DSDHT Project: Dialysis Treatment Adequacy across the United States

Amruta Chavan
MS in Data Science at
Indiana University of Bloomington
amchavan@iu.edu

ABSTRACT
This project involves analyzing patient database across all the facilities in the US and discovering success rate of dialysis treatment. The goal is to study patient characteristics, treatment patterns and patient outcomes for chronic dialysis patients. The research will also include trends of type of dialysis treatment opted by the patients nation-wide, compare the success rates and scores of clinics offering the procedure region-wise and overall scores. By identifying clinics that provide better treatment procedure we can improve the success rate region-wise and prolong the patientâĂŹs survival rate.

1. INTRODUCTION
Kidney disease often has no symptoms in its early stages and can go undetected until it is very advanced. Hence, kidney disease is also referred to as a 'silent disease'. The adjusted incidence rate of ESRD(end-stage renal disease) in the United States rose sharply in the 1980s and 1990s, leveled off in the early 2000s, and has declined slightly since its peak in 2006. Each year, kidney disease kills more people than breast or prostate cancer. In 2013, more than 47,000 Americans died from kidney disease.
The overall prevalence of CKD in the general population is approximately 14 percent. High blood pressure and diabetes are the main causes of CKD. Almost half of individuals with CKD also have diabetes and/or self-reported cardiovascular disease (CVD).More than 661,000 Americans have kidney failure. Of these, 468,000 individuals are on dialysis, and roughly 193,000 live with a functioning kidney transplant.[1]

2. BACKGROUND
Dialysis is a process for removing waste and excess water from the blood and is used primarily as an artificial replacement for lost kidney function in people with kidney failure either until suitable transplant is made available or in case of acute kidney failure.Dialysis care is intricate and multiple factors may influence patient survival. As a result finding good facility and treatment measures becomes crucial for patients to avoid further complications and/or prolong survival rate.Also the cost for providing care for patients on hemo-dialysis due to end stage kidney disease is high. Finding ways to improve patient outcomes and reduce the cost of dialysis is important. To ensure that proper care is offered nation-wide an annual survey and analysis is necessary.To ensure that proper care is offered nation-wide an annual survey and analysis is necessary. This could help improve the quality of treatment provided at the clinics.[2]

3. OBJECTIVE
The goal is to study patient characteristics, treatment patterns and patient outcomes for chronic dialysis patients. The research will also include trends of type of dialysis treatment opted by the patients nation-wide, compare the success rates and scores of clinics offering the procedure region-wise and overall scores. By identifying clinics that provide better treatment procedure we can improve the success rate region-wise and prolong the patient's survival rate.
This report will compare the characteristics of a facility's patients, patterns of treatment, and patterns in transplantation, hospitalization, and mortality to local and national averages. Such comparisons help you to evaluate patient outcomes and to account for important differences in the patient mix - including age, sex, race, and patients' diabetic status - which in turn enhances each facility's understanding of the clinical experience relative to other facilities in the state, Network, and nation.

4. DATA SETS
This project uses data set published by University of Michigan Kidney Epidemiology and Cost Center (UM-KECC) alongwith the Centers for Medicare and Medicaid Services (CMS) and is based primarily on Medicare claims and data collected for CMS.[3][4]

5. ABOUT THE DATA
This dataset includes summaries of patient characteristics, treatment patterns, and patient outcomes for chronic dialysis patients who were treated in this facility between January 2011 and December 2014. Mortality, hospitalization, and transplantation statistics are reported.

Regional and national averages were calculated for comparisons. Several count of patient mortality, hospitalization, and transplantation are adjusted to account for the characteristics of the patient mix at these facilities, such as age, sex and diabetes as a cause of ESRD. Data includes patients that received either hemodialysis (HD) and peritoneal dialysis (PD) combined.

6. TECHNOLOGIES
Following technologies were used in the project

6.1 MySQL
For combining datasets.

6.2 R 3.3.4

For Data Cleaning/Wrangling, Exploratory Data Analysis and modeling data.

6.3 Tableau

Visualizing the graphs we generated via the exploratory data analysis phase using Tableau tool.

7. DATA WRANGLING

The dataset was processed and cleaned for futher analysis. The AnnualRecords.csv and FacilityDetails.csv datasets were consolidated into one single dataset for easier analysis based on the prov ID column using R. Upon processing the data for missing and NA values, dataset consisting of 210 features and 6337 instances was obtained.

8. DESCRIPTIVE ANALYSIS AND DATA TRENDS

The dataset was extensively analysed to study 3 types of trends in the data set namely demographic variability in the data, overall patient characteristic trend (for eg: age,ethnicity etc) and trends in treatment across the facilities to derive conclusions.Yearly variability analysis deals with year-to-year of various factors leading to raise/drop in dialysis treatment. Overall Analysis of patient data will help find definite correlations between factors causing rise or fall in the patient count across the facilities whereas studying demographic aspect of dialysis treatment will be understand the quality of treatment provided at centers across the country.

9. DATA VISUALIZATION

9.1 Pre-Analysis Data Plots

Preliminary analysis graphs describe the treatment and quality of the dialysis procedure in every state across the US.

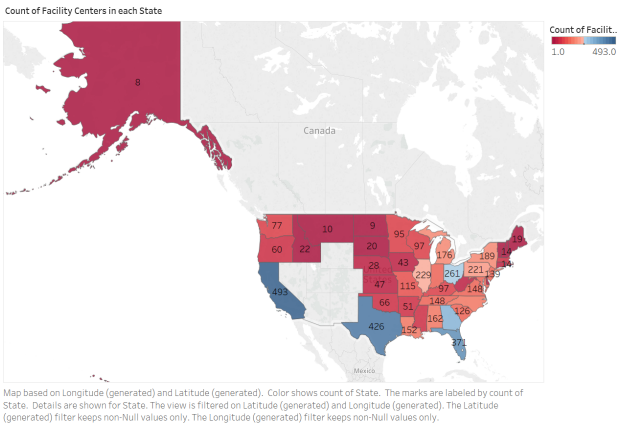

Figure 1: No.of Dialysis Centers across the US

Figure 1 shows total no. of dialysis centers in each state across the US. As seen from the plot, states of California,Missouri,Florida,Georgia and Illinois have more number of dialysis treatment centers available as compared to rest of the States. One of the reasons could be due to urbanization and/or awareness in those states. More centers could also mean proportionately more no. of patients with renal disorder in those states.

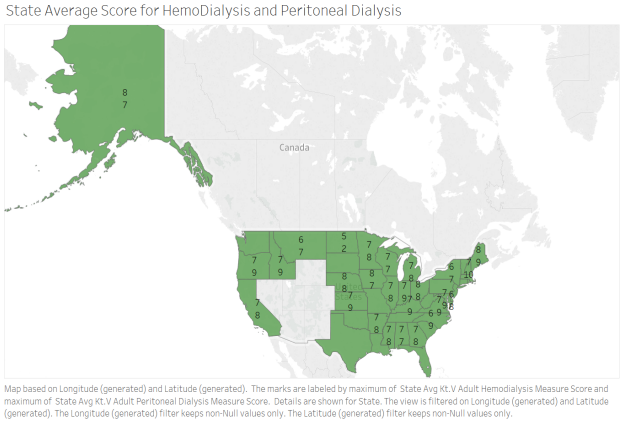

Figure 2: State Avg. Score for Dialysis Treatment

Figure 2 shows average score of dialysis treatment centers in each state. The score for both hemo-dialysis (top score) and peritoneal dialysis(bottom score) are listed in the graph show.

9.2 Patient Analysis

To estimate count of patients with renal disorder that facility centers could possibly expect annually is calculated using a Cox model, adjusting for patient age, sex, diabetes, duration of ESRD, nursing home status, patient comorbidities at incidence, body mass index (BMI) at incidence, and calendar year. Duration of ESRD is divided into six intervals with cut points at 6 months, 1 year, 2 years, 3 years and 5 years and hospitalization rates are estimated separately within each interval. For each patient, the time at risk in each ESRD interval is multiplied by the (adjusted) national admissions rate for that interval, and a sum over the intervals gives the expected number of admissions for each patient.

This expected count of patients was compared with actual count of patients admitted annually over the span of two years 2013-2014 as shown in the figure below, the expected count of patients was fairly accurate to the no. of patients admitted every year during 2013 and 2014.

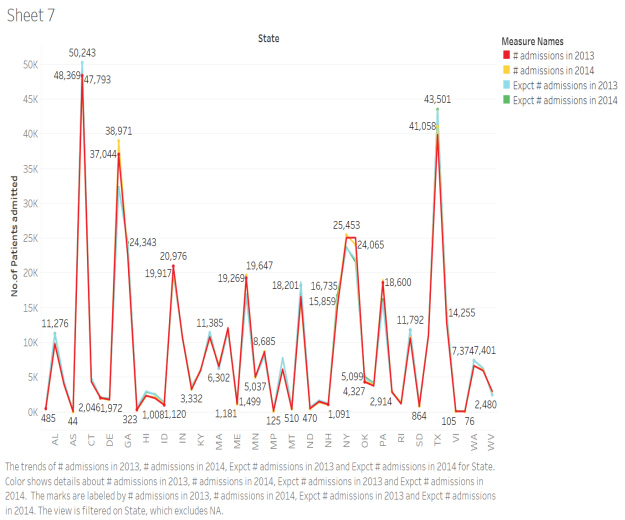

Figure 3: Count of Patients admitted annually

The dialysis data was then analysed to study patients based on gender,ethnicity,age etc.

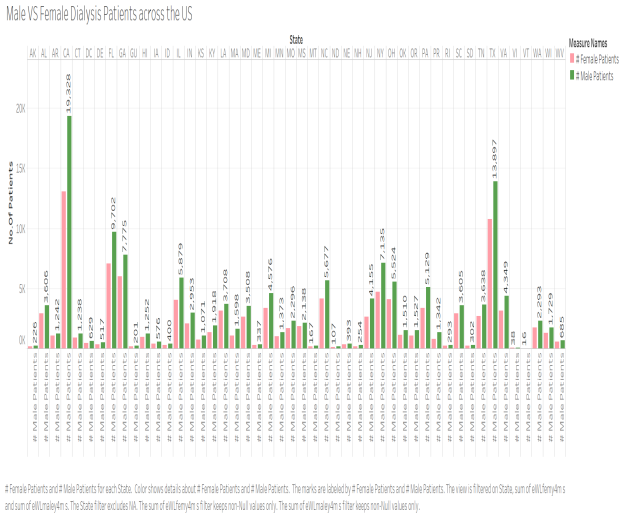

Figure 4: Male VS Female Dialysis Patient Count across US

Figure 4 shows male versus female count of dialysis patients across the country.Analysis shows no. of female dialysis patients is less than the males in all of the states of US. States of California,Texas,New York,Ohio have significantly large no. of male dialysis patients compared to female.

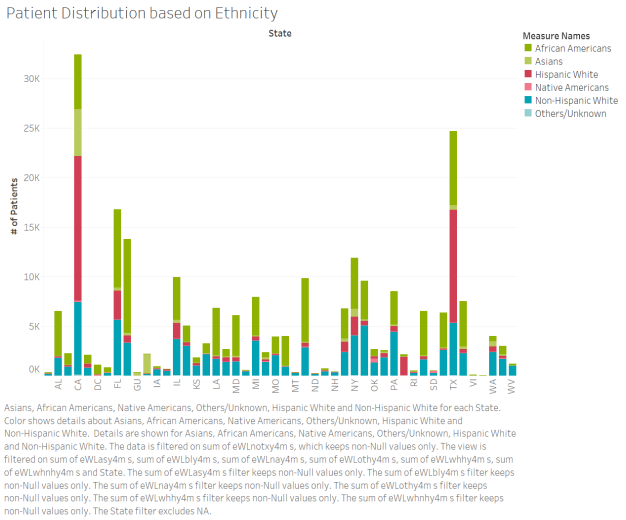

Figure 5: Patient distribution based on Ethnicity

Figure 5 shows proportion of patients from various ethnic background in every states that receive dialysis treatment at the facility centers. Kidney failure disorders appeared to be more common amongst African Americans compared to rest of the patients.

9.3 Mortality

The graphs below study the mortality across different states and analyzes the cause and nature of death in the dialysis treatment patients.

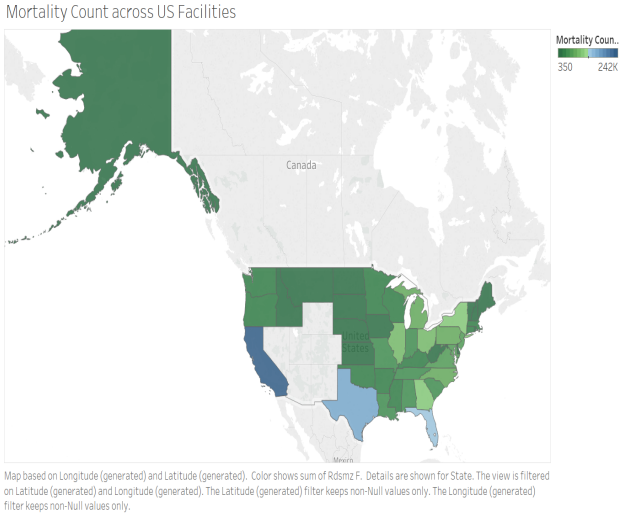

Figure 6: Mortality Count across US

Figure 6 shows overall mortality count across the US on an average every year. States of California and Texas have higher mortality rate compared to rest of the states.

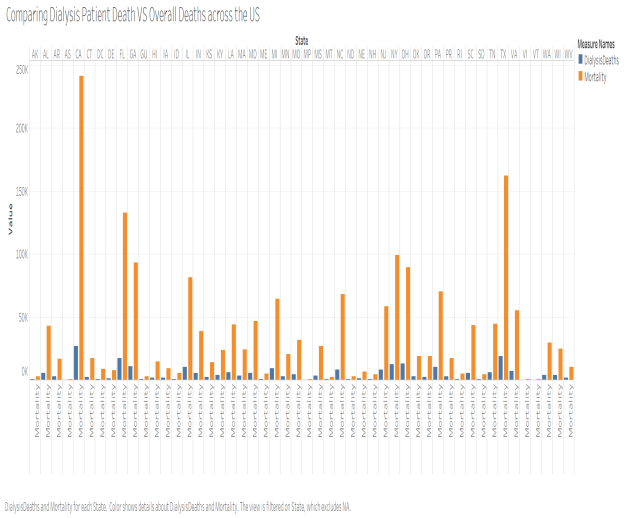

Figure 7: Dialysis Patient deaths VS Overall Death count

Figure 7 compares deaths of dialysis patients with overall death count in every state.State such as Missouri,Maryland,Florida,LosAngeles,Arizona have proportionately large no. of deaths of dialysis patients. On an average approximately 20-30 percent of overall death count consisted of dialysis patients.

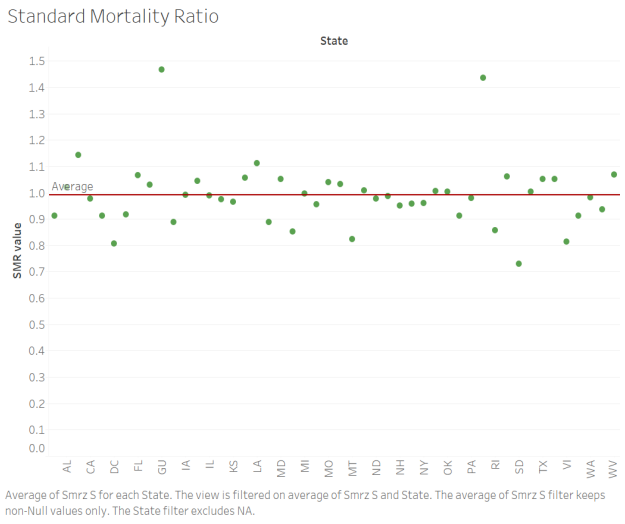

Figure 8: Standard Mortality Ratio Value across US

Figure 8 shows the average SMR value in each state across the US. Standardize Mortality Ratio is ratio of observed deaths over expected death count.Ideally SMR value should be less than 1.00. Value 1.00 represents observed death count equal to expected death count.SMR value for states namely Arizona,Porto Rico,Guam and LA are exceptionally higher than 1.00. Porto Rico and Guam showed approximately 4.5 percent higher mortality count than expected count.

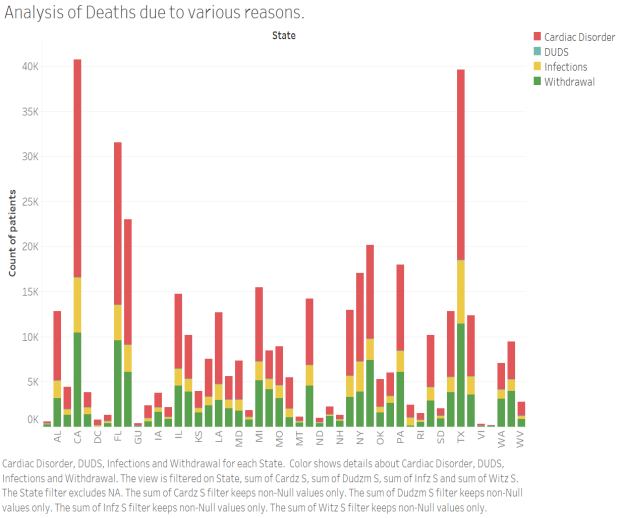

Figure 9: Causes of Death in Dialysis Patients

Figure 9 studies various causes of death in dialysis patients in different parts of country namely cardiac disorder,diabetes,infections or death due to/after withdrawal from dialysis treatment. Cardiac disorder and deaths due to withdrawal from treatment seem to be major cause of death in the patients with renal disorder. The data is sufficient enough to prove that renal failure could lead to cardiac disorders in the patients.

9.4 Duration of Dialysis treatment

Many people chose to opt out of dialysis treatment for various reasons such as lack of improvement, affordability, death etc. The duration of treatment for dialysis of patients was recorded with the cause of terminating the treatment. Three significant durations recorded were for duration of hospitalization until opting out,death of patient during the treatment and dialysis treatment duration in infants.

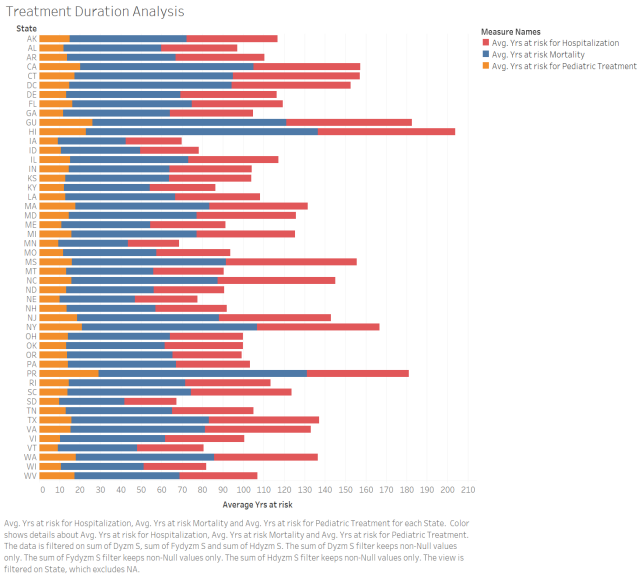

Figure 10: Duration of Hospitalization

Figure 10 shows that patients generally continued to receive dialysis until the very last moment across the country. The patients logged under the hospitalization category do not sufficient describe the reason to opt out of the treatment. The reason for withdrawal are unknown.

9.5 Transplant Treatment

The data of patients under age 70 that opted to be put on waitlist for renal transplant during their dialysis treatment were studied for the options that they chose. The graph below shows waitlisted patients with various severity of ESRD.

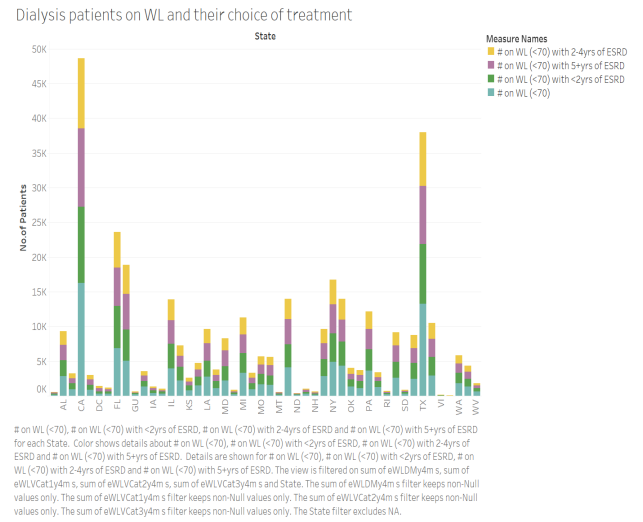

Figure 11: Causes of Death in Dialysis Patients

The results of numerous studies have indicated that the recipients of renal transplants have better survival than comparable dialysis patients. The first step in the transplant process is getting placed on the transplant waitlist.Figure 6 shows waitlisted patients for transplant.This waitlist statistics refers only to those patients less than 70 years of age because transplants in people aged 70 or greater occur with much less frequency than do transplants in younger patients.

10. HIGHLIGHTS

Few significant revelations that were made while analysis dialysis patient data across US were as follows:

- State of California,Texas,Ohio,Florida and Illinois recorded maximum number of dialysis treatment facility centers compared to rest of the states.
- SMR value for Porto Rico and Guam were higher than the expected count.
- Urbanized and/or highly populated states such as California , Los Angeles, New York, Ohio, Florida shows comparatively higher count of patients with renal failure.
- Males were observed to possess higher risk of renal failure compared to females across the country.
- The count of patients belonging to African American ethnicity was recorded higher than rest of the ethnicities.

11. CONCLUSION

The report compared the characteristics of a facilityâĂŹs patients, patterns of treatment, and patterns in transplantation, hospitalization, and mortality to local and national averages to evaluate patient outcomes and to account for important differences in the patient mix âĂŤ including age, sex, race, and patientsâĂŹ diabetic status âĂŤ which in turn enhances each facilityâĂŹs understanding of the clinical experience relative to other facilities in the state, and nation.

12. ACKNOLEDGEMENTS

I would like to express our sincere gratitude to Prof. Joanne Luciano, Prof. Jeremy Yang and the course TA, John Wolohan for their constant guidance and suggestions throughout the project completion process.

13. REFERENCES

[1] DaVita.Inc, “Kidneydisorder,” *DaVitaWebsite*, daVita.Inc. [Online]. Available: https://www.davita.com/kidney-disease/dialysis/the-basics

[2] AmericanKidneyFund, “Renaldisorder.” [Online]. Available: http://www.kidneyfund.org/kidney-disease/kidney-failure/?referrer=https://www.google.com/

[3] “Data link.” [Online]. Available: https://dialysisdata.org/content/dialysis-facility-report-data

[4] [Online]. Available: https://www.healthdata.gov/dataset/esrd-qip-dialysis-adequacy-payment-year-2016