2011.
16. National Vital Statistics System: Multiple Causes of Death Data [Internet]. CDC WONDER. 2018. Available from: <https://wonder.cdc.gov/mcd.html>.
17. Median Household Income by State [Internet]. 2018. Available from: <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=2ahUKEwjXtPmyrYjAhUoq1kKHf0eBg0QFjAAegQIARAB&url=https%3A%2F%2Fwww2.census.gov%2Fprograms-surveys%2Fcps%2Ftables%2Ftime-series%2Fhistorical-income-households%2Fh08.xls&usq=AOvVaw0MqyTNSJP8VVd2wWKeoHeo>.

18. Health Insurance Historical Tables - HIC Series [Internet]. 2019. Available from: <https://www.census.gov/data/tables/time-series/demo/health-insurance/historical-series/hic.html>.
19. Centers for Disease Control and Prevention. Alcohol Consumption. 2018.
20. Centers for Disease Control and Prevention. Tobacco Use. 2018.
21. AffirmHealth. Opioid Prescribing Guidelines: A State-by-State Overview 2019 [Available from: <https://www.affirmhealth.com/blog/opioid-prescribing-guidelines-a-state-by-state-overview>].
22. Prescription Drug Abuse Policy System. Pain Management Clinic Laws 2018 [Available from: <http://pdaps.org/datasets/pain-management-clinic-laws>].
23. Centers for Disease Control and Prevention. Public Health Law. 2014.
24. Bao Y, Pan Y, Taylor A, Radakrishnan S, Luo F, Pincus HA, et al. Prescription drug monitoring programs are associated with sustained reductions in opioid prescribing by physicians. *Health Affairs*. 2016;35(6):1045-51.
25. Wen H, Hockenberry JM, Jeng PJ, Bao Y. Prescription drug monitoring program mandates: impact on opioid prescribing and related hospital use. *Health Affairs*. 2019;38(9):1550-6.
26. Chen Q, Larochelle MR, Weaver DT, Lietz AP, Mueller PP, Mercaldo S, et al. Prevention of Prescription Opioid Misuse and Projected Overdose Deaths in the United States. *JAMA network open*. 2019;2(2):e187621-e.
27. Wettstein G. Health insurance and opioid deaths: Evidence from the Affordable Care Act young adult provision. *Health economics*. 2019;28(5):666-77.
28. Johnson EM, Lanier WA, Merrill RM, Crook J, Porucznik CA, Rolfs RT, et al. Unintentional prescription opioid-related overdose deaths: description of decedents by next of kin or best contact, Utah, 2008–2009. *Journal of general internal medicine*. 2013;28(4):522-9.
29. Mojtabai R, Mauro C, Wall MM, Barry CL, Olfson M. Medication treatment for opioid use disorders in substance use treatment facilities. *Health Affairs*. 2019;38(1):14-23.

Chapter 4. Machine Learning-Enabled 30-day Readmission Model for Stroke Patients

Abstract

Importance— At a personalized level, identification of patients with higher chances of 30-day readmission and in need of special clinical attention could lower their chances of readmission. While at a system's level, reducing hospital readmissions improves the overall quality of care delivery and reduces the associated cost burden.

Objective—To enhance understanding of the risk factors of 30-day readmission in ischemic stroke patients and identify high risk individuals, we applied four machine learning techniques in addition to logistic regression. We aimed to compare the performance and the predictive power of these methods and identify the best model.

Method—The electronic health records (EHR) of stroke patients were extracted from two stroke centers within the Geisinger Health System between January 1, 2015, and October 7, 2018. A total of 61 variables, including clinical variables, demographical characteristics, discharge status, and type of health insurance were extracted and used in this study. Patients were randomly split for model development (80%) and testing (20%). Imputation was performed separately on the training and testing set. Feature selection and sampling strategies were applied during the model development. Five different models, random forest, gradient boosting machine, extreme gradient boosting (XGBoost), support vector machine, and logistic regression, were developed to predict 30-day readmission after stroke. The models were evaluated based on area under the curve (AUC), sensitivity or recall, specificity, and positive predictive value (PPV).

Result— A total of 3,184 stroke patients (mean age: 71±13.9 years, men: 51.06%) were included in this study. From the 3,184, 301 (9.4%) were readmitted within 30 days. The best performance was obtained when XGBoost was used with ROSE-sampling (max-depth= 4, gamma=0, colsample_bytree=0.8, min_child_weight=10, and subsample=0.5). The AUC for the test set was 0.74 (95% CI: 0.64-0.78) with PPV of 0.43. The top four predictors of the 30-day readmission model were National Institutes of Health Stroke Scale score above 24, insert indwelling urinary catheter, hypercoagulable state, and percutaneous gastrostomy.

Conclusion—Machine learning-based (ML) model can be designed to predict 30-day readmission after stroke using structured data from EHR. Use of ML models improved prediction on 30-day stroke readmissions compared to logistic regression. Among the five algorithms analyzed, the model developed using XGBoost had the best performance.

4.1 Introduction

Readmission rate is one of the commonly used indicators for the quality of care and efficiency of hospitals. Hospital readmissions may reflect issues that were not resolved at the time of discharge or during an inpatient visit. Hospital readmissions impose a substantial financial burden on the healthcare system in the U.S., costing Medicare about \$26 billion annually. (1) “Centers for Medicare and Medicaid Services” defined 30-day readmission as an indicator of poor hospital care and has made reducing readmissions rates a national healthcare reform goal. (2) Consequently, readmission within 30 days after discharge imposes penalties on hospitals in the U.S. (3). On the other hand, stroke is an important cause of long-term disability that can reduce mobility in survivors aged 65 or above. (4) Every year more than 795,000 people experience one type of stroke in the U.S., which results in \$34 billion costs for the healthcare system. (4) Reducing readmissions after stroke, and minimizing its associated cost burden is valuable and could lead to improve quality of care.

Rate of readmission after stroke varies depending on different variables, including patients’ profile, the severity of the stroke, and the quality of care that the patient has received. Studies focusing on finding the potential predictors of readmissions after stroke considered different demographical, clinical, and health-related variables. Past studies found that stroke severity (3, 5), being discharged to skilled nursing, intermediate care facility, or left against doctor’s advice (2, 3, 6, 7), being enrolled in Medicaid or Medicare (5, 7-9), and being married (6) were associated with higher readmissions. Also longer length of hospital stay was associated with lower readmissions among stroke patients. (6) Some of the diagnoses that were found significantly correlated with stroke readmissions are heart failure (2, 7, 9), coronary artery disease (10), and dysphagia (5). Additionally, patients with anemia, dementia, malnutrition, and diabetes had higher chances of 30-day readmission (2, 6, 7, 9). However, past studies provided mixed results regarding the association between readmission rate and age, sex, and hypertension (2, 3, 6, 8, 11, 12).

The literature surrounding potential risk factors of stroke readmissions included baseline clinical variables (demographics, stroke severity, less than 15 comorbidities and clinical interventions) in their models. These studies mainly chose readmission with a long follow-up period (more than 30 days) as a response variable. Moreover, the majority of past studies used multivariate logistic regression which limited the number of included interactions among variables (12, 13) and did not perform well in the presence of collinearity among variables, especially when number of variables included in the models were increased. Machine learning (ML) algorithms that are able to include higher-order interactions among variables and handle multicollinearity, could improve readmission predictions when applied to large and high-dimensional datasets. Despite the success of ML to predict readmissions after heart failure (14, 15), heart attack (16), and other causes of readmissions (17, 18), ML algorithms have not yet been explored for the development of prediction models for the 30-day readmission after ischemic stroke. The objective of this work was to develop prediction models of 30-day readmission in stroke patients. We aimed to identify associated predictors using large and high-dimensional dataset from an integrated healthcare system.

4.2 Methods

4.2.1 Database Description

stroke patients at Geisinger Health System between January 1, 2015, and October 7, 2018. Patients younger than 18 years of age were excluded from this study. As a part of the de-identification process, the age of patients older than 89 years old was masked and changed to 89. Geisinger is a fully integrated health system that serves central, south-central, and northeast Pennsylvania. Geisinger comprises seven acute care hospitals, and an insurance operation, serving a population of 3 million with more than 95% Caucasians and a low (<5%) outmigration rate. This work was exempted by the Geisinger Institutional Review Board for using de-identified information. The data cleaning and preparation were performed in STATA 14.0 and the analyses were performed using R 3.6.0 in R studio. As part of the data pre-processing, units were verified and reconciled if needed.

4.2.2 Study Variables

The outcome measure was readmissions within 30 days after discharge in patients with ischemic stroke. Independent variables included patient age, length of stays (LOS), gender, marital status

(married, single, and previously married), and the “National Institutes of Health Stroke Scale” (NIHSS). A four-level categorical variable for the NIHSS was created (0 to 4, 5 to 11, 12 to 23, and above 24). The types of health insurance at the time of first admission (Medicare, Medicaid, Private, direct employer contact, self-pay, worker compensation, and other government payers) was also included. Other variables in this study were discharge destinations (discharged to the home health organizations; discharged to home, court, or against medical advice; discharged to hospice-home/hospice-medical facility; discharged or transferred to other facilities; discharged or transferred to Skilled Nursing Facility, SNF; discharged or transferred to another rehab facility), and five clinical interventions (received tissue plasminogen activator, TPA; insert indwelling urinary catheter; insert endotracheal tube; percutaneous gastrostomy; and hemodialysis). A total of 47 comorbidities were extracted from the electronic health records of the patient and included in this study. The complete list of variables along with their descriptive statistics and level of missingness was provided in Table 12.

4.2.3 Data Processing, Feature Selection, and Sampling

There were several variables suffering from missing observations. In total we had 3.5% missing observations and except NIHSS, all the other variables had up to 15.4% missingness. Missing data imputation was performed separately so that the test dataset provides an unbiased evaluation of the model fitted on the training data. We performed the imputation using “Multivariate Imputation by Chained Equations” (MICE) package in R (19). Imputed values were computed with 5 imputations and 5 iterations to improve the efficiency of the estimate.

In this study we used two sets of variables, set 1 was the comprehensive set including all the variables and set 2 included only variables selected using a random forest classification algorithm by Boruta package in RStudio (20). By adding randomness, the algorithm created a corresponding shadow attribute for each feature in the data. A random forest classifier was trained on this extended data to compute Z-score for each attribute. All attributes with higher Z-scores than the maximum Z-score among the shadow attributes were considered as important variables.

Further, to avoid poor performance of minority class compare to the dominant class, we balanced the data by applying a class balancing techniques named ROSE on the minority class(21). Because, in case of data imbalanced, many of the classification models are biased towards majority class. Random over sampling examples (ROSE) is a hybrid method which is suitable for

binary imbalanced data set. This method generates an artificial balanced samples using a smoothed bootstrap approach (21).

4.2.4 Model Development

The study sample was randomly split to the train (80%) and the test set (20%). We developed models to predict 30-day readmission of ischemic stroke. We built models using logistic regression (LR) and four ML algorithms to predict the outcomes and compared their performance and model discrimination. The four ML algorithms included random forest (RF), gradient boosting machine (GBM), extreme gradient boosting (XGbosst), and support vector machines (SVM). Parameter tuning was performed by an automatic grid search with 10 different values to randomly try for each algorithm parameter. All the hyperparameter evaluation and model development were performed using the Caret package in R Studio (22). Each model was trained, and their performance were evaluated using the test set.

4.2.5 Comparison Measures

To compare the performance of applied models, the area under the curve (AUC) was calculated which measures the model discrimination. AUC is a performance measure that indicates the ability of a classifier in distinguishing true positives (TP) and true negatives (TN). In our study, the number of patients who were readmitted and correctly classified in the model were TP and the number of stroke patients who were not readmitted and correctly identified by the model were TN. We also used other performance measures such as sensitivity or recall, specificity and positive predictive value (PPV) which can be calculated as:

$$PPV = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} = \frac{TP}{TP + FP}$$

Further, to identify the list of the most significant variables associated with the response variable in descending order, we used variable importance. Variable importance represents the significance of each variable with respect to its effect on the model. In the variable importance list, the variables' scores were scaled between zero and 100, higher the scores were associated with higher the variables contribution to model prediction.

4.3 Results

4.3.1 Study Design and Population Characteristics

A total number of 3,184 records were used in this study. There were 1,960 patient records for Geisinger Medical Center (GMC) and 1,224 records for Geisinger Wyoming Valley (GWV) in the dataset. The data de-identification was performed by an honest data broker from the Phenomic Analytics and Clinical Data Core (PACDC) team at Geisinger. The train set and test set included 2,548 (80%) and 636 (20%) observation respectively.

Among 3,184 patients with ischemic stroke, 301 (9.4%) were readmitted within 30 days. The average age of patients was 71 [interquartile range, IQR: 18-89] and 1,611 (50.6%) patients were male. There was a significant difference among patients that were readmitted and those that were not in terms of median LOS, being married or previously married, discharged to SNF or against medical advice, and being Medicare or private insurance users (see Table 12). This study included 47 patients' comorbidities and 5 procedures—which were selected based on expert opinion and past studies in the literature—, and 31 out of these 52 (59.6%) were significantly different among two groups. Table 12 included the patients' demographics, past medical history, and outcome.

The patients were compared regarding their demographical characteristics, medical history prior to ischemic stroke event, and stroke severity using the “National Institutes of Health Stroke Scale” (NIHSS) for a subset of patients with NIHSS data available. The categorical variables were compared using the chi-square test. Also, we performed pearson's correlation coefficient for continuous variables. The correlation matrix between all the predictors along with a list of correlations above 30% and 50% is provided in Figure C1 and Table C1 in the appendix respectively. In Table 12, continuous variables were presented as average and standard deviation and as median with interquartile range (IQR). Also, for binary and categorical variables, their frequency and percentage were reported for each group. Student's t-test was used to identify the significant difference between two groups of patients (i.e., readmitted and not readmitted) for each predictor and the test statistic and P-values were reported in Table 12. Moreover, the generalized linear model (GLM) with binomial link function, which is known as logistic regression, was performed to find the log odds and 95% confidence interval for the significant predictors of 30-day readmission in stroke patients.

Table 12. Descriptive statistics of continuous variables.

	Variables	Missing	Not Readmitted (n=2883)	Readmitted (n=301)	t statistics	P-value
	Age (y), Mean (SD)	12	71.1 (13.9)	71.5(12.9)	-0.52	0.60
	LOS* (d), Median (IQR)	319	3 [1, 76]	4 [1, 24]	-4.03	0.00
Gender, n (%)	All gender	12	2873	299	-	-
	Female		1406 (48.8)	155 (51.5)	-0.90	0.39
	Male		1467 (50.9)	144 (47.8)	1.00	0.31
Marital Status, n (%)	Total	12	2873	299		
	Married		1349 (46.8)	124 (41.2)	1.85	0.06
	Single		425 (14.7)	45 (15.0)	-0.10	0.92
	Previously married		1099 (38.1)	130 (43.2)	-1.72	0.08
NIHSS Score, n (%)	0 to above 24	2545	580	59	-	-
	0 to 4		330 (11.4)	31 (10.3)	0.60	0.55
	5 to 11		150 (5.2)	17 (5.6)	-0.33	0.74
	12 to 23		75 (2.6)	10 (3.3)	-0.74	0.46
	Above 24		25 (0.9)	1 (0.3)	0.98	0.33
Procedures, n (%)	Total	-	291	36	-	-
	Received TPA**		71 (3.0)	2 (0.8)	2.06	0.04
	Insert indwelling urinary catheter		3 (0.1)	1 (0.3)	-1.07	0.28
	Insert endotracheal tube		148 (5.1)	10 (3.3)	1.35	0.17
	Percutaneous gastrostomy		42 (1.5)	16 (5.4)	-4.81	0.00
	Hemodialysis		27 (0.9)	7 (2.3)	-2.25	0.02
Hospital, n (%)	All centers	-	2883	301	-	-
	GMC***		1784 (61.9)	176 (58.5)	1.08	0.27
	GWV◇		1099 (38.1)	125 (41.5)	-1.16	0.25
Discharge Status, n (%)	Total	492	2418	274	-	-
	Discharged to home health organization		346 (12.0)	37 (12.3)	-0.15	0.88
	Discharged to home, court, or against medical advice		902 (31.3)	57 (18.9)	4.46	0.00
	Discharged to hospice-home/hospice-medical facility		82 (2.8)	4 (1.3)	1.54	0.12
	Discharged/transferred to other facilities		27 (0.9)	2 (0.7)	0.47	0.63
	Discharged/transferred to SNF ‡		447 (15.5)	91 (30.2)	-6.53	0.00
	Discharged/transferred to another rehab facility		614 (21.3)	83 (27.6)	-2.51	0.01
	Total	51	2836	297	-	-
Payer, n (%)	Direct employer contract		65 (2.3)	9 (3.0)	-0.80	0.42
	Medicaid		216 (7.5)	22 (7.3)	0.11	0.91
	Medicare		2037 (70.7)	230 (76.4)	-2.10	0.03
	Other government payers		58 (2.0)	4 (1.3)	0.81	0.41
	Private		425 (14.7)	30 (1.0)	2.25	0.02
	Self-pay		32 (1.1)	1 (0.3)	1.27	0.20
	Workers compensation		3 (0.1)	1 (0.3)	-1.06	0.29
Diagnoses, n (%)	Anemia	-	319 (13.7)	67 (26.4)	-5.42	0.00
	Atrial fibrillation		691 (29.4)	93 (36.3)	-2.29	0.02
	Anxiety disorders		328 (14.0)	53 (20.7)	-2.90	0.00
	Cerebral arterial dissection		23 (1.0)	6 (2.3)	-1.98	0.05
	Coronary artery disease		684 (29.1)	77 (30.1)	-0.32	0.75
	Delirium		44 (1.9)	12 (4.7)	-2.95	0.00
	Dementia		222 (9.5)	35 (13.7)	-2.15	0.03
	Diabetes		642 (27.5)	86 (33.9)	-2.13	0.03
	Dysphagia		150 (6.4)	29 (11.3)	-2.97	0.00
	Heart failure		426 (18.3)	64 (25.2)	-2.67	0.00
	Hypercoagulable state		17 (0.7)	7 (2.7)	-3.20	0.00

Hypertension	1308 (56.1)	131 (51.6)	1.38	0.17
Hypotension	55 (2.3)	13 (5.1)	-2.61	0.00
Kidney disease	670 (28.5)	101 (39.5)	-3.65	0.00
Malignancy	286 (12.2)	47 (18.4)	-2.82	0.00
Malnutrition	105 (4.5)	32 (12.5)	-5.49	0.00
Migraine	69 (2.9)	13 (5.1)	-1.86	0.06
Overweight	58 (2.5)	16 (6.2)	-3.46	0.00
Tobacco use	1171 (49.9)	143 (55.9)	-1.83	0.07
Venous thrombosis	81 (3.4)	18 (7.0)	-2.85	0.00
Acute myocardial infarction	36 (1.5)	9 (3.5)	-2.31	0.02
Alcohol use	95 (4.0)	10(3.9)	0.11	0.91
Arrhythmias	99 (4.2)	15 (5.9)	-1.22	0.22
Blindness	14 (0.6)	1 (0.4)	0.41	0.68
Cardiac valvular disease	159 (6.8)	22 (8.6)	-1.09	0.27
Cardiomyopathy	131 (5.6)	23 (9.0)	-2.20	0.03
Cerebral atherosclerosis	93 (4.0)	13 (5.1)	-0.86	0.39
Chronic kidney disease	583 (25.0)	89 (35.0)	-3.47	0.00
Chronic liver disease	49 (2.1)	2 (0.8)	1.43	0.15
Chronic lung disease	476 (20.3)	69 (27.0)	-2.50	0.01
Dysautonomia	16 (0.7)	2 (0.8)	-0.18	0.85
Hyperlipidemia	1537 (65.4)	177 (69.1)	-1.19	0.23
Intracerebral hemorrhage	521 (22.2)	51 (19.9)	0.83	0.41
Inflammatory disorders	66 (2.8)	5 (2.0)	0.80	0.42
Mood disorders	372 (15.8)	58 (22.7)	-2.79	0.00
Non-compliance	107 (4.6)	12 (4.7)	-0.09	0.92
Normal weight	44 (1.9)	11 (4.3)	-2.56	0.01
Obese	466 (19.8)	62 (24.2)	-1.65	0.09
Palliative care on board	254 (10.8)	12 (4.7)	3.08	0.00
Peripheral vascular disease	125 (5.3)	17 (6.6)	-0.88	0.38
Respiratory failure	164 (7.0)	19 (7.4)	-0.26	0.79
Seizure disorders	96 (4.1)	15 (5.9)	-1.33	0.18
Sleep apnea	216 (9.2)	25 (9.8)	-0.30	0.76
Systemic infection	63 (2.7)	16 (6.2)	-3.17	0.00
Thyroid disease	445 (18.9)	54 (21.1)	-0.83	0.41
Underweight	34 (1.4)	9 (3.5)	-2.47	0.01
Use of steroids	72 (3.1)	11 (4.3)	-1.06	0.29
Year, Median (IQR)	-	2016 [2015, 2018]	2016 [2015, 2018]	-0.36 0.72

Abbreviations: *Length of Stay, ** Received Tissue Plasminogen activator, *** Geisinger Medical Center, ◇ Geisinger Wyoming Valley, † Skilled Nursing Facility

4.3.2 Results of the Logistic Regression

The result of logistic regression with feature selection and ROSE sampling was provided in Table 13. The odd ratios (OR), log odds, 95% CI, and P-values were reported in this table. This analysis revealed that being discharged to SNF, malignancy, and malnutrition were significantly associated with stroke readmission within 30 days (p-value < 0.0001). Also, being discharged to rehab facility and stroke severity above twelve were significantly associated with 30-day readmission at 0.001 significance level. According to these results, being discharged to SNF increased the odds of readmission by 2.76. Also, malignancy and malnutrition increased the odds of 30-day readmission by 2.10 and 2.51 respectively.

Table 13. Logistic regression analysis results for predictors of 30-day readmission in stroke patients.

Variables	OR	Log Odds	95% CI (2.5%, 97.5%)		P-value
(Intercept)	0.250	-1.387	-2.548	-0.300	0.015 *
Age	0.987	-0.013	-0.027	0.001	0.072
Discharged to home health organization	1.606	0.473	-0.007	0.943	0.050
Discharged to hospice-home or hospice-medical facility	0.437	-0.828	-2.041	0.276	0.157
Discharged/transferred to other facilities	0.495	-0.704	-2.621	0.665	0.380
Discharged/transferred to another rehab facility	1.788	0.581	0.172	0.995	0.006 **
Discharged/transferred to SNF	2.766	1.017	0.558	1.482	0.000 ***
Medicaid	0.410	-0.892	-1.797	0.067	0.058
Medicare	0.575	-0.553	-1.313	0.302	0.175
Other government payers	0.394	-0.930	-2.334	0.317	0.159
Private insurance	0.401	-0.914	-1.752	-0.003	0.039 *
Self-pay	0.291	-1.235	-4.211	0.599	0.267
Workers compensation	1.821	0.600	-2.586	3.044	0.649
Chronic kidney disease	1.276	0.244	-0.360	0.891	0.444
Hypercoagulable state	3.097	1.130	0.182	1.993	0.013 *
Kidney disease	1.150	0.140	-0.492	0.719	0.649
Malignancy	2.102	0.743	0.375	1.098	0.000 ***
Malnutrition	2.515	0.922	0.395	1.425	0.000 ***
Palliative care on board	1.134	0.126	-0.727	0.903	0.761
Respiratory failure	0.738	-0.304	-1.065	0.386	0.409
Underweight	1.005	0.005	-1.012	0.910	0.991
Insert endotracheal tube	0.864	-0.147	-1.114	0.745	0.756
Percutaneous gastrostomy	1.391	0.330	-0.512	1.095	0.416
NIHSS 12 to 23	1.765	0.568	0.142	0.986	0.008 **
NIHSS 5 to 11	0.651	-0.429	-0.792	-0.076	0.019 *
NIHSS above 24	0.171	-1.766	-3.250	-0.648	0.006 **

Significance codes for P-values: 0 **** 0.001 *** 0.01 **

4.3.3 Results of the Machine Learning Methods

The performance metrics of the machine learning methods were provided in Table 14, Table 15, and Table 16. Sensitivity, specificity, PPV, and the computation time for all of the methods with and without ROSE sampling and feature selection were reported in these tables. We provided the results for both training and testing sets for each model. The results showed that applying ROSE for class imbalance correction improved testing AUC, PPV, and specificity of ML models. However, feature selection did not upgrade the results.

In the absence of sampling and feature selection, GBM provided the highest testing AUC (0.68), specificity (0.95), and PPV (0.33) when compared to the other models (see Table 14). However, the best testing AUC, PPV, sensitivity, and specificity were reached when ROSE sampling was applied (see Table 15). According to results, the best prediction model of 30-day readmission in stroke patients was achieved by XGBoost with the highest testing AUC (0.74). The reported PPV

in this model was 0.43, with 0.20 sensitivity, and 0.98 specificity. The optimal parameters for ROSE sampled XGBoost were max-depth=4, subsample=0.5, colsample_bytree=0.8, gamma=0, and min_child_weight=10. Based on AUC and PPV, logistic regression had poor performance compared to other models with or without sampling and feature selection. Further, we used ROSE sampled XGBoost model as the best predictive model to identify the most important predictors of 30-day readmission in ischemic stroke patients.

Variable importance identifies the associated variables with chances of 30-day readmission after stroke. According to the variable importance scores for XGBoost with ROSE class imbalance correction (Table 17), the top 15 predictors of 30-day readmission were NIHSS above 24, insert indwelling urinary catheter, hypercoagulable state, percutaneous gastrostomy, using workers compensation as insurance, hemodialysis, overweight, cerebral arterial dissection, malnutrition, received TPA, venous thrombosis, discharged to hospice, home, or hospice-medical facility, palliative care on board, delirium, and self-pay insurance type. The importance scores of all these 15 variables were reported in Table 17.

Table 14. Performance metrics for machine learning models.

Method	Train set					Test set			
	AUC	Sensitivity	Specificity	PPV	Time (sec.)	AUC	Sensitivity	Specificity	PPV
GLM (LR)	0.76	0.26	0.94	0.46	3	0.60	0.32	0.86	0.19
RF	0.82	0.63	0.94	0.48	42	0.57	0.09	0.93	0.12
GBM	0.70	0.23	0.95	0.33	48	0.68	0.23	0.95	0.33
XGBoost	0.76	0.32	0.94	0.36	1,752	0.62	0.30	0.88	0.21
SVM	0.98	0.45	0.98	0.98	990	0.62	0.30	0.86	0.18

Table 15. Performance metrics, with ROSE sampling, for machine learning models.

Method	Train set					Test set			
	AUC	Sensitivity	Specificity	PPV	Time (sec.)	AUC	Sensitivity	Specificity	PPV
GLM (LR)	0.74	0.65	0.71	0.19	3	0.63	0.38	0.72	0.12
RF	0.74	0.12	0.97	0.47	33	0.67	0.09	0.97	0.26
GBM	0.74	0.10	0.98	0.43	48	0.70	0.09	0.98	0.45
XGBoost	0.76	0.18	0.98	0.54	2,340	0.74	0.20	0.98	0.43
SVM	0.83	0.55	0.91	0.38	1,689	0.67	0.38	0.89	0.27

Table 16. Performance metrics, with feature selection and ROSE sampling, for machine learning models.

Method	Train set					Test set			
	AUC	Sensitivity	Specificity	PPV	Time (sec.)	AUC	Sensitivity	Specificity	PPV
GLM (LR)	0.70	0.62	0.68	0.17	3	0.64	0.53	0.69	0.15
RF	0.70	0.29	0.92	0.27	12	0.65	0.30	0.89	0.24
GBM	0.69	0.18	0.94	0.24	30	0.66	0.17	0.95	0.26
XGBoost	0.70	0.18	0.94	0.26	2,130	0.65	0.17	0.95	0.27
SVM	0.72	0.48	0.80	0.20	960	0.64	0.42	0.77	0.16

Table 17. Variable importance scores of XGBoost model with ROSE sampling.

No.	Variable	Importance Score (out of 100)
1	NIHSS above 24	100.00
2	Insert Indwelling Urinary Catheter	89.94
3	Hypercoagulable State	61.95
4	Percutaneous Gastrostomy	40.04
5	Payer Workers Compensation	37.78
6	Hemodialysis	35.13
7	Overweight	33.35
8	Cerebral Arterial Dissection	30.61
9	Malnutrition	25.74
10	Received TPA	24.15
11	Venous Thrombosis	22.94
12	Discharged to Hospice-Home or Hospice-Medical Facility	17.15
13	Palliative Care on Board	14.92
14	Delirium	14.54
15	Payer Self-Pay	14.42

4.3.4 Model Comparison and Performance

We used a test set to compare the performance of logistics regression and four machine learning techniques in predicting the chances of 30-day readmission. The ROC curves for logistics regression, random forest (RF), gradient boosting machines (GBM), extreme gradient boosting (XGBoost), and support vector machines (SVM) without feature selection and sampling were shown in Figure 6 and with ROSE-sampling were presented in Figure 7. In the absence of feature selection and sampling, the best model was GBM with AUC of 0.68 for the test set (Figure 6) while after ROSE-sampling XGBoost was the best performing model with 0.74 testing AUC (Figure 7). In terms of the computation time for training the models, logistic regression, RF and GBM models performed faster than XGBoost and SVM models (see Table 14, Table 15, and Table 16 for more details). In these tables, we reported the computation time for each model which performed on a macOS with 3.1 GHz Dual-Core Intel Core i5. XGBoost and SVM models

had longer computation time compared to other models. However, the computation time for SVM was shorter than XGBoost models. For example, SVM with ROSE sampling was 39% faster than XGBoost.

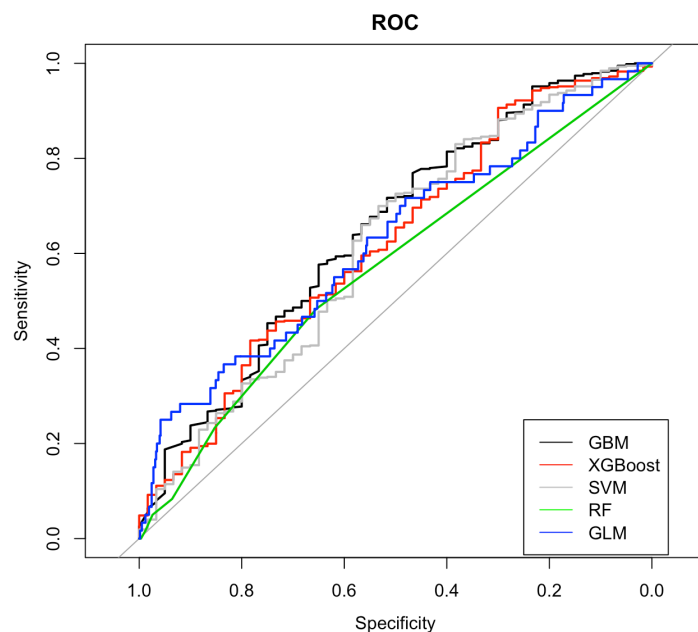


Figure 6. ROC curves for ML models.

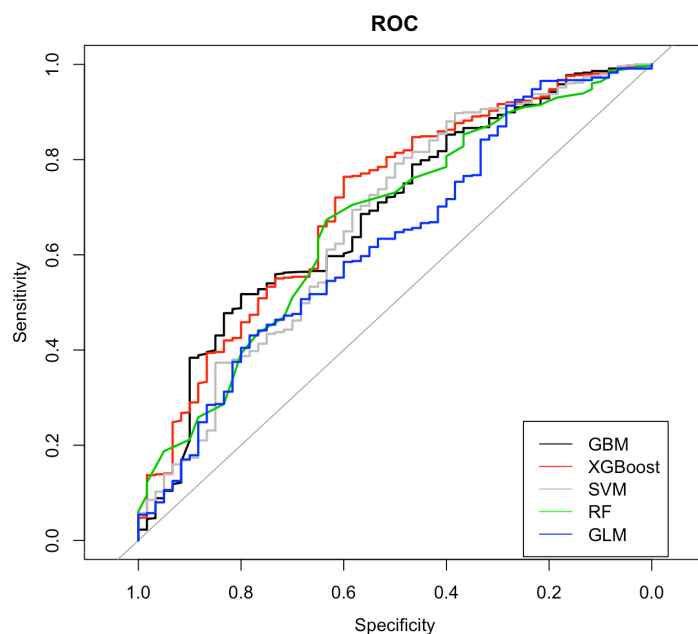


Figure 7. ROC curves for ML models with ROSE-sampling.

4.4 Discussion

Readmissions within 30 days were associated with higher costs and a higher risk of mortality compared to the first admissions (23). In this study, we aimed to develop prediction model for 30-days readmission among ischemic stroke patients and identify associated variables. We focused on 30-day readmission which is an important criterion for hospitals because Medicare has introduced penalties for readmission within 30 days. Also, 30-day readmission were associated with higher costs and a higher risk of mortality compared to the first admission (23). In this work, statistical and ML techniques were applied and their effectiveness were compared to find the best predictive model. Finding the most suitable methods for our dataset was a crucial step. Our dataset had four specific characteristics that guided us through the method selection.

First, the distribution of the dependent variable was imbalanced. One of the best techniques for predicting the minority category was decision trees. Second, there were more than 50% collinearities between some of the predictors (see Table C1 in the appendix). Logistic regression had the assumption of little or no multicollinearity among predictors; thus, we needed to use other models that were able to address the collinearity. Third, we were interested in including factors with nonlinear and high dimensional relationships among a multitude of variables. Fourth, our dataset had 58 categorical variables. This limited the application of distance-based methods such as k-nearest neighbors. Therefore, considering these four characteristics of the dataset, we applied random forest (RF), gradient boosting machines (GBM), extreme gradient boosting (XGBoost), and support vector machines (SVM).

Past studies mostly used logistic regression and included none or a limited number of diagnoses in their analysis (12). Also, papers that used other machine learning techniques to improve the prediction power, either performed their analysis on readmission more than 30-day or studied other causes of readmission such as heart failure (14, 15, 17). This study was the first in predicting the associated variables of 30-day ischemic stroke readmission using machine learning techniques.

Our findings showed that stroke severity (NIHSS score above 24), hypercoagulable state, and malnutrition were significantly associated with the 30-day readmission in both logistic regression (with feature selection and sampling) and XGBoost (with sampling). Additionally, malignancy, NIHSS scores between 5 to 23, private insurance type, and being discharged to rehab facility or SNF were only significant in the logistic regression model, and they had low importance score in the XGBoost model. Among all variables, stroke severity and malnutrition were found significant

predictors of 30-day readmission in ischemic stroke patients in past studies and our results corroborated the previous findings (2, 3, 5-7, 9). Despite other studies, results from this study showed that insert indwelling urinary catheter, hemodialysis, and cerebral arterial disease were significantly associated with readmission based on the XGBoost model. It has been shown in previous studies that heart failure and being a Medicare or Medicaid user were significantly correlated with 30-day readmission (2, 5, 7-9). However, we found no evidence in favor of these assertions. Past studies provided mixed results on the importance of age, hypertension, and gender; some studies found that patients with older age were more likely to be readmitted (2, 6) while others showed that age was not a significant predictor (3, 11). Also, hypertension was found as a significant risk factor of readmission in a study (8) while in other works authors claimed that hypertension was not significantly associated with 30-day readmission (11, 12). There were several studies conducted on data from Taipei, China, and western Australia that found that gender of patient was not significantly associated with chances of being readmitted (3, 6, 11), however past analyses of the U.S. data showed that women were significantly at higher risk of readmission after stroke (2, 8, 12). The results of the ROSE-sampled XGBoost model indicated that age, hypertension, and gender were not significantly associated with 30-day readmission.

Based on our analysis, the best performing predictive model, which was ROSE sampled XGBoost, had 17.5% improvement in AUC compared to logistic regression. The XGBoost model performed better in comparison with other models of 30-day readmission in the literature (14, 15, 17). We improved the AUC up to 0.74 (95% CI: 0.64, 0.78) for the test set with 0.43 PPV (see XGBoost model in Table 15). In the absence of sampling and feature selection, GBM returned very close AUC for the training and testing sets, corroborating that the models did not suffer from overfitting. XGBoost and GBM with ROSE sampling returned close AUC for the testing and training sets, confirming that they did not overfit as well. However, the SVM-based models had the largest difference between testing and training AUC, leading to the possibility of overfitting given this dataset. The provided results in Table 14, Table 15, and Table 16 and the ROC curves shown in Figure 6 and Figure 7 revealed that machine learning methods such as GBM and XGBoost improved the prediction of 30-day readmission in stroke patients compared to traditional logistic regression.

4.5 Strengths, Limitations, and Future Directions

One of the important strengths of this study was that we analyzed a diverse list of potential predictors including an extensive number of clinical interventions and patient's comorbidities. Our

data set had unique characteristics; it was obtained from a large medical system and it mainly included data from non-urban area in which the majority of patients were white. To the best of our knowledge, this was the first attempt to apply machine learning techniques to the 30-day readmission data for ischemic stroke patients. Considering the large number of included variables in our dataset, these ML techniques were able to include higher order interactions among variables, yet improve the prediction power compared to logistic regression.

Our analysis had several limitations. Although our dataset was rich in number of variables, the number of patients was relatively small compared to the included independent variables. Therefore, the small number of observations might result in overfitting in the models. Another limitation of this work was missing data specifically for NIHSS score. The most missing data points belonged to NIHSS score before 2016 and we applied imputation to not lose any observation. However, we had around 3.5% missingness in the rest of the data. Additionally, due to the demographic characteristics of this dataset, the results may not be generalizable to other medical systems.

In this study we only considered ischemic stroke as the cause of readmission. Therefore, future avenues of research can be done by considering other stroke types. However, considering the size of our dataset which came from two health centers, further work needs to focus on a larger population with diverse demographics to introduce a generalizable model. Additionally, to improve the prediction power, future study may apply deep learning techniques and compare their performance with machine learning techniques.

4.6 Conclusion

This study identified the most significant predictors of 30-day readmission in patients with ischemic stroke. We developed different predictive models using ML techniques and compared them together and with the traditional logistic regression. Our results supported the hypothesis that machine learning predictive models perform better compared to the traditional logistic regression. These ML models had a good predictive power while including a comprehensive list of variables. The insights of this work can assist with the identification of ischemic stroke patients who are at higher risk of readmission. These results can be used to inform practitioners and healthcare managers to enhance the quality of care, pay special attention to high-risk patients, and reduce the chances of 30-day readmission. As a result, medical systems using predictive models in care management could more effectively improve care and outcome for their stroke

patients. Moreover, this study can increase the understanding of patients and caregivers about the factors that are associated with higher chances of hospital readmissions. Once validated in independent cohort, this model can be integrated into EHR and provide decision support system to the healthcare providers.

References

1. LaPointe J. 3 Strategies to Reduce Hospital Readmission Rates, Costs 2018 [Available from: <https://revcycleintelligence.com/news/3-strategies-to-reduce-hospital-readmission-rates-costs>].
2. Lichtman JH, Leifheit-Limson EC, Jones SB, Wang Y, Goldstein LB. Preventable Readmissions Within 30 Days of Ischemic Stroke Among Medicare Beneficiaries. *Stroke*. 2013;44(12):3429-35.
3. Chuang K-Y, Wu S-C, Ma A-HS, Chen Y-H, Wu C-L. Identifying factors associated with hospital readmissions among stroke patients in Taipei. *Journal of Nursing Research*. 2005;13(2):117-28.
4. Center for Disease Control and Prevention. Stroke Facts 2017 [Available from: <https://www.cdc.gov/stroke/facts.htm>].
5. Jia H, Zheng Y, Reker DM, Cowper DC, Wu SS, Vogel WB, et al. Multiple system utilization and mortality for veterans with stroke. *Stroke*. 2007;38(2):355-60.
6. Wen T, Liu B, Wan X, Zhang X, Zhang J, Zhou X, et al. Risk factors associated with 31-day unplanned readmission in 50,912 discharged patients after stroke in China. *BMC neurology*. 2018;18(1):218.
7. Smith MA, Liou J-I, Frytak JR, Finch MD. 30-day survival and rehospitalization for stroke patients according to physician specialty. *Cerebrovascular Diseases*. 2006;22(1):21-6.
8. Kennedy BS. Does race predict stroke readmission? An analysis using the truncated negative binomial model. *Journal of the National Medical Association*. 2005;97(5):699.
9. Smith MA, Frytak JR, Liou J-I, Finch MD. Rehospitalization and survival for stroke patients in managed care and traditional Medicare plans. *Medical care*. 2005;43(9):902.
10. Heller RF, Fisher JD, O'Este CA, Lim LLY, Dobson AJ, Porter R. Death and readmission in the year after hospital admission with cardiovascular disease: the Hunter Area Heart and Stroke Register. *Medical journal of Australia*. 2000;172(6):261-5.
11. Lee AH, Yau KK, Wang K. Recurrent ischaemic stroke hospitalisations: A retrospective cohort study using Western Australia linked patient records. *European journal of epidemiology*. 2004;19(11):999-1003.
12. Lichtman JH, Leifheit-Limson EC, Jones SB, Watanabe E, Bernheim SM, Phipps MS, et al. Predictors of hospital readmission after stroke: a systematic review. *Stroke*. 2010;41(11):2525-33.
13. Ouwerkerk W, Voors AA, Zwinderman AH. Factors influencing the predictive power of models for predicting mortality and/or heart failure hospitalization in patients with heart failure. *JACC: Heart Failure*. 2014;2(5):429-36.
14. Mortazavi BJ, Downing NS, Bucholz EM, Dharmarajan K, Manhapra A, Li S-X, et al. Analysis of machine learning techniques for heart failure readmissions. *Circulation: Cardiovascular Quality and Outcomes*. 2016;9(6):629-40.
15. Golas SB, Shibahara T, Agboola S, Otaki H, Sato J, Nakae T, et al. A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. *BMC medical informatics and decision making*. 2018;18(1):44.
16. Francisco A, Stabler ME, Hisey W, Mackenzie TA, Dorn C, Denton J, et al. Using Machine Learning to Predict 30-Day Readmission of Patients Hospitalized With an Acute Myocardial Infarction. *Circulation*. 2018;138(Suppl_1):A15808-A.
17. Wolff P, Graña M, Ríos SA, Yarza MB. Machine Learning Readmission Risk Modeling: A Pediatric Case Study. *BioMed research international*. 2019;2019.
18. Kalagara S, Eltorai AE, Durand WM, DePasse JM, Daniels AH. Machine learning modeling for predicting hospital readmission following lumbar laminectomy. *Journal of Neurosurgery: Spine*. 2018;30(3):344-52.

19. Zhang Z. Multiple imputation with multivariate imputation by chained equation (MICE) package. *Annals of translational medicine*. 2016;4(2).
20. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. *J Stat Softw*. 2010;36(11):1-13.
21. Lunardon N, Menardi G, Torelli N. ROSE: A Package for Binary Imbalanced Learning. *R journal*. 2014;6(1).
22. Kuhn M. Building predictive models in R using the caret package. *Journal of statistical software*. 2008;28(5):1-26.
23. Nouh AM, McCormick L, Modak J, Fortunato G, Staff I. High mortality among 30-day readmission after stroke: predictors and etiologies of readmission. *Frontiers in neurology*. 2017;8:632.

Chapter 5. Conclusion

Many public health challenges in the United States exist that impact individual well-being and overall population health., high infant mortality and readmission rate of stroke patients, and the opioid crisis have been among the list of top public health concerns in the U.S. in recent years (1-5). The major goal of this doctoral dissertation is to enhance understanding of these three critical challenges of the U.S. healthcare system using different levels of analysis in three stand-alone essays. In particular, I studied disparities in the performance of states' healthcare systems in terms of their birth outcomes (e.g., infant mortality rate). I also provided an explanation of variation in opioid overdose mortalities and explored the risk factors of 30-day readmissions in stroke patients. I employed several industrial engineering and data analytics approaches including data envelopment analysis, statistical analysis, and machine learning techniques to address these existing challenges in the U.S. healthcare system.

In Chapter 2, I focused on the performance of the healthcare system of each state in terms of their birth outcomes. I investigated the relative efficiency of health systems in reducing unfavorable outcomes such as infant mortality, preterm birth, and low birthweight. Considering that a health system's environment affects its operation, I tested the impact of some environmental factors that might obstruct or facilitate the improvement of healthcare systems' outcome. This work was the first study that applied data envelopment analysis to assess, compare and contrast the performance of healthcare systems through quantitative measures at the state level. I constructed a unique data set by obtaining data from several resources. Our findings revealed that in comparison to the states' health systems performance, it is important to not limit the analysis to only resources and expenditures of the healthcare systems. Socioeconomic and demographic factors affect the effectiveness of health systems and need to be considered in the model. The results of our model showed that states with a lower rate of African-American women and poverty are more successful in reducing unfavorable birth outcomes. However, this study is a starting point to apply well-established mathematical modeling in the arena of efficiency measurement of health systems at the state-level. Further research is required to enhance the understanding of the impact of social and personal factors on the birth outcomes.

In Chapter 3, I studied the trends of opioid overdose deaths and opioid prescribing rates at the state-level between 2008 to 2017. I exploited state variation in opioid overdose mortality rates to identify factors contributing the most to mortality reduction. Further, I investigated the effect of

several state laws and programs enacted to combat the opioid epidemic. These programs include dose and duration limitations on the initial prescription, pain management clinic laws, mandated use of prescription drug monitoring programs (PDMPs), and medical cannabis laws. Three research questions guide the second essay: (1) What factors explain mortality from opioid painkillers? (2) What are the determinants of state variations in mortalities from an opioid overdose? (3) Is there any correlation between states' laws and programs and opioid prescription rates and overdose mortality rates? I applied two-way fixed-effect regressions to address these questions. The results of our study contribute to the body of literature in the opioid epidemic in several ways. It provides new insights on the main drivers of opioid overdose death rates and opioid prescription rates in states. Understanding the effect of different factors on lowering prescribing rates and consequently overdose mortalities may help curb the opioid epidemic in the U.S. Our finding shows that a higher share of uninsured people and binge drinkers are associated with higher overdose deaths. Although we found significant negative associations between opioid prescription rates and mandated PDMPs, limitations on dose/duration of the initial prescription, and pain clinic laws, they were not negatively associated with overdose mortalities. However, once we considered a delay for mandated PDMPs, a significant association appeared. In a nutshell, state laws are unlikely to result in an immediate reduction in overdose deaths, and comprehensive interventions are needed to curb the epidemic.

In the third essay (Chapter 4), I aimed to advance the understanding of the potential risk factors of 30-day readmission in stroke patients. I applied several statistical and machine learning techniques and compared their performance to identify the best predictive model. In this study, I used the electronic health records (EHR) of stroke patients extracted from two stroke centers within the Geisinger Health System from 2015 to 2018. I used logistic regression and proper machine learning techniques, which is the first time that such analysis has been done for 30-day readmission in stroke patients. Also, we included a wide list of potential predictors including the patient's comorbidities and clinical interventions in the models. One of the major findings was that machine learning-based models can be designed to better predict 30-day readmissions after stroke. We showed that stroke severity, indwelling urinary catheter, and hypercoagulable state are more important than generally known diagnoses such as diabetes and hypertension.

5.1 Contributions

This dissertation provides a better understanding of the several health-related problems and their complexities. Our study may assist higher-level decision-makers to make proper decisions

regarding the design and implementation of interventions and programs that can address healthcare challenges. To do so, we proposed important research questions and introduced innovative industrial engineering and data analytics methods to assess the effectiveness of healthcare systems, evaluate the impact of adopted policies, and explore the main predictors of an important challenge of medical systems. Scholars in the healthcare arena have studied similar problems; however, the questions that we answered, and our modeling approaches bring novel insights at state- and individual-level. Therefore, this dissertation provides several major contributions to the body of research.

The results of the first essay contribute to the literature of birth outcomes and infant mortality by testing two theoretical explanations for the healthcare system's success in improving the birth outcomes at the state level. While many studies that applied DEA, failed to consider the effect of contextual variables or only focused on the performance of the healthcare system at the national level, our study took one further step. We investigated the characteristics of every healthcare system at the state-level by focusing on their birth outcomes while considering the impact of environmental factors. We applied a mathematical modeling approach to a unique healthcare domain with a unique dataset that has never been studied before. We also identified best practices or peers for each state's health system using data envelopment analysis. This is helpful especially in determining the key components of successful healthcare systems.

Essay 2 contributes to the literature of mortalities from opioid overdose in the U.S. The novelty of this work lies in clarifying the effect of enacted laws, programs, and policies (i.e., dose and duration limitations on the initial prescription, pain management clinic laws, mandated use of PDMPs, and medical cannabis laws) in states to curb the opioid epidemic. Scholars in this field have studied the impact of some of these state programs; however, their results are contradictory. On the other hand, their study timeframes range from 1999 to 2013 which limits their ability to investigate long term impacts of the policies because most opioid-related policies were implemented after 2013. However, our analysis is applied to the most recent data to investigate the correlation between these policies—with and without delays—and opioid prescription and mortality rates. Unlike previous studies, we control for multiple factors including median household income, insurance coverage type, binge drinking, and tobacco use. Our results imply that in addition to upstream prevention policies, states need to implement downstream programs to curb the epidemic.

Essay 3 contributes to the literature of stroke readmissions and advances the application of machine learning techniques to provide a better predictive model in this arena. To our knowledge, it is the first attempt to improve the prediction power of 30-day readmissions in ischemic stroke patients including an extensive list of variables as well as prioritizing the importance of each variable. These variables are a comprehensive number of potential factors that are found to be important in the literature and by experts' opinions. Also, our data set is unique and it is representative of a non-urban population in which the majority of patients were white. The proposed model may facilitate the identification of patients who are at greater chances of readmission and need special clinical attention. Further, it can enhance the quality of care delivery and reduce stroke-related readmissions and its noticeable financial costs burden on the medical systems.

5.2 Broader Impacts

This dissertation is defined based on three major public health challenges in the U.S. I hope that the insights of this work help policymakers and practitioners at state and organizational level in a way to enhance the population health, quality of healthcare, and societal well-being. Data envelopment analysis provides policymakers with the opportunity of policy testing on how to reduce unfavorable birth outcomes such as infant mortality. This could be helpful especially for states that are suffering from inefficiencies in their healthcare system. The results of data envelopment analysis can assist states' decision-makers to find their best practices and realize where they can spend or save their resources to improve their performance. We also provided insights into the impact of interventions targeting opioid prescribing rates and opioid-related mortalities to enable states' policymakers to combat this national crisis. We found that regulations and laws are necessary as they are associated with a reduction in prescription rates with a short delay. However, it takes time to observe impacts on the mortality rates. In addition to introducing such regulations, policymakers and health managers need to implement evidence-based interventions and researchers should investigate other possible ways to reduce overdose death rates. Further, we identified and prioritized the main risk factors of an acute condition (i.e., ischemic stroke readmissions). Our findings can assist with the identification of ischemic stroke patients who are at higher risk of readmission. These results can be used to inform practitioners and healthcare managers to enhance the quality of care, pay special attention to high-risk patients, and reduce the chances of 30-day readmissions. As a result, medical systems will be able to save more lives and reduce the associated readmission costs. Moreover, this study can

increase the understanding of patients and caregivers about the factors that are associated with higher chances of hospital readmissions. Future avenue of research could be extensions of these three studies to further explore the underlying complexities of the abovementioned public health challenges.

5.3 Implications

The result of our work provides policymakers new insights on how states' healthcare systems are performing and how efficient they are in achieving desirable birth outcomes. Also, our results bring policymakers and health providers' attention to the importance of considering contextual factors in their interventions. Furthermore, our findings in the first essay can be informative for practitioners at the health systems in the identification of those who are at higher risk of giving unfavorable birth outcomes and need special clinical attention. The results of the second essay provide states' policymakers a better understanding of the impact of different enacted laws and programs on the opioid epidemic in the U.S. Finally, the findings of the third essay can be used to inform practitioners and healthcare managers to enhance the quality of care, pay special attention to high-risk patients, and reduce the chances of 30-day readmission. As a result, medical systems using predictive models in care management could more effectively improve care and outcome for their stroke patients. Moreover, this study can increase the understanding of patients and caregivers about the factors that are associated with higher chances of hospital readmissions. Ultimately, once validated on an independent cohort, the provided model can be integrated into EHR (electronic health record) and provide a decision support system to the healthcare providers.

References

1. CDC. Infant Mortality: What is CDC doing? 2017.
2. Managed Healthcare Exclusive. Top Public Health Issues Impacting Healthcare Organizations. 2018.
3. CDC. Public Health Professionals Gateway. 2017.
4. Center for Disease Control and Prevention. Stroke Facts 2017 [Available from: <https://www.cdc.gov/stroke/facts.htm>].
5. Liotta EM, Singh M, Kosteva AR, Beaumont JL, Guth JC, Bauer RM, et al. Predictors of 30 day readmission after intracerebral hemorrhage: a single-center approach for identifying potentially modifiable associations with readmission. *Critical care medicine*. 2013;41(12).

Appendices

Appendix A Bootstrapping, Robust PCA, and Slack Analysis

Bootstrapping Approach

Bootstrapping is the process of resampling with replacement several times (e.g., M times) from the original dataset to obtain a consistent estimation of empirical distribution. The process of drawing independent and identically distributed (iid) samples (i.e., pseudo-data) from all DMUs, which includes input, output, and environmental variables, is called data generating process (DGP). Hence, bootstrapping computes a new set of efficiency scores using the original estimator of the pseudo-data, which are supposed to follow the same distribution as the original efficiency scores.

Assume $x_{(1 \times p)}$ is the vector of input variables for DMUs, $y_{(1 \times q)}$ is the vector of outputs, and $z_{(1 \times r)}$ is the vector of environmental variables. Each sample of (x_i, y_i, z_i) , which represents the set of input, output, and environmental variables of DMU_i , are iid with density function $f(x, y, z)$. Let s_i represents the true efficiency score of DMU_i and $\mathcal{P}$ denotes the production frontier, which is constructed only by inputs and outputs:

$$\mathcal{P} = \{(x, y) | x \text{ can produce } y\} \quad (1)$$

Every DMU that is located on this production frontier has efficiency score of one (i.e., $s_i = 1$) and any DMU which is located inside the production frontier has efficiency score of less than one (i.e. $s_i < 1$). Suppose we have $\delta_i = \delta(x_i, y_i | \mathcal{P})$ and it represents the Euclidean distance from (x_i, y_i) , therefore we conclude that $\delta_i = \frac{1}{s_i}$. Since an efficiency score is always a number between zero and one ($0 \leq s_i \leq 1$), then $\delta_i \geq 1$. Also, δ_i can be represented as a linear function of environmental variables plus an error (1):

$$\delta_i = \beta z_i + \varepsilon_i \geq 1 \quad (2)$$

In the equation (2) ε_i represents the error term, which is a random variable with $\mathcal{N}(0, \sigma)$ distribution and β is the vector of parameters. If we consider $\hat{\delta}_i$ as the estimator of δ_i , we can write it as follows ($\epsilon_i \equiv \mathcal{N}(0, \sigma)$ is the error term):

$$\hat{\delta}_i = \hat{\beta} z_i + \epsilon_i \geq 1 \quad (3)$$

On the other hand, we can write an estimator as:

$$\hat{\delta}_i = E(\hat{\delta}_i) + \xi_i \quad (4)$$

Where ξ_i is the error term with expected value of zero ($E(\xi_i) = 0$). Also, the bias of $\hat{\delta}_i$ for DMU_i and the estimated bootstrap bias for DMU_i will be:

$$bias \equiv E(\hat{\delta}_i) - \delta_i \quad (5)$$

In the equation (4) the error term (ξ_i) is negligible in the *bias* because $\hat{\delta}_i$ is a consistent estimator of δ_i . Despite ξ_i which has expected value of zero and cannot be estimated, the *bias* does not have a zero mean and can be estimated. Bootstrapping method introduces a bias estimation ($\widehat{bias}$) as *bias* plus residuals. Therefore, considering this definition of *bias* provided in equation (5) we can estimate the *bias* as follows:

$$\widehat{bias} = E(\hat{\hat{\delta}}_i) - \hat{\delta}_i \quad (6)$$

Therefore, to find the bias-corrected estimator of δ_i (i.e., estimate the bias-corrected efficiency score (i.e., $\hat{\hat{\delta}}_i$)) we can use the following formula (1):

$$\hat{\hat{\delta}}_i = \hat{\delta}_i - \widehat{bias} \quad (7)$$

Then we can substitute $\widehat{bias}$ using equations (4) and (6) the bias-corrected efficiency scores can be calculated as:

$$\hat{\hat{\delta}}_i = 2\hat{\delta}_i - E(\hat{\hat{\delta}}_i) = 2\hat{\delta}_i - \frac{\sum_M \hat{\hat{\delta}}_i}{M} \quad (8)$$

Where M denotes number of sampling with replacement in the DGP that we repeated.

Simar (2) obtains a confidence interval for each efficiency score and we will construct the confidence intervals based on his approach and the definition of confidence intervals from Ross (3) book. For any random variable (like $\hat{\hat{\delta}}_i - \delta_i$), if the distribution of that random variable is known, we can find u_α and v_α such that ($\alpha \in [0,1]$):

$$\Pr\left\{u_{\frac{\alpha}{2}} \leq \hat{\hat{\delta}}_i - \delta_i \leq v_{\frac{\alpha}{2}}\right\} = 1 - \alpha \quad (11)$$

Therefore, a $(1 - \alpha)\%$ confidence interval for δ_i will be obtained by:

$$\hat{\delta}_i - v\frac{\alpha}{2} \leq \delta_i \leq \hat{\delta}_i - u\frac{\alpha}{2} \quad (12)$$

Now if we substitute δ_i by $\frac{1}{s_i}$ then the $(1 - \alpha)\%$ confidence interval for DMU_i 's efficiency score will be as it is shown in equation (13) and in this way we can find the confidence interval for DMUs.

$$\frac{1}{\hat{\delta}_i - u\frac{\alpha}{2}} \leq s_i \leq \frac{1}{\hat{\delta}_i - v\frac{\alpha}{2}} \quad (13)$$

Robust PCA

Principal Component Analysis (PCA) is a well-known statistical method of transforming observations into a hyperplane. This hyperplane which represents the linear combination (i.e., principal components) of observations accounts for variability among dataset. Every observation in the dataset can be defined by one or more principal components (PCs). Therefore, PCs can be computed as the eigenvector of the covariance matrix.

Although PCA is a useful statistical method, it is very sensitive to outliers. To resist sensitivity toward these anomalous data, Hubert and Engelen (4) introduced a robust PCA (ROBPCA) method. ROBPCA is a method to identify the outliers (influential data points). The main purpose of this technique is to capture as much variability as possible through the resulting hyper-planes. Therefore, it is a very suitable method for analyzing high-dimensional data. ROBPCA applies the projection pursuit approach combined with robust covariance estimations (4). The robust estimators called minimum covariance determinant (MCD). MCD calculates the distance between each observation and the hyperplane. ROBPCA aims to minimize this distance to explain more variability and make the PCA more robust.

ROBPCA flags four types of observations: regular, good leverage, orthogonal, and bad leverage points (4). As it is shown in Figure A 1, the regular observations (black points) are located on the PC hyperplane and they are relatively close together with no extreme observation. Good leverage observations (G1 and G2) are located on the hyperplane but horizontally they are located far from the center of regular observations. Orthogonal outliers (O1 and O2) are located outside of the PC hyperplane with a large vertical distance, however, their projection will place among the regular observations. Bad leverage points (B1 and B2) are also located outside of the PC hyperplane with large vertical distances. Also, their projection will have a large horizontal distance from the center of regular observations.

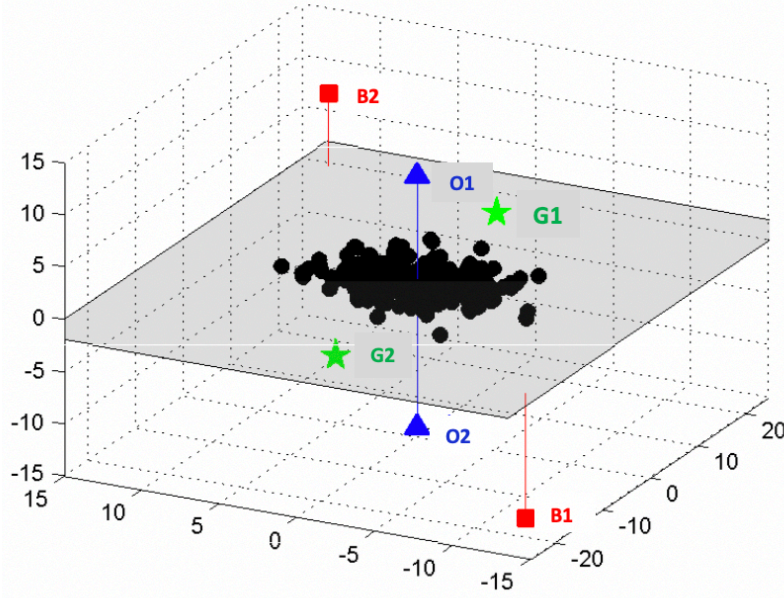


Figure A 1. Different types of observations.

To measure the outlyingness two types of distances can be calculated by ROBPCA; “orthogonal distance” and “robust score distance”. The orthogonal distance measures the distance between a point and its projection on the PC hyperplane and can be calculated as:

$$OD_i = \|x_i - \hat{x}_i\| \quad (1)$$

In equation (1), x_i is the i th observation, $\hat{x}_i$ represents the projection of x_i on the hyperplane, and OD_i denotes the orthogonal distance between x_i and its projection.

According to the ROBPCA, a set of robust scores can be calculated as follow:

$$s_i = P'(x_i - \hat{\mu}) \quad (2)$$

In this equation, s_i denotes the robust scores, P is robust loading matrix, and $\hat{\mu}$ represents the robust center or center of the data. Further, the robust score distance (SD_i) which measures the horizontal distance of observations from the center of the principal hyperplane and can be calculated as follow (e_i denotes to the eigenvalues of MCD matrix):

$$SD_i = \sqrt{\sum_{j=1}^k \frac{s_{ij}^2}{e_i}} \quad (3)$$

Bootstrapping Results

Table A 1. Bootstrapping results for 50 states of the U.S.

	State	Meta-frontier Efficiency Scores	Bias-Corrected Efficiency Scores	90% Confidence Intervals	
1	Alabama	0.873	0.825	0.794	0.859
2	Alaska	0.903	0.809	0.743	0.885
3	Arizona	0.974	0.888	0.827	0.958
4	Arkansas	0.920	0.859	0.818	0.895
5	California	1.000	0.869	0.782	0.978
6	Colorado	1.000	0.893	0.825	0.956
7	Connecticut	0.824	0.742	0.690	0.804
8	Delaware	0.831	0.764	0.729	0.802
9	Florida	0.857	0.776	0.715	0.826
10	Georgia	0.913	0.839	0.783	0.904
11	Hawaii	0.941	0.867	0.813	0.912
12	Idaho	1.000	0.872	0.782	0.977
13	Illinois	0.754	0.694	0.657	0.735
14	Indiana	0.807	0.738	0.695	0.776
15	Iowa	1.000	0.915	0.851	0.977
16	Kansas	0.845	0.788	0.755	0.822
17	Kentucky	0.819	0.775	0.747	0.803
18	Louisiana	0.765	0.712	0.670	0.764
19	Maine	0.759	0.686	0.633	0.736
20	Maryland	0.841	0.783	0.740	0.829
21	Massachusetts	1.000	0.857	0.767	0.955
22	Michigan	0.772	0.718	0.691	0.748
23	Minnesota	0.775	0.701	0.650	0.744
24	Mississippi	0.918	0.839	0.789	0.890
25	Missouri	0.789	0.748	0.719	0.773
26	Montana	0.896	0.819	0.765	0.904
27	Nebraska	0.769	0.705	0.662	0.741
28	Nevada	1.000	0.884	0.809	0.946
29	New Hampshire	0.893	0.799	0.731	0.871
30	New Jersey	0.819	0.732	0.674	0.799
31	New Mexico	0.921	0.858	0.809	0.901

32	New York	0.724	0.654	0.617	0.704
33	North Carolina	0.861	0.802	0.763	0.846
34	North Dakota	0.970	0.861	0.795	0.941
35	Ohio	0.731	0.688	0.658	0.715
36	Oklahoma	0.956	0.874	0.811	0.935
37	Oregon	1.000	0.887	0.833	0.966
38	Pennsylvania	0.700	0.646	0.616	0.675
39	Rhode Island	0.879	0.787	0.720	0.855
40	South Carolina	0.854	0.795	0.752	0.826
41	South Dakota	0.778	0.709	0.657	0.760
42	Tennessee	0.819	0.766	0.728	0.812
43	Texas	1.000	0.901	0.837	0.979
44	Utah	1.000	0.883	0.799	0.961
45	Vermont	1.000	0.864	0.781	0.982
46	Virginia	0.792	0.722	0.668	0.772
47	Washington	1.000	0.866	0.775	0.988
48	West Virginia	0.778	0.723	0.680	0.765
49	Wisconsin	0.860	0.793	0.750	0.833
50	Wyoming	1.000	0.871	0.782	1.003

Slack Analysis Results

Table A 2. Slack analysis results for 50 states.

State	Input target					% of input slacks				
	X1	X2	X3	X4	X5	X1	X2	X3	X4	X5
Alabama	0.088	1.999	2.227	6353.672	56.146	-17.043	-12.634	-39.021	-12.712	-43.373
Alaska	0.087	2.140	2.492	7549.531	99.069	-14.134	-9.701	-9.732	-31.771	-21.310
Arizona	0.093	2.077	2.200	6281.249	48.725	-6.313	-10.281	-3.829	-2.646	-2.654
Arkansas	0.074	1.775	2.334	6813.390	82.969	-11.342	-18.853	-38.486	-8.027	-46.677
California	0.120	2.850	2.180	7550.000	19.780	0.000	-0.161	0.167	0.000	0.015
Colorado	0.120	2.780	2.280	6805.000	64.130	-2.797	-0.028	-0.012	0.000	-0.004
Connecticut	0.123	2.880	2.109	7675.137	38.318	-36.763	-27.688	-17.635	-22.159	-17.608
Delaware	0.083	2.054	2.227	6955.338	76.027	-18.429	-18.362	-16.777	-32.170	-16.891
Florida	0.086	1.996	2.464	6498.213	47.279	-72.480	-76.399	-60.735	-45.594	45.334
Georgia	0.099	2.164	2.148	6014.675	36.367	-4.278	-17.212	-26.695	-25.524	-34.117
Hawaii	0.111	2.522	2.070	6868.388	51.233	-7.327	6.282	-26.405	4.272	2.882
Idaho	0.070	1.720	2.360	6926.000	89.540	-49.749	-45.397	7.301	-5.136	64.410
Illinois	0.104	2.274	2.111	6228.398	41.450	38.371	32.324	-10.397	-10.072	-53.705
Indiana	0.077	1.832	2.307	6697.000	76.178	-42.850	-40.947	-17.640	-18.923	14.088
Iowa	0.070	1.720	2.360	6926.000	89.540	-28.125	-24.073	-22.135	-16.554	-33.930
Kansas	0.085	1.943	2.253	6468.628	62.854	27.400	1.168	-29.164	-21.114	-49.383
Kentucky	0.082	1.902	2.273	6552.599	67.753	-14.076	-20.496	-42.776	-14.367	-34.335
Louisiana	0.100	2.180	2.140	5982.000	34.460	2.234	-9.331	-40.833	-25.262	-81.459
Maine	0.083	2.146	2.201	7233.224	82.268	-42.951	-25.951	-42.903	-7.456	13.136
Maryland	0.117	2.685	2.070	7237.409	49.901	6.980	-6.842	-28.661	-24.073	-62.142
Massachusetts	0.150	5.100	2.990	10559.000	47.780	-7.723	17.947	21.590	22.751	-19.440
Michigan	0.093	2.066	2.194	6215.394	48.078	-38.473	-59.447	-26.651	-41.137	0.628
Minnesota	0.093	2.259	2.177	6873.259	66.559	-25.505	-19.920	-23.977	-14.681	-44.600
Mississippi	0.072	1.753	2.403	6892.427	83.295	-37.975	-45.029	-14.505	-22.313	-4.963
Missouri	0.087	1.978	2.237	6397.536	58.705	-17.387	3.362	-53.616	-16.339	-35.308
Montana	0.072	1.725	2.448	7002.289	93.643	-37.103	-35.632	-30.468	-13.637	-38.800
Nebraska	0.085	1.943	2.253	6467.687	62.799	0.183	-10.672	-46.614	-21.308	-59.510
Nevada	0.080	1.930	2.630	6713.000	49.920	-28.706	-26.015	-33.168	-20.207	-48.192
New Hampshire	0.102	2.741	2.080	8019.812	92.206	21.089	42.040	-20.758	19.467	84.701
New Jersey	0.110	2.604	2.207	7254.399	33.607	-18.659	-15.141	-5.297	-24.355	-71.861

New Mexico	0.092	2.182	2.164	6643.632	58.850	-36.450	-31.359	-22.929	-25.007	43.408
New York	0.116	2.710	2.197	7080.163	33.388	16.968	8.075	-6.604	-1.869	-50.497
North Carolina	0.103	2.281	2.110	6255.035	42.009	-37.814	-44.772	-33.995	-36.030	-8.902
North Dakota	0.078	1.984	2.264	7176.433	87.267	-34.429	-25.810	-7.673	-1.205	0.526
Ohio	0.088	1.992	2.230	6368.208	56.994	12.699	-17.277	-54.746	-35.348	-55.004
Oklahoma	0.070	1.720	2.360	6926.000	89.540	-40.000	-42.282	-27.235	-20.491	-35.881
Oregon	0.130	2.970	1.900	8044.000	92.960	65.605	64.730	-45.310	5.454	-17.905
Pennsylvania	0.084	1.935	2.257	6484.057	63.754	-34.219	-34.866	18.509	-19.393	-31.419
Rhode Island	0.130	3.213	2.196	8393.001	69.981	9.214	0.067	-34.952	-9.343	-43.029
South Carolina	0.092	2.051	2.202	6247.185	49.933	-47.424	-50.333	-18.869	-34.578	-37.286
South Dakota	0.070	1.720	2.360	6926.000	89.540	-37.764	-28.323	-15.207	-5.266	5.330
Tennessee	0.098	2.153	2.153	6036.923	37.665	14.912	-8.963	-60.223	-32.412	-70.741
Texas	0.110	2.330	2.850	6998.000	30.090	-11.425	-15.698	-14.269	-5.086	-77.780
Utah	0.100	2.180	2.140	5982.000	34.460	-10.236	-6.334	-24.813	-14.518	14.522
Vermont	0.160	4.060	2.290	10192.000	165.520	55.476	86.425	7.200	70.378	380.394
Virginia	0.100	2.180	2.140	5982.000	34.460	-38.159	-46.266	-6.725	-41.307	-79.180
Washington	0.100	2.760	1.980	7913.000	80.580	-22.113	-3.123	-28.299	4.711	20.595
West Virginia	0.070	1.720	2.360	6926.000	89.540	-30.019	-37.657	19.285	-12.473	11.115
Wisconsin	0.086	2.274	2.158	7451.297	84.771	-8.481	-5.921	-52.930	-21.259	-66.417
Wyoming	0.100	1.820	3.970	8321.000	164.570	-0.195	-32.538	58.250	-4.400	41.873

Appendix B Pairwise Regression

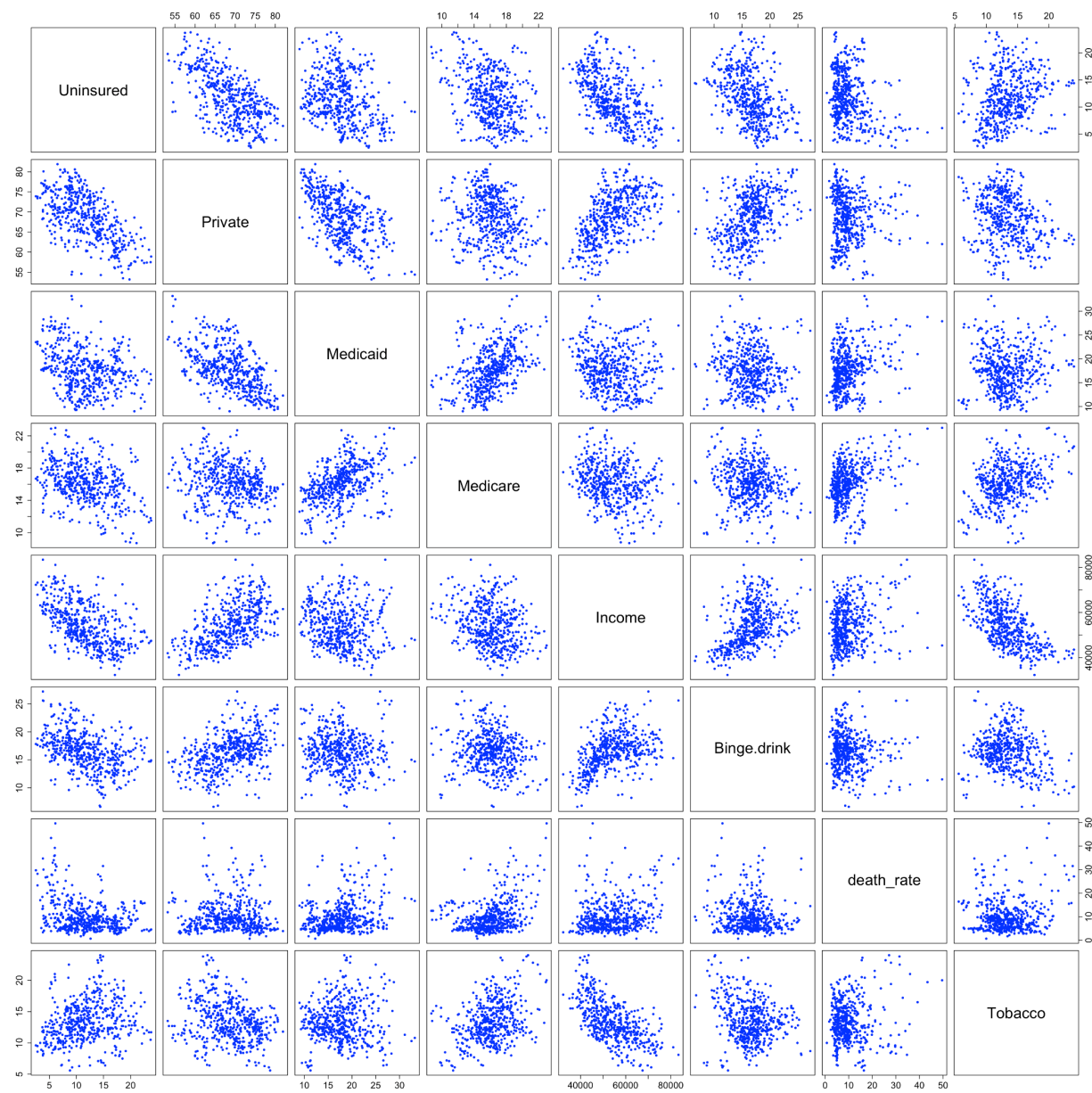


Figure B 1. Pairwise regression among all variables.

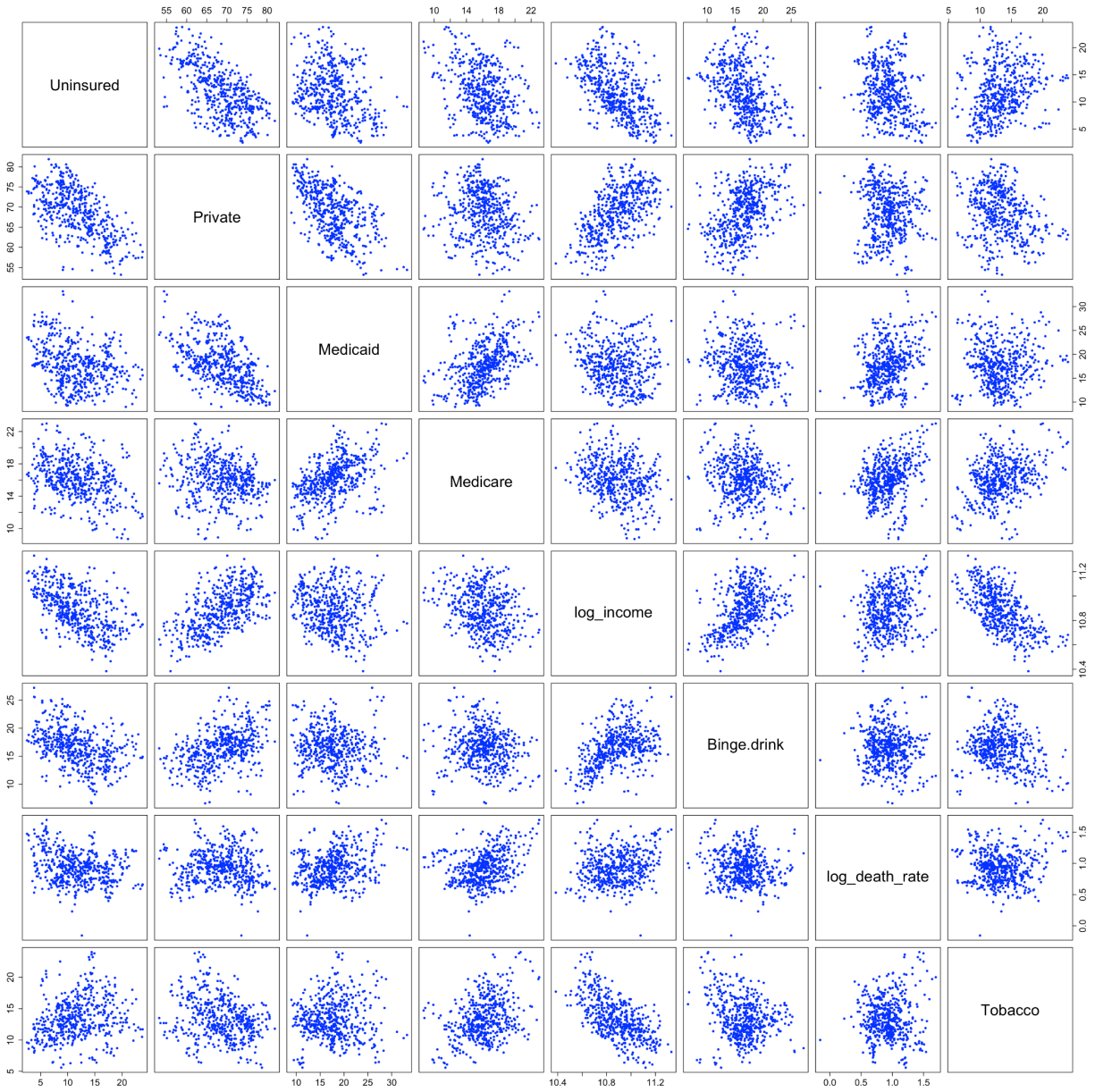


Figure B 2. Pairwise regression among all variables after log transformations.

Appendix C Correlation Matrix

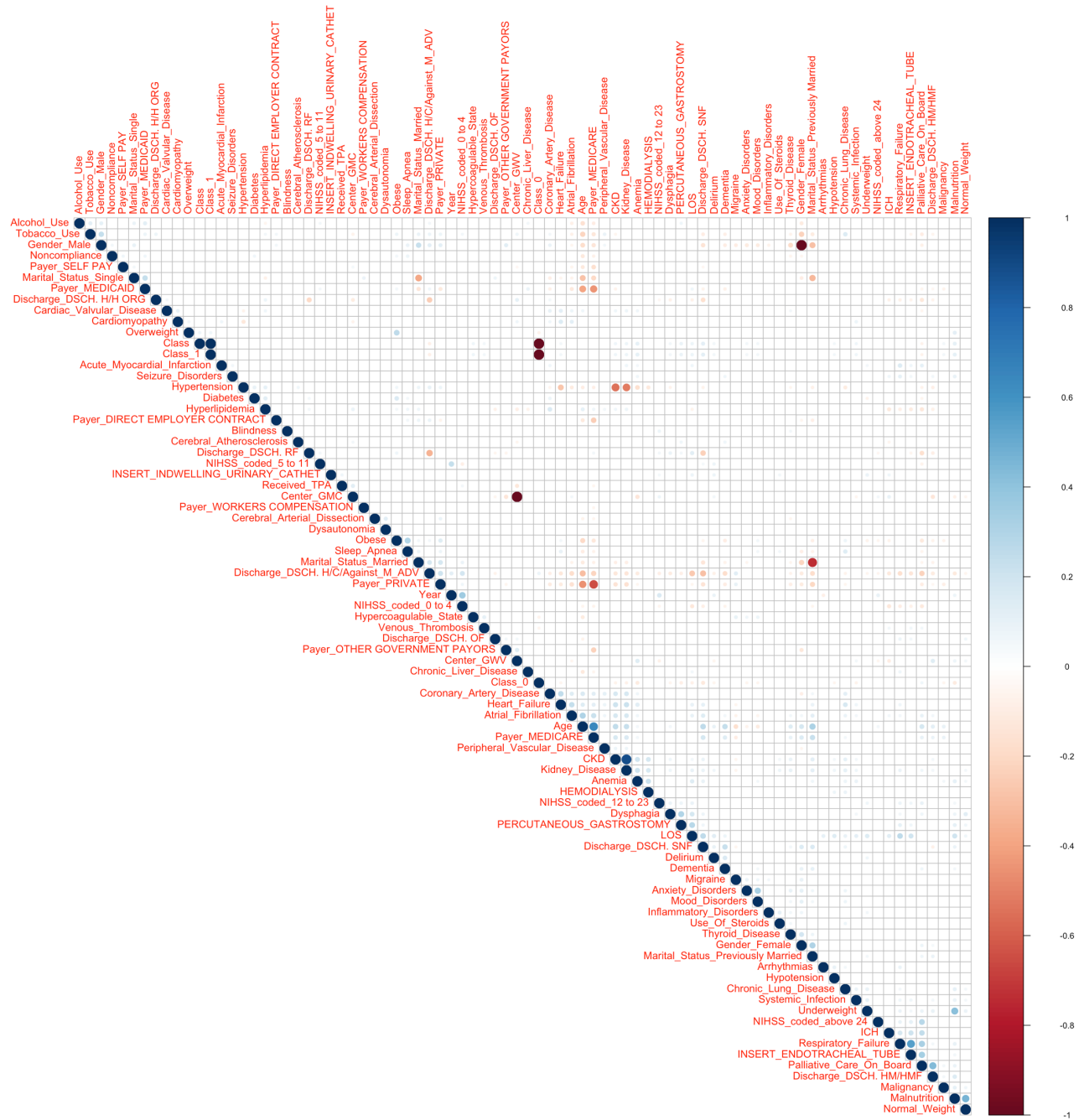


Figure C 1. Correlation matrix between all the variables.

Table C 1. Correlations above 30% and 50% between independent variables.

Variables	Correlations 50% or more	Correlations 30% or more
Age & Medicare	0.64	0.64
CKD & Hypertension	-0.55	-0.55
CKD & Kidney Disease	0.88	0.88
Hypertension & Kidney Disease	-0.51	-0.51
Respiratory Failure & Insert Endotracheal Tube	0.54	0.54
Female & Male	-0.99	-0.99
Married & Previously Married	-0.74	-0.74
Medicare & Private	-0.64	-0.64
GMC & MWV	-1	-1
Age & Atrial Fibrillation		0.31
Age & Previously Married		0.32
Age & DISCHARGED TO HOME, COURT, OR AGAINST MEDICAL ADVICE		-0.3
Age & Medicaid		-0.4
Age & Private		-0.45
Anxiety Disorders & Mood Disorder		0.32
Malnutrition & Normal Weight		0.47
Malnutrition & Underweight		0.41
Palliative Care on Board & Respiratory Failure		0.32
Palliative Care on Board & Insert Endotracheal Tube		0.33
Palliative Care on Board & Discharged to Hospice-Home/Hospice- Medical Facility		0.43
Year & NIHSS 0 to 4		0.36
Female & Previously Married		0.3
Married & Single		-0.39
Medicaid & Medicare		-0.45
Discharged to Home Court or Against Medical Advice & Discharged/Transferred to SNF		-0.3
Discharged to Home Court or Against Medical Advice & Dschrgd/Tnsfrd To Another Rehab Facility		-0.35

References

1. Simar L, Wilson PW. Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics*. 2007;136(1):31-64.
2. Simar L. Estimating efficiencies from frontier models with panel data: a comparison of parametric, non-parametric and semi-parametric methods with bootstrapping. *International Applications of Productivity and Efficiency Analysis*: Springer; 1992. p. 167-99.
3. Ross SM. *Introduction to probability and statistics for engineers and scientists*: Academic Press; 2014.
4. Hubert M, Engelen S. Robust PCA and classification in biosciences. *Bioinformatics*. 2004;20(11):1728-36.