Regional Variation in Health Care Utilization in Diabetes

Introduction

Previous studies have documented substantial variation in health care spending, quality, and utilization across the United States (Fisher et al., 2003). For instance, the Dartmouth Atlas of Health Care reports more than a two-fold variation in per capita Medicare spending in different regions of the county, even after adjusting for price (Skinner and Fisher, 2010). Regional variation in health care usage can arise for a number of reasons, including differences in the health of the population (i.e., comorbidities and health needs), overuse of marginally effective medical services such as unnecessary screening or diagnostic testing, or differences in the price of services. But it can also arise from individuals not getting needed or necessary care. This latter possibility is supported by evidence of strong gradients across the United States in health, including variations in health care utilization, cost and quality across and within states (Gittelsohn, A., & Powe, N. R. 1995).

Using the Andersen behavioral model of health service use, we can see that there are three main categories that explain why people use or do not use health services (Andersen, 1995). The first category is called predisposing characteristics. Andersen describes this as the demographic characteristics of age and sex as "biological imperatives", social factors, and mental factors in regards to health beliefs and knowledge related to health that people may have at the individual level. Enabling factors, the second category of the theory, includes factors dealing with financial and organizational factors that affect utilization of health services. Andersen describes that individual financial factors include income/wealth and health insurance status of a person. Regarding organizational factors, these are described as factors that entail whether the person has a regular source of care and the nature of that source. Organizational factors also include characteristics of health care services facilities and personnel in the area. The third category, need factors, is composed of two areas. Andersen differentiates these two areas of perceived need and evaluated need. Perceived need for health related services is described as how people view and experience their own health, illness symptoms, and functional state. Evaluated need is individual need which is assessed by a professional in health care for a patients' health status and medical need for care (Babitsch, Gohl, & Lengerke, 2012).

These three categories proposed by Andersen seem to cover all of the reasons for regional variation in health care utilization that were found in the reviewed literature. Under the first category, predisposing factors, one factor that was found in the literature was differences in hospital capacities. What this study found was that if a hospital had greater capacity then there was substantially increased use of that hospital, with no benefit to mortality rates (Fisher et al., 2000). In a systematic literature review, environmental variables, such as characteristics of the health care delivery system, external environment, and community, were found in nearly fifty percent of studies to be causes of variation in utilization of health services (Phillips, Morrison, Andersen & Aday, 1998). In this same literature review, provider related variables, such as patient factors that may be influenced by providers, were found in over fifty-percent of the reviewed studies. For example, whether individuals have a regular source of care influences utilization. Regions with high minority populations was found to be a reason for variation. In a study by Weinick, Zuvekas, and Cohen (2000) researchers found that disparities increased from 1977 to 1996 particularly for Hispanic Americans. They also found that about fifty to seventy-five percent of observed disparities would exist even if racial and ethnic disparities in income and health insurance coverage were eliminated. Multiple studies found that demography is a factor that explains why regional variation exists. Gittelsohn and Powe (1995) found that hospital use is directly related to several factors, one of which is demography. They explain that uniform delivery of health care rarely holds because these factors, such as demography, vary and are heterogeneous. Some other predisposing factors that were found in the literature were individual SES, health related behaviors, population size, social class, poverty and unemployment. In a study that examined geographic variation of dental utilization in low income children, researchers found that rural vs urban was a factor effecting utilization of health care (McKernan et al., 2015). This is due to how easily accessible health care is in the region, where people who live in rural areas are usually much further away from health care centers. Level of education is a factor that was found in some studies. A study by Parslow, Jorm, Christensen and Jacomb (2002) showed that people who visited a general practitioner were more likely to be educated. Most studies seem to show that people with higher education levels utilize health care more, however one study found that individuals with less than a high school education actually utilized health care for treatment of mental problems more than those with some college education (Dhingra, Zack, Pearson & Balluz, 2010). Age was also a factor that was found in several of the studies to influence health care utilization. Certain types of health care are used more by different age groups. For example, elderly people were found to not use drug or mental health services nearly as often as younger individuals (Stockdale, Tang, Zhang, Belin & Wells, 2007). In a study by Fonseca, Otahal, Wiggins and Marwick (2015), researchers examined growth and geographic variations in the use of cardiac imaging in Australia. Interestingly, they found that one factor that influenced regional use of testing was associated with the proportion of females. Females utilized cardiac imaging more than males in this study. Another predisposing factor that was found in the literature to possibly effect usage of health care was cultural differences. In a study, researchers examined the determinates of regional differences in the utilization of ambulatory services in Germany and found that there was a proportion of regional differences that remained unexplained (Kopetsch & Schmitz, 2014). They concluded that the unexplained area could be due to cultural differences incapable of being measured between patients or doctors.

The second category regarding factors that influence usage of health care from the Andersen model is enabling factors. Enabling factors consists of two sub categories, financing factors and organizational factors. Under financing factors, health insurance and income were found in the literature to influence usage of health services. If people do not have health insurance, obviously they would be less likely to use health services. Use also varies by the type of health insurance, as deductibles, copayments, and coinsurance rates all vary. Income is somewhat related to this, as using health care services is not always cheap financially. Higher income people tend to utilize services more, with the exception of drug and mental health care services (Dhingra, Zack, Strine, Pearson & Balluz, 2010). Lower income people are more likely to struggle with drug and alcohol addiction which explains why they would use these services more often. Regarding organizational factors, several factors from the literature can be placed under this sub category. Higher physician density as well as lower patient travel times were both explanations for variation in usage of cancer screening in Germany (Vogt, Siegal & Sundmacher, 2014). Medical resources was also found in other studies to influence usage of health care. One study concluded that if a region has more hospital beds, then usage will go up (Camenzind, 2012). This concept is a familiar one, as it has been proven that building another hospital in an area because the current hospital is at maximum capacity will not change the usage of the old hospital, but rather completely fill up both the new and old hospital. Other studies have found that differences in service availability is a factor as well. Also differences in clinical training was found to be a factor in a study that examined regional variations in lumbar spine surgery (Weinstein et al., 2006). All of these enabling factors explain why regional variation exists in the usage of health care.

The third category, need factors, proposed by Andersen in his model is broken into the two categories perceived need and evaluated need. Many factors found in the literature can be classified as perceived needs. Morbidity, mortality, individual health status, and individual health behavior are all factors that different studies found to explain variation. A study that explored regional variations in health care and inequalities in the provision of hip and knee replacements found that one explanation for variation in these types of surgery could be due to the numbers with limiting long term illness (Dixon, Shaw & Dieppe, 2006). If people are dealing with pain for long periods of time, they are likely to seek care. In areas with more people like this, knee and joint replacement surgery will be more common. Evaluated needs that some of the studies mentioned were provider characteristics that interact with patients (such as physician gender), physician practice patterns, financial incentives, and differences in professional opinion. Physicians will not always be consistent with the way they treat certain conditions. In a study that investigated regional variations in lumbar spine surgery, researchers found that financial incentives and disincentives to surgical interventions was a factor causing variations (Weinstein et al., 2006). Sometimes physicians will avoid performing surgeries, or perform surgeries on certain cases more often because of the financial benefits to them. Also, physicians have different opinions in the way they go about providing care. One physician might always perform surgeries for a torn ACL, while another might opt to put the patient in a brace. These differences come from the provider, and will always be a reason for variation.

There are numerous studies that examine regional variation and causes of variation in health care utilization across the United States, and even in different countries, but there has not been research specifically comparing different counties in California. Examining regional variation in diabetes, we may find that there is a lack of access to needed care. Regional variations in hospital usage specifically for diabetes has not been examined. Diabetes is a growing Public Health concern. In 2015, over 30 million Americans, or 9.4% of the population had diabetes (American Diabetes Association, 2017). A little over 6% of people living in California have been diagnosed with diabetes, and it accounts for over 5.5% of total health care expenditures (Brown et al., 2015).

The challenge of identifying possible factors for regional variations in hospital usage for diabetes across California counties is trying to control for as many of these factors as possible. Using secondary hospital data, we are limited in what we can adjust for. Instead of controlling for all of these possible factors found in the literature, we can use proxy variables to try and explain the reasons for variations. This study aims to determine if there are variations in hospital admissions, average length of stay, and average cost per visit for diabetes across all 58 California counties.

Methods

Data Sources

The study used data from the following sources:

- Diagnoses code for diabetes – A complete list of the codes used can be found in table 1. The codes were obtained from the Centers for Medicare and Medicaid Services (CMS) website.

- Hospital data – Data from the Office of Statewide Health Planning and Development (OSHPD) patient discharge and emergency department data for 2013 to examine regional variations across California counties for health care utilization for diabetes. The OSHPD datasets contain facility level data from more than 6,000 California Department of Public Health (CDPH) licensed healthcare facilities and include financial, utilization, patient characteristics, and service information. Office of Statewide Health and Planning and Development definitions of rural, urban, and frontier were used to determine patient's rurality by ZIP code.
- Cost data – Centers for Medicare and Medicaid Services data was used in order to link Diagnosis Related Group (DRG) information to average national Medicare inpatient payment for discharge data (2014 data), and to link Current Procedural Terminology (CPT) codes to national average payment for a facility (2017 data) in order to estimate costs.
- Census data – Median income estimates by ZIP code for 2010, as well as population estimates were obtained from the Census Bureau.
- Cases of diabetes by county – Prevalence estimates for diabetes for each county were obtained from the California Health Interview Survey. The estimate came from the 2015 survey question, "ever diagnosed with diabetes."
- California county health rankings – 2017 county health rankings for health behaviors, clinical care, social and economic factors, and the physical environment were obtained from the University of Wisconsin Population Health Institute.

Table 1: Complete list of diagnosis codes used

ICD-9	Description
25000	Diabetes mellitus without mention of complication, type II or unspecified type, not stated as uncontrolled
25001	Diabetes mellitus without mention of complication, type I [juvenile type], not stated as uncontrolled
25002	Diabetes mellitus without mention of complication, type II or unspecified type, uncontrolled
25003	Diabetes mellitus without mention of complication, type I [juvenile type], uncontrolled
25010	Diabetes with ketoacidosis, type II or unspecified type, not stated as uncontrolled
25011	Diabetes with ketoacidosis, type I [juvenile type], not stated as uncontrolled
25012	Diabetes with ketoacidosis, type II or unspecified type, uncontrolled
25013	Diabetes with ketoacidosis, type I [juvenile type], uncontrolled
25020	Diabetes with hyperosmolarity, type II or unspecified type, not stated as uncontrolled
25021	Diabetes with hyperosmolarity, type I [juvenile type], not stated as uncontrolled
25022	Diabetes with hyperosmolarity, type II or unspecified type, uncontrolled
25023	Diabetes with hyperosmolarity, type I [juvenile type], uncontrolled
25030	Diabetes with other coma, type II or unspecified type, not stated as uncontrolled
25031	Diabetes with other coma, type I [juvenile type], not stated as uncontrolled
25032	Diabetes with other coma, type II or unspecified type, uncontrolled
25033	Diabetes with other coma, type I [juvenile type], uncontrolled
25040	Diabetes with renal manifestations, type II or unspecified type, not stated as uncontrolled
25041	Diabetes with renal manifestations, type I [juvenile type], not stated as uncontrolled
25042	Diabetes with renal manifestations, type II or unspecified type, uncontrolled

25043	Diabetes with renal manifestations, type I [juvenile type], uncontrolled
25050	Diabetes with ophthalmic manifestations, type II or unspecified type, not stated as uncontrolled
25051	Diabetes with ophthalmic manifestations, type I [juvenile type], not stated as uncontrolled
25052	Diabetes with ophthalmic manifestations, type II or unspecified type, uncontrolled
25053	Diabetes with ophthalmic manifestations, type I [juvenile type], uncontrolled
25060	Diabetes with neurological manifestations, type II or unspecified type, not stated as uncontrolled
25061	Diabetes with neurological manifestations, type I [juvenile type], not stated as uncontrolled
25062	Diabetes with neurological manifestations, type II or unspecified type, uncontrolled
25063	Diabetes with neurological manifestations, type I [juvenile type], uncontrolled
25070	Diabetes with peripheral circulatory disorders, type II or unspecified type, not stated as uncontrolled
25071	Diabetes with peripheral circulatory disorders, type I [juvenile type], not stated as uncontrolled
25072	Diabetes with peripheral circulatory disorders, type II or unspecified type, uncontrolled
25073	Diabetes with peripheral circulatory disorders, type I [juvenile type], uncontrolled
25080	Diabetes with other specified manifestations, type II or unspecified type, not stated as uncontrolled
25081	Diabetes with other specified manifestations, type I [juvenile type], not stated as uncontrolled
25082	Diabetes with other specified manifestations, type II or unspecified type, uncontrolled
25083	Diabetes with other specified manifestations, type I [juvenile type], uncontrolled
25090	Diabetes with unspecified complication, type II or unspecified type, not stated as uncontrolled
25091	Diabetes with unspecified complication, type I [juvenile type], not stated as uncontrolled
25092	Diabetes with unspecified complication, type II or unspecified type, uncontrolled
25093	Diabetes with unspecified complication, type I [juvenile type], uncontrolled

Independent variables

Two levels of variables were used in the subsequent analysis: individual level factors taken from the OSHPD data, and county level factors. Following the categories suggested by the Andersen model, three categories of factors were included in the individual analysis as predictor variables. The predisposing factors in our model were:

- Age – As recoded in the OSHPD. This variable was mean centered.
- Ethnicity – OSHPD reports seven categories: unknown/invalid/blank, White, Black, Hispanic, Asian/Pacific Islander, Native American/Eskimo/Aleut, and other. These were recoded into NH White, NH Black, Hispanic, NH Asian, and NH other.
- Gender – Male or female.

The predisposing factors in our county level model were:

- Age – Average age for diabetes admissions by county. This variable was mean centered.
- NH White – Percentage of diabetes admissions who are NH White by county.
- NH Black – Percentage of diabetes admissions who are NH Black by county.

- NH Asian – Percentage of diabetes admissions who are NH Asian by county.
- Hispanic – Percentage of diabetes admissions who are Hispanic by county.
- Female – Percentage of diabetes admissions who are female by county.
- Social and Economic Factors – County ranking based on education, employment, income, family and social support, and community safety. This variable was mean centered. The best county is ranked 1, meaning the social and economic factors are the best in this county.

The enabling factors in our model were:

- Insurance type – OSHPD reports ten categories in the PDD: Medicare, Medi-Cal, private coverage, workers' compensation, county indigent programs, other government, other indigent, self-pay, other payer, and invalid/blank. OSHPD reports twenty-one categories in the ED: self-pay, other non-federal programs, preferred provider organization, point of service, exclusive provider organization Medicare risk, automobile medical, Blue Cross/Blue Shield, Champus (Tricare), commercial insurance company, disability, health maintenance organization, Medicare Part A, Medicare Part B, Med-Cal, other federal program, title V, veterans affairs plan, workers' compensation health claim, other, and invalid/unknown. These were recoded into five categories: Medicare, Medi-Cal, private, other, and no insurance.
- Rurality – OSHPD reports the patient's five digit ZIP code of residence. These were recoded into three categories, rural, urban and frontier. Urban, defined as a population range 75,000 to 125,000 with similar demographic and socio-economic characteristics. Rural is defined as a population density of less than 250 persons per square mile, not exceeding 50,000. Frontier is defined as a population density of less than 11 persons per square mile.
- Income – Median income by ZIP code from the Census Bureau was assigned to each admission based on the reported ZIP code. This variable was square root transformed and grand mean centered, as this is the only centering option for level two variables, noted by Enders and Tofighi (2007).

The enabling factors in our county level model were:

- Medicare – Percentage of diabetes admission covered by Medicare by county.
- Medi-Cal – Percentage of diabetes admissions covered by Medi-Cal by county.
- Private – Percentage of diabetes admissions covered by private insurance by county.
- No insurance – Percentage of diabetes admissions with no insurance by county.
- Income – Median income by county from the Census Bureau. This variable was mean centered.
- Clinical care – County ranking based on access to care and quality of care. This variable was mean centered. The best county is ranked 1, meaning the clinical care is the best in this county.

The need factors in our model were:

- Charlson Comorbidity Index – The Charlson Comorbidity Index (Charlson et al., 1987) was used in order to evaluate comorbidities, using ICD-9 codes. This variable was mean centered.

The need factors in our county level model were:

- Charlson Comorbidity Index –The sum of the Charlson Comorbidity Index for each diabetes admission, by county, divided by the total number of admissions for each county to get the average Charlson Comorbidity Index for each county. This variable was mean centered.
- Health behaviors – County ranking based on tobacco use, diet and exercise, alcohol and drug use, and sexual activity. This variable was mean centered. The best county is ranked 1, meaning health behaviors is the best in this county.
- Physical environment – County ranking based on air and water quality and housing and transit. This variable was mean centered. The best county is ranked 1, meaning the physical environment is best in this county.

Outcome variables

Two outcome variables were used for the individual level analysis:

- Cost of care – For PDD, DRG information from the OSHPD is linked with Medicare inpatient payment data to get the cost per admission for treating diabetes. For ED, CPT codes from the OSHPD are linked with national average Medicare payment for a facility in order to get the cost per admission for treating diabetes.
- Length of stay – An original PDD OSHPD variable, which is the total number of days from admission date to discharge date.

Two outcomes variables were used for the county level analysis:

- Admissions rates – Admissions for a principal diagnosis of diabetes (ICD-9 code 250) divided by an estimate of the number of people with diabetes in the county, per 10,000 people. This was done for both the PDD and ED.
- Average length of stay –The sum of length of stay in days for each diabetes admission, by county, divided by the total number of admissions for each county to get the average length of stay (in days) for each county.

Data Analysis

Our main goal is to model regional variations in various healthcare utilization outcomes. Because we are dealing with different types of outcomes and data structures, for our small-area analyses we will be using a variety of general linear models to analyze our data. We made statistical inferences using an alpha level of .05 and used Stata version 14 to estimate each model.

Admissions

Multiple regression was used to investigate whether variations in admission rates for diabetes in patient discharge data and emergency department data are due to differences in the makeup of the population at the county level. Diabetes admissions rates were computed using the following formula:

$\text{rate}_{\text{admissions}}$= # of diabetes visits / total # of people with diabetes x 10,000

Rates was modeled using a weighted binomial generalized linear model. Age, gender, insurance type, median income, comorbidity, and Wisconsin health factors ranking all at the county level were controlled for. Admission rates was modeled using a weighted binomial generalized liner model described in the following equation:

$$y_i = \beta_0 + \beta_1 * AGE_{CENTERED_i} + \beta_2 * FEMALE_i + \beta_3 * PRIVATEINSURANCE_i + \beta_4 * MEDICAREINSURANCE_i + \beta_5 * MEDICALINSURANCE_i + \beta_6 * NOINSURANCE_i + \beta_7 * NHWHITE_i + \beta_8 * NHBLACK_i + \beta_9 * NHASIAN_i + \beta_{10} * HISPANIC_i + \beta_{11} * CHARLSON_{CENTEREDi} + \beta_{12} * INCOME_{GRAN\ CENTEREDi} + \beta_{13} * HEALTHBEHAVIORRANK_{CENTEREDi} + \beta_{14} * CLINICALCARERANK_{CENTEREDi} + \beta_{15} * SOCIALECONRANK_{CENTEREDi} + \beta_{16} * PHYSICALENVIRONRANK_{CENTEREDi} + e_i$$

where y_i is the admission rate for the ith county, β_0 is the overall intercept, β_1 is the effect of age on admission rate, β_2 is the effect of female on admission rate, β_3 is the effect of private insurance on admission rate, β_4 is the effect of Medicare on admission rate, β_5 is the effect of Medi-Cal on admission rate, β_6 is the effect of age on admission rate, β_7 is the effect of NH White on admission rate, β_8 is the effect of NH Black on admission rate, β_9 is the effect of NH Asian on admission rate, β_{10} is the effect of Hispanic on admission rate, β_{11} is the effect of comorbidities on admission rate, β_{12} is the effect of income on admission rate, β_{13} is the effect of the health behavior county rank on admission rate, β_{14} is the effect of the clinical care county rank on admission rate, β_{15} is the effect of the social and economic factors county rank on admission rate, β_{16} is the effect of the physical environment county rank on admission rate, and $e_i$is random error distributed binomially.

Length of stay

Poisson regression was used to examine if there are regional variations across ZIP codes for length of stay (LOS) for admissions for diabetes. The outcome, LOS, is measured in number of days. Because individual admissions are nested within ZIP codes, a generalized mixed effects approach was utilized to model the multilevel structure of the data. Individual admissions level covariates were included. Level-one covariates include age, gender, ethnicity, insurance type, and comorbidities. Length of stay was modeled using a multi-level model using the following formula:

Level-one model:

$$y_{ij} = \beta_{0j} + \beta_{1j} * AGE_{CENTERED_{ij}} + \beta_{2j} * FEMALE_{ij} + \beta_{3j} * PRIVATEINSURANCE_{ij} + \beta_{4j} * MEDICAREINSURANCE_{ij} + \beta_{5j} * MEDICALINSURANCE_{ij} + \beta_{6j} * OTHERINSURANCE_{ij} + \beta_{7j} * NHBLACK_{ij} + \beta_{8j} * NHASIAN_{ij} + \beta_{9j} * HISPANIC_{ij} + \beta_{10j} * NHOTHER_{ij} + \beta_{11j} * CHARLSON_{CENTEREDij} + e_{ij}$$

Level-two model:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + u_{2j}$$

$$\dots$$
$$\beta_{nj} = \gamma_{n0} + u_{nj}$$

where y is length of stay in days, β_0 is the overall intercept, β_1 is the effect of age on LOS, β_2 is the effect of female on LOS, β_3 is the effect of private insurance on LOS, β_4 is the effect of Medicare on LOS, β_5 is the effect of Medi-Cal on LOS, β_6 is the effect of other insurance on LOS, β_7 is the effect of NH Black on LOS, β_8 is the effect of NH Asian on LOS, β_9 is the effect of Hispanic on LOS, β_{10} is the effect of NH Other on LOS, β_{11} is the effect of having a comorbidity on LOS, γ_{n0} is the intercept for the nth random slope, u_{nj} is the variance component of the nth random slope distributed normally, and e_i is random error distributed binomially.

Next, multiple regression was used to investigate whether variations in average LOS for diabetes in patient discharge data is due to differences in the makeup of the population at the county level using the following equation:

$$\begin{aligned} y_i = \beta_0 + \beta_1 * AGE_{CENTERED_i} + \beta_2 * FEMALE_i + \beta_3 * PRIVATEINSURANCE_i + \beta_4 \\ * MEDICAREINSURANCE_i + \beta_5 * MEDICALINSURANCE_i + \beta_6 * NOINSURANCE_i \\ + \beta_7 * NHWHITE_i + \beta_8 * NHBLACK_i + \beta_9 * NHASIAN_i + \beta_{10} * HISPANIC_i + \beta_{11} \\ * CHARLSON_{CENTEREDi} + \beta_{12} * INCOME_{GRAN\ CENTEREDi} + \beta_{13} \\ * HEALTHBEHAVIORRANK_{CENTEREDi} + \beta_{14} * CLINICALCARERANK_{CENTEREDi} \\ + \beta_{15} * SOCIALECONRANK_{CENTEREDi} + \beta_{16} * PHYSICALENVIRONRANK_{CENTEREDi} \\ + e_i \end{aligned}$$

where y_i is the admission rate for the ith county, β_0 is the overall intercept, β_1 is the effect of age on admission rate, β_2 is the effect of female on admission rate, β_3 is the effect of private insurance on admission rate, β_4 is the effect of Medicare on admission rate, β_5 is the effect of Medi-Cal on admission rate, β_6 is the effect of age on admission rate, β_7 is the effect of NH White on admission rate, β_8 is the effect of NH Black on admission rate, β_9 is the effect of NH Asian on admission rate, β_{10} is the effect of Hispanic on admission rate, β_{11} is the effect of comorbidities on admission rate, β_{12} is the effect of income on admission rate, β_{13} is the effect of the health behavior county rank on admission rate, β_{14} is the effect of the clinical care county rank on admission rate, β_{15} is the effect of the social and economic factors county rank on admission rate, β_{16} is the effect of the physical environment county rank on admission rate, and $e_i$is random error distributed normally.

Cost

Multilevel regression was used to investigate the relationship between cost and individual, as well as county level predictors in both discharge and emergency department data. This analysis also examined if these covariates vary across ZIP codes. Level-one covariates include age, gender, ethnicity, insurance type, and comorbidities. Level-two covariates include those in level one as well as median ZIP code level income and rurality. The multilevel model with no level-two predictors for cost is modeled using the following model:

Level-one model:

$$y_{ij} = \beta_{0j} + \beta_{1j} * AGE_{CENTERED_{ij}} + \beta_{2j} * FEMALE_{ij} + \beta_{3j} * PRIVATEINSURANCE_{ij} + \beta_{4j} * MEDICAREINSURANCE_{ij} + \beta_{5j} * MEDICALINSURANCE_{ij} + \beta_{6j} * OTHERINSURANCE_{ij} + \beta_{7j} * NHBLACK_{ij} + \beta_{8j} * NHASIAN_{ij} + \beta_{9j} * HISPANIC_{ij} + \beta_{10j} * NHOTHER_{ij} + \beta_{11j} * CHARLSON_{CENTEREDij} + e_{ij}$$

where y is cost of treating diabetes, β_0 is the overall intercept, β_1 is the effect of age on cost, β_2 is the effect of female on cost, β_3 is the effect of private insurance on cost, β_4 is the effect of Medicare on cost, β_5 is the effect of Medi-Cal on cost, β_6 is the effect of other insurance on cost, β_7 is the effect of NH Black on cost, β_8 is the effect of NH Asian on cost, β_9 is the effect of Hispanic on cost, β_{10} is the effect of NH Other on cost, β_{11} is the effect of having a comorbidity on cost, and e_i is random error distributed binomially.

Level-two model:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + u_{2j}$$

$$\dots$$

$$\beta_{nj} = \gamma_{n0} + u_{nj}$$

The multilevel model with level-two predictors for cost was modeled using the following equation:

Level-one model:

$$y_{ij} = \beta_{0j} + \beta_{1j} * AGE_{CENTERED_{ij}} + \beta_{2j} * FEMALE_{ij} + \beta_{3j} * PRIVATEINSURANCE_{ij} + \beta_{4j} * MEDICAREINSURANCE_{ij} + \beta_{5j} * MEDICALINSURANCE_{ij} + \beta_{6j} * OTHERINSURANCE_{ij} + \beta_{7j} * NHBLACK_{ij} + \beta_{8j} * NHASIAN_{ij} + \beta_{9j} * HISPANIC_{ij} + \beta_{10j} * NHOTHER_{ij} + \beta_{11j} * CHARLSON_{CENTEREDij} + e_{ij}$$

where y is cost of treating diabetes, β_0 is the overall intercept, β_1 is the effect of age on cost, β_2 is the effect of female on cost, β_3 is the effect of private insurance on cost, β_4 is the effect of Medicare on cost, β_5 is the effect of Medi-Cal on cost, β_6 is the effect of other insurance on cost, β_7 is the effect of NH Black on cost, β_8 is the effect of NH Asian on cost, β_9 is the effect of Hispanic on cost, β_{10} is the effect of NH Other on cost, β_{11} is the effect of having a comorbidity on cost, and e_i is random error distributed binomially.

Level-two model:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} * INCOME_{\ Gran\ Centered j} + \gamma_{02} * URBAN_{\ j} + \gamma_{03} * FRONTIER_{\ j} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + u_{2j}$$

$$\dots$$

$$\beta_{nj} = \gamma_{n0} + u_{nj}$$

where y is cost, γ_{00} is the overall intercept, γ_{01} is the effect of median income on cost, γ_{02} is the effect of an Urban ZIP code on cost, γ_{03} is the effect of a Frontier ZIP code on cost, and u_{0j} is random error distributed binomially.

Results

The original 2013 OSHPD patient discharge included 3,806,911 admissions, and the emergency department included 10,897,885 admissions. The analyses were done using two different data sets. The cost analyses were done using the patient discharge and emergency department data, but included only admissions for the principal diagnosis of diabetes (ICD-9 code 250). One length of stay analysis was done using patient discharge data as well. After dropping all other admissions data, the patient discharge data included 54,526 admissions. The emergency department data included 72,776 admissions. The other length of stay analysis was done using a patient discharge county level estimates dataset. The admissions analyses were done using this same patient discharge county level estimates dataset as well as an emergency department county level estimates dataset. The county level estimates dataset included 58 observations (one for each county).

Table 2 shows average California demographic information for diabetes admissions for the patient discharge and emergency department data set for 2013. Across counties, 56.86% of all diabetes admissions for the PDD are male, while 52.82% are male for the ED. The average age of PDD admissions is 51.28 years (SD=19.98) and 52.78 years (SD=18.54) for the ED. Roughly 38.94% are White for the PDD, and 35.85% for the ED. 15.04% of admissions in the PDD are Black and 16.14% of admissions for the ED. Only 5.4% of admissions are Asian in the PDD, and 5.03% in the ED. Hispanics account for 37.76% of all diabetes admissions for the PDD, and 36.45% of admissions in the ED. 36.45% of PDD admissions for diabetes have Medicare and 33.08% for the ED have Medicare. 26.76% of admissions in the PDD and 21.63% in the ED have Medi-Cal. 19.26% of admissions have private insurance for the PDD and 18.39% for the ED. The average Charlson Comorbidity Index across counties for diabetes admissions is 0.77 (SD=1.18) for the PDD, and 0.82 (SD=0.95) for the ED.

Table 2: CA demographics for Patient discharge and emergency department diabetes admissions, 2013

	Average Age	% Male	% NH White	% NH Black	% NH Asian	% Hispanic	% Medicare	% Medical	% Private	Average Charlson Comorbidity Index
California PDD	51.28	56.86	38.94	15.04	5.40	37.76	36.45	26.76	19.26	0.77
California ED	52.78	52.82	35.85	16.14	5.03	39.56	33.08	21.63	18.39	0.82

Admissions

Table 3 shows admissions divided by the number with diabetes rate per 10,000 for diabetes for the top and bottom five counties for both the PDD and ED. The BIC for the PDD model is -161.70, and for the ED model is -161.69. The AIC for the PDD model is 0.73 and 0.76 for the ED. Insurance status is the only significant predictor in the PDD model. We found that for every 1% increase in the uninsured rate leads to a 0.03 (p=.01) increase in diabetes admissions. This suggests that counties with a higher percent of people with no insurance are more likely to have higher rates of diabetes admissions. In the ED the only significant predictor is the social and economic factors county ranking. Counties that are above the average social and economic factors ranking, meaning they have a lower score, tend to have higher rates of diabetes admissions. Results of these models can be found in table 4.

Table 3: Top and bottom 5 counties diabetes admissions rates for patient discharge and emergency department (per 10,000), 2013

Bottom 5 Counties	PDD Admission Rates	Bottom 5 Counties	ED Admission Rates
Mono	64.01	Alpine	0
San Mateo	74.07	Sierra	48.99
Yolo	77.28	Mono	49.79
Alpine	85.97	Lassen	126.96
Madera	86.08	Nevada	130.83
Top 5 Counties	PDD Admission Rates	Top 5 Counties	ED Admission Rates
Lake	264.88	Tehama	393.93
Stanislaus	282.12	Siskiyou	431.34
San Benito	295.81	Lake	439.43
Mendocino	298.84	Mendocino	471.97
Trinity	321.37	Del Norte	508.81
CA	149.35	CA	199.34

Table 4: Diabetes admission county level regression models patient discharge and emergency department, 2013

	PDD			ED		
	B	SE	P	B	SE	P
Percent Female	0.01	0.01	0.63	-0.01	0.01	0.46
Age (mean centered)	0.04	0.02	0.11	-0.06	0.03	0.07
Percent NH White	0.00	0.01	0.85	0.01	0.02	0.70
Percent NH Black	0.00	0.01	0.77	0.01	0.02	0.49
Percent NH Asian	-0.01	0.02	0.74	0.00	0.02	0.81
Percent Hispanic	-0.01	0.01	0.31	-0.01	0.02	0.71
Percent No Insurance	0.03	0.01	0.01	0.02	0.01	0.06
Percent Medicare	-0.01	0.01	0.60	0.00	0.01	0.84
Percent Medi-Cal	0.01	0.01	0.12	-0.01	0.01	0.24
Percent Private	-0.03	0.01	0.07	0.01	0.01	0.18
County Median Income (mean centered)	0.00	0.00	0.90	0.00	0.00	0.91
Charlson Comorbidity Score (mean centered)	-0.36	0.30	0.24	-0.13	0.34	0.70
Health Behaviors Rank (mean centered)	0.01	0.00	0.24	0.00	0.00	0.79
Clinical Care Rank (mean centered)	0.00	0.01	0.76	-0.01	0.01	0.19
Social and Economic Factors Rank (mean centered)	-0.01	0.01	0.39	0.01	0.01	0.02
Physical Environment Rank (mean centered)	0.00	0.00	0.24	0.00	0.00	0.23
Constant	-3.48	1.23	0.01	-3.90	1.29	0.00
N	57			57		
BIC	-161.70			-161.69		

Length of stay

Table 5 shows the average LOS (in days) for the top and bottom five counties for PDD. The overall model Chi-Square for the Poisson regression model is 16526.62 (p<.001). The BIC is 425669. The significant predictors are female, age, NH Black, NH Asian, Hispanic, Medicare, Medi-Cal, other insurance, and Charlson Comorbidity Index. We found that women tend to have a shorter length of stay than men by 0.05 (p<.001) days. Every year above the average age increases length of stay by 0.01 (p<.001) days. NH Blacks, compared to NH Whites, stay 0.12 (p<.001) days less. NH Asians stay 0.16 (p<.001) days less than NH Whites. Compared to NH Whites, Hispanics tend to say 0.02 (p<.001) days less. Those covered with Medicare stay 0.02 (p=.022) days more than those with no insurance, and those with Medi-Cal stay 0.34 (p<.001) days more than uninsured individuals. For every point above the average Charlson Comorbidity Index score, length of stay increases by 0.08 (p<.001) days. Although no random effects were used in the model, the model shows that there is variation in length of stay at the ZIP code level. Results of this model can be found in table 6.

Results of the county level length of stay analysis can be found in table 7. The BIC for this model is -138.49, and the AIC is 2.54. There are no significant predictors in this model.

Table 5: Top and bottom 5 counties length of stay for patient discharge data, 2013

Bottom 5 Counties	PDD Average LOS (days)
Sierra	1.67
Mono	3
Nevada	3.23
Calaveras	3.37
Butte	3.45
Top 5 Counties	**PDD Average LOS (days)**
San Francisco	6.4
Humboldt	7.21
Trinity	14.46
Modoc	17
Mariposa	30.05
CA	4.823

Table 6: Diabetes length of stay multi-level model patient discharge, 2013

	B	SE	P	95% Confidence Interval	
Female	-0.05	0.00	0.00	-0.06	-0.05
Age (mean centered)	0.01	0.00	0.00	0.01	0.01
NH Black	-0.12	0.01	0.00	-0.13	-0.10
NH Asian	-0.16	0.01	0.00	-0.18	-0.14
NH Other	0.01	0.01	0.65	-0.02	0.03
Hispanic	-0.02	0.01	0.00	-0.03	-0.01
Private Insurance	-0.02	0.01	0.07	-0.04	0.00
Medicare	0.02	0.01	0.02	0.00	0.04
Medi-Cal	0.34	0.01	0.00	0.32	0.36
Other Insurance	0.16	0.01	0.00	0.14	0.18
Charlson Comorbidity Score (mean centered)	0.08	0.00	0.00	0.08	0.08
Constant	1.39	0.01	0.00	1.36	1.42
Patient ZIP (constant)	0.15	0.01		0.14	0.17
N	55,203				
BIC	425669				

Table 7: Diabetes length of stay county level regression model patient discharge, 2013

	B	SE	P
Percent Female	-0.02	0.03	0.60
Age (mean centered)	0.08	0.07	0.26
Percent NH White	0.00	0.05	0.96
Percent NH Black	-0.02	0.05	0.73
Percent NH Asian	0.07	0.08	0.39
Percent Hispanic	-0.01	0.05	0.80
Percent No Insurance	-0.03	0.05	0.46
Percent Medicare	0.04	0.03	0.30
Percent Medi-Cal	-0.02	0.02	0.49
Percent Private	-0.06	0.05	0.22
County Median Income (mean centered)	0.00	0.00	0.89
Charlson Comorbidity Score (mean centered)	0.10	1.10	0.93
Health Behaviors Rank (mean centered)	0.01	0.02	0.60
Clinical Care Rank (mean centered)	0.02	0.02	0.25
Social and Economic Factors Rank (mean centered)	0.01	0.02	0.80
Physical Environment Rank (mean centered)	-0.01	0.01	0.37
Constant	6.57	4.45	0.14
N	57		
BIC	-138.49		

Cost

Table 8 shows average cost for the top and bottom five counties for treating diabetes in the PDD and ED. Inferences for the factors that are in both models in the PDD and ED are the same, except for NH Black is significant in the model that does not include any county level predictors in the PDD. Other than this, they did not change, however by including counties we significantly improved our model fit based on the BIC for both the PDD and ED models and therefore decided to report and interpret the final model for each data set which includes level two predictors. For both the PDD and ED, the overall model was significant. The overall model Chi-Square is 6250.06 ($p<.001$) for the PDD and 951.67 ($p<.001$) for the ED. The BIC for the PDD model is 497563.20, and for the ED model is 213206.20. The significant predictors of the PDD model for first level fixed effects are female, age, Hispanic, private insurance, Medicare, Medi-Cal, other insurance, and Charlson Comorbidity Index. The significant predictors of the ED model for first level fixed effects were all the same except that other insurance is not significant. For second level fixed effects in the PDD, frontier was significant, and in the ED income and urban. In the PDD, we found that females have a cost of 24.56 ($p<.001$) dollars less than males, and in the ED females have a cost of 0.06 ($p<.001$) dollars higher than males. For PDD, for every year above the average age cost increased by 0.07 ($p<.001$) dollars, and for the ED every year above the average age decreased cost by 0.03 ($p<.001$) dollars. Compared to NH Whites, Hispanic's cost 11.09 ($p<.001$) dollars more in the PDD, and 0.01 ($p=.034$) dollars less in the ED. Having private insurance increased cost by 17.72 ($p<.001$) dollars compared to not having insurance in the PDD, and increased cost by 1.61 ($p<.001$) dollars in the ED. Medicare, compared to no insurance, cost 41.78 ($p<.001$) dollars more in the PDD and 2.35 ($p<.001$) dollars more in the ED. Medi-Cal, compared to no insurance, cost 8.98 ($p<.001$) dollars more in the PDD and 0.0004 ($p=.007$) dollars more in the ED. Other insurance, compared to no insurance, cost 2.09 ($p=.008$) dollars more in the PDD. For every point above the average Charlson Comorbidity Index score, the cost increases by 14.96 ($p<.001$) dollars in the PDD and decreases by 0.04 ($p<.001$) dollars in the ED. Frontier, compared to rural areas, has a lower cost by 5.53 ($p=.030$) dollars in the PDD. In the ED, urban, compared to rural, has a lower cost by 0.28 ($p<.001$) dollars. For income, every dollar above the average income increased cost by 0.0000086 ($p=.005$) dollars in the ED. Examining the results of the random effects parameters, we see that all of the variance components in the model are significant. This shows that sex, age, race, insurance type and comorbidities vary across ZIP codes in both PDD and ED. These findings emphasize that regional variation exist. Results can be found in tables 9 and 10.

Table 8: Top and bottom 5 counties average cost for patient discharge and emergency department, 2013

Bottom 5 Counties	PDD Average Cost	Bottom 5 Counties	ED Average Cost
Sierra	$4,262.93	Alpine	$0.00
Lassen	$5,991.65	Trinity	$31.46
Plumas	$6,262.58	Imperial	$81.66
Siskiyou	$6,678.46	San Bernardino	$100.83
Butte	$6,735.54	Lake	$106.33
Top 5 Counties	**PDD Average Cost**	**Top 5 Counties**	**ED Average Cost**
Sonoma	$9,281.10	Mono	$188.30
Tuolumne	$9,405.13	San Diego	$190.30
Colusa	$10,105.31	Plumas	$204.61
Modoc	$10,454.91	Mariposa	$238.46
Alpine	$27,133.44	Sierra	$307.56
CA	$8,228.99	CA	$136.53

Table 9: Diabetes cost multi-level models patient discharge, 2013

	Model 1			Model 2		
	B	SE	P	B	SE	P
Female	-4.92	0.22	0.00	-4.96	0.22	0.00
Age (mean centered)	0.26	0.01	0.00	0.26	0.01	0.00
NH Black	-0.75	0.36	0.04	-0.56	0.37	0.14
NH Asian	-0.50	0.62	0.42	-0.71	0.62	0.25
NH Other	1.15	0.70	0.10	1.02	0.74	0.17
Hispanic	3.27	0.27	0.00	3.33	0.28	0.00
Private Insurance	4.31	0.48	0.00	4.21	0.50	0.00
Medicare	6.53	0.50	0.00	6.46	0.52	0.00
Medi-Cal	3.09	0.47	0.00	3.00	0.49	0.00
Other Insurance	1.65	0.52	0.00	1.45	0.54	0.01
Charlson Comorbidity Score (mean centered)	3.89	0.11	0.00	3.87	0.11	0.00
ZIP Median Income (grand mean centered)				0.00	0.00	0.95
Urban				-0.44	0.39	0.27
Frontier				-2.35	1.09	0.03
Constant	83.28	0.45	0.00	83.78	0.57	0.00

	Model 1				Model 2			
Random-effects Parameters	Estimate	SE	95% Confidence Interval		Estimate	SE	95% Confidence Interval	
Female	0.00	0.02	0.00	5.82E+14	0.00	0.08	0.00	6.96E+14
Age (mean centered)	0.00	0.00	0.00	984828.50	0.00	0.00	0.00	2.17E+10
NH Black	4.56	2.54	1.53	13.56	4.46	2.62	1.41	14.08
NH Asian	54.96	12.32	35.42	85.30	51.10	11.68	32.64	79.98
NH Other	29.36	19.44	8.02	107.47	41.79	16.45	19.32	90.39
Hispanic	6.21	1.67	3.67	10.51	5.09	1.67	2.67	9.69
Private Insurance	1.84	2.24	0.17	19.84	1.57	2.41	0.08	31.67
Medicare	18.27	2.76	13.59	24.56	17.39	2.68	12.85	23.53
Medi-Cal	4.52	1.78	2.10	9.77	4.52	1.83	2.05	9.98
Other insurance	0.00	0.01	0.00	8.44E+16	0.00	0.00	0.00	0.00
Charlson Comorbidity Score (mean centered)	1.72	0.48	0.99	2.99	1.65	0.48	0.94	2.90
Constant	0.02	0.10	0.00	205.96	0.00	0.00	0.00	1.81E+18
Residual	611.41	3.86	603.89	619.04	614.46	3.93	606.81	622.20
N	55203				53583			
BIC	512332.60				497563.20			

Table 10: Diabetes cost multi-level models emergency department, 2013

	Model 1			Model 2		
	B	SE	P	B	SE	P
Female	0.24	0.04	0.00	0.24	0.04	0.00
Age (mean centered)	-0.02	0.00	0.00	-0.02	0.00	0.00
NH Black	-0.01	0.07	0.87	0.03	0.07	0.62
NH Asian	-0.18	0.10	0.07	-0.17	0.10	0.09
NH Other	-0.10	0.12	0.44	-0.09	0.13	0.46
Hispanic	-0.12	0.05	0.01	-0.11	0.05	0.03
Private Insurance	1.26	0.07	0.00	1.27	0.07	0.00
Medicare	1.54	0.07	0.00	1.53	0.07	0.00
Medi-Cal	0.21	0.07	0.00	0.20	0.07	0.01
Other Insurance	-0.10	0.08	0.22	-0.13	0.09	0.15
Charlson Comorbidity Score (mean centered)	-0.20	0.03	0.00	-0.19	0.03	0.00
ZIP Median Income (grand mean centered)				0.00	0.00	0.01
Urban				-0.53	0.12	0.00
Frontier				-0.37	0.33	0.27
Constant	10.29	0.07	0.00	10.75	0.12	0.00

	Model 1				Model 2			
Random-effects Parameters	Estimate	SE	95% Confidence Interval		Estimate	SE	95% Confidence Interval	
Female	0.00	0.02	0.00	5.82E+14	0.00	0.08	0.00	6.96E+14
Age (mean centered)	0.00	0.00	0.00	984828.50	0.00	0.00	0.00	2.17E+10
NH Black	4.56	2.54	1.53	13.56	4.46	2.62	1.41	14.08
NH Asian	54.96	12.32	35.42	85.30	51.10	11.68	32.64	79.98
NH Other	29.36	19.44	8.02	107.47	41.79	16.45	19.32	90.39
Hispanic	6.21	1.67	3.67	10.51	5.09	1.67	2.67	9.69
Private Insurance	1.84	2.24	0.17	19.84	1.57	2.41	0.08	31.67
Medicare	18.27	2.76	13.59	24.56	17.39	2.68	12.85	23.53
Medi-Cal	4.52	1.78	2.10	9.77	4.52	1.83	2.05	9.98
Other insurance	0.00	0.01	0.00	8.44E+16	0.00	0.00	0.00	0.00
Charlson Comorbidity Score (mean centered)	1.72	0.48	0.99	2.99	1.65	0.48	0.94	2.90
Constant	0.02	0.10	0.00	205.96	0.00	0.00	0.00	1.81E+18
Residual	611.41	3.86	603.89	619.04	614.46	3.93	606.81	622.20
N	55203				53583			
BIC	512332.60				497563.20			

Conclusion

The purpose of this study was to identify variations in hospital admissions, average length of stay, and average cost per visit for diabetes across all 58 California counties. The results suggest that significant variations in hospital admissions, length of stay, and costs for diabetes exists at the county level, and in some cases even at the ZIP code level. For the admissions analysis, the results suggest that counties with higher levels of uninsured people are likely to have higher rates of hospital admissions for diabetes. The uninsured are using the hospital to treat their diabetes, instead of primary care. Not having insurance can lead to uncontrolled diabetes because of a lack of going to their primary care provider for checkups and taking medication. This suggests that it is important to keep more people insured in order to prevent hospital admissions due to diabetes. In the ED admissions analysis, the results suggest that counties that were ranked higher than the average rank for social and economic factors tend to have higher rates of diabetes related admissions. In other words, counties that have worse social and economic factors have higher rates of diabetes admissions. The measures included in this ranking are education level, unemployment rate, children in poverty, income inequality, children in single parent households, social associations, violent crimes, and injury deaths (County Health Rankings, 2017). More stressful living conditions caused by these measures can explain why counties with worse social and economic factors visit the emergency department more frequently.

In regards to the length of stay analyses, the results suggest that women stay less days than men. As age increases length of stay increases. NH Blacks, NH Asians, and Hispanics all stay less days than NH Whites. This is an interesting finding because according to the American Diabetes Association (2017), NH Whites have the lowest rates of diagnosed diabetes, yet they are staying more days. Those with Medicare and Medi-Cal stay more days than uninsured. Although not statistically significant at the p=.05 level, those with private insurance stay less days for diabetes compared to those with no insurance (p=.073). This suggests that healthier and younger people stay less days for diabetes. Private insurance is considered to be high quality insurance, meaning that those covered under these types of insurance plans will have more primary care visits, checkups, and access to medications resulting in more control of their diabetes. Those covered under Medicare have similar benefits to those with private insurance, but tend to be above 65 years old, which means that they tend to have more health issues. Those covered with Medi-Cal tend to be people of less socioeconomic status, and might not have access to the services that someone covered with private insurance or Medicare would have, thus would maybe have to stay in the hospital more days because their diabetes is not managed (Dayaratna, 2012). This goes along the lines of the finding where more comorbidities results in a longer length of stay. The more complications to one's health, the longer it takes to treat them, resulting in more days spent in the hospital. No significant findings were found when examining length of stay at the county level. This could be because the analysis contained only 58 observations, one per county, and lacked statistical power. In the multi-level model, we found that length of stay varied at the ZIP code level, which is interesting because within a county there are several ZIP codes. This suggests that in the county level analysis, with more statistical power we would find significant differences at the county level.

Regarding the cost analyses, the results suggest there are some differences between inpatient and outpatient costs of diabetes. It is important to note that costs are national level estimates from the CMS, meaning that the same DRG or CPT code will be paid the same amount, regardless of insurance type or county. DRG's are used for PDD, and CPT's are used for ED. The differences in cost are strictly due to differences in the DRG or CPT codes used to treat diabetes. In PDD, females tend to cost less than males, but in ED females cost more than males. It could be due to in hospitals females are generally getting less complicated DRG's, while in the ED they are getting more complicated CPT's. As age increases, cost decreases in ED. This could be due because those 65 years and older are covered under Medicare. In the PDD, as age increases, cost increases, which is expected. Interestingly, in the PDD Hispanics cost more than NH Whites, while in the ED Hispanics cost less. This can be explained, once again, with different DRG and CPT codes being used, where Hispanics are getting different types of treatment in the hospital that are more costly than NH Whites, and getting different types of treatment in the emergency department that are less costly than NH Whites. Private insurance, Medicare and Med-Cal costs for treating diabetes are higher than uninsured costs in both PDD and ED. Other insurance, which includes any type of insurance which is not private, Medicare, or Medi-Cal, is more costly than uninsured costs for diabetes treatment in patient discharge. This suggests that more costly codes are being used to treat those with insurance. It could be that they are getting better treatment, which is more costly. In the PDD, the higher the comorbidity score the higher the cost. This is expected because sicker people require more expensive treatment. Interestingly, in the ED, the higher the comorbidity score the lower the cost is for treating diabetes. One possible explanation could be that these people are seen in the emergency department and are then admitted to the hospital due to not being able to sufficiently treat them in the emergency department. This would explain why people who have an above average Charlson Comorbidity Index have lower costs in the ED, and higher costs in the hospital. Frontier, compared to rural areas, cost less in the PDD. In the ED, urban ZIP codes have a lower cost compared to rural areas. The higher the income above the average, the more diabetes care costs in the ED. The results of the multi-level model suggests that there is variation in cost across ZIP codes. This is interesting as counties often consist of multiple ZIP codes. This would suggest that even in a single county, variations could exist in cost of diabetes depending on the ZIP code.

The results reported here are consistent with previous studies, showing that regional variations exist in access, quality (LOS), and cost regarding health care. Although many studies have documented such regional variations, no study has specifically examined regional variations for diabetes across California counties. Many of the factors we tested in our models that possibly explain why regional variations exist are consistent with previous studies.

It is important to emphasize the limitations of this study. As with any analysis of secondary hospital data, many important variables of interest are not available, including income, cost of care and family history of diabetes. Because we are limited in the variables we can adjust for, and the difficulty in assessing potential confounders as an explanation for small effect sizes which are statistically significant, analyses of large secondary data should be interpreted with caution. Also, we did not have linked hospital data so we could not track if the same person came in for multiple diabetes admissions. We were not able to track patients across discharge and emergency department data either. Prevalence estimates for diabetes by county were gathered from the CHIS. This is all self-reported survey data, which has limitations of its own. Future research should use linked hospital data to see whether the same individual utilizes both the emergency department and hospital for diabetes treatment, and also if they are utilizing either service multiple times for diabetes treatment.

References

American Diabetes Association, 2017. Retrieved from http://www.diabetes.org/diabetes-basics/statistics/

Andersen, R. M. (1995). Revisiting the behavioral model and access to medical care: does it matter?. *Journal of health and social behavior*, 1-10.

Babitsch, B., Gohl, D., & von Lengerke, T. (2012). Re-revisiting Andersen's Behavioral Model of Health Services Use: a systematic review of studies from 1998–2011. *GMS Psycho-Social-Medicine*, *9*.

Brown, P. M., Gonzalez, M., & Dhaul, R. S. (2015). Cost of chronic disease in California: estimates at the county level. *Journal of Public Health Management and Practice*, *21*(1), E10-E19.

Camenzind, P. A. (2012). Explaining regional variations in health care utilization between Swiss cantons using panel econometric models. *BMC health services research*, *12*(1), 62.

Charlson, M. E., Pompei, P., Ales, K. L., & MacKenzie, C. R. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of chronic diseases*, *40*(5), 373-383.

Dayaratna, K. D. (2012). Studies show: Medicaid patients have worse access and outcomes than the privately insured. *The Heritage Foundation Backgrounder*, *2740*.

Dhingra, S. S., Zack, M., Strine, T., Pearson, W. S., & Balluz, L. (2010). Determining prevalence and correlates of psychiatric treatment with Andersen's behavioral model of health services use. *Psychiatric Services*, *61*(5), 524-528.

Dixon, T., Shaw, M. E., & Dieppe, P. A. (2006). Analysis of regional variation in hip and knee joint replacement rates in England using Hospital Episodes Statistics. *Public health*, *120*(1), 83-90.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological methods*, *12*(2), 121.

Fisher, E. S., Wennberg, D. E., Stukel, T. A., Gottlieb, D. J., Lucas, F. L., & Pinder, E. L. (2003). The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. *Annals of internal medicine*, *138*(4), 273-287.

Fisher, E. S., Wennberg, J. E., Stukel, T. A., Skinner, J. S., Sharp, S. M., Freeman, J. L., & Gittelsohn, A. M. (2000). Associations among hospital capacity, utilization, and mortality of US Medicare beneficiaries, controlling for sociodemographic factors. *Health services research*, *34*(6), 1351.

Fonseca, R., Otahal, P., Wiggins, N., & Marwick, T. H. (2015). Growth and geographical variation in the use of cardiac imaging in Australia. *Internal medicine journal*, *45*(11), 1115-1127.

Gittelsohn, A., & Powe, N. R. (1995). Small area variations in health care delivery in Maryland. *Health Services Research*, *30*(2), 295.

Kopetsch, T., & Schmitz, H. (2014). Regional variation in the utilisation of ambulatory services in Germany. *Health economics*, *23*(12), 1481-1492.

McKernan, S. C., Kuthy, R. A., Hanley, P. F., Jones, M. P., Momany, E. T., McQuistan, M. R., & Damiano, P. C. (2015). Geographic variation of dental utilization among low income children. *Health & place*, *34*, 150-156.

Parslow, R., Jorm, A., Christensen, H., & Jacomb, P. (2002). Factors associated with young adults' obtaining general practitioner services. *Australian Health Review*, *25*(6), 109-118.

Phillips, K. A., Morrison, K. R., Andersen, R., & Aday, L. A. (1998). Understanding the context of healthcare utilization: assessing environmental and provider-related variables in the behavioral model of utilization. *Health services research*, *33*(3 Pt 1), 571.

Skinner, J., & Fisher, E. (2010). Reflections on geographic variations in US health care. *Dartmouth Institute for Health Policy and Clinical Practice, May*, *12*.

Stockdale, S. E., Tang, L., Zhang, L., Belin, T. R., & Wells, K. B. (2007). The effects of health sector market factors and vulnerable group membership on access to alcohol, drug, and mental health care. *Health services research*, *42*(3p1), 1020-1041.

University of Wisconsin Population Health Institute. *County Health Rankings 2017.*

Vogt, V., Siegel, M., & Sundmacher, L. (2014). Examining regional variation in the use of cancer screening in Germany. *Social science & medicine*, *110*, 74-80.

Weinick, R. M., Zuvekas, S. H., & Cohen, J. W. (2000). Racial and ethnic differences in access to and use of health care services, 1977 to 1996. *Medical Care Research and Review*, *57*(1_suppl), 36-54.

Weinstein, J. N., Lurie, J. D., Olson, P., Bronner, K. K., Fisher, E. S., & Morgan, M. T. S. (2006). United States trends and regional variations in lumbar spine surgery: 1992–2003. *Spine*, *31*(23), 2707.