# Trends and patterns of disparities in diabetes and chronic kidney disease mortality among US counties, 1984-2014.

Benard Moranga

December 13, 2023


## 1 Abstract

**Background**

Diabetes and chronic kidney diseases pose substantial health challenges both within the United States and globally. This study aims to assess age-standardized mortality rates related to diabetes mellitus and chronic kidney disease across different counties.

**Methodology and Setting**

Utilizing validated small area estimation models, I analyzed de-identified death records from the National Center for Health Statistics (NCHS) and population data from the census bureau, NCHS, and the Human Mortality Database. The study focused on estimating county-level mortality rates from 1980 to 2014 for diabetes mellitus and chronic kidney disease.

**Exposure and Outcomes**

The primary focus was on the county of residence, and the study measured age-standardized mortality rates, considering factors such as year, sex, and cause of death.

**Results**

Between 1980 and 2014, the USA documented 2,067,805 deaths attributed to diabetes. The diabetes-related mortality rate surged by 33.6 percent from 1980 to 2000 and subsequently reduced by 26.4 percent from 2000 to 2014.

A total of 1,659,045 deaths associated with Chronic Kidney Disease (CKD) occurred during the same period. Among these, 477,332 were linked to diabetes mellitus, 1,056,150 to hypertension, 122,795 to glomerulonephritis, and 2,768 to other causes. CKD mortality exhibited geographical variations, with notably lower rates in central Colorado, certain counties in southern Florida, California, and some Midwest regions. Conversely, higher mortality rates from CKD were prevalent in numerous Deep South counties, particularly those surrounding the Mississippi River.

**Conclusions**

This study reveals significant disparities in diabetes and CKD mortality across US counties. The findings underscore the need for addressing root causes, enhancing risk factors, improving access to medical care, and ensuring the quality of healthcare services.


**Keywords**: Diabetes, Chronic kidney disease, Disparities, Mortality.

## 2 Background

Diabetes exhibits associations with various long-term medical conditions, among them chronic kidney disease (CKD). In 1990, chronic kidney disease (CKD), a condition preventable with adequate medical care, was the 14th leading cause of death in the u.S. Despite prior calls for action, the prevalence of diabetes is on the rise. Recent findings from the Behavioral Risk Factor Surveillance System (BRFSS), a comprehensive state-based surveillance system, indicate that the self-reported prevalence of diagnosed diabetes among adults aged 18 or older reached 10.6 percent in 2020. The number of people dying from diabetes varies between states and counties in the United States, and this conceals some important differences. Figuring out the areas that are most affected is important since it will help health professionals focus their efforts on communities at higher risk, given the limited available resources. In their study titled "US County-Level Trends in Mortality Rates for Major Causes of Death, 1980-2014," Laura Dwyer-Lindgren et al. systematically analyzed county-level patterns in mortality rates for 21 major causes of death in the United States. Using innovative methods, they addressed implausible cause-of-death codes and employed small-area estimation techniques on National Vital Statistics System data to estimate annual county-level mortality rates from 1980 to 2014. Their results reveal insights into geographic variations in age-standardized mortality rates for the 10 highest-burden causes, providing a nuanced understanding of how mortality patterns differ across U.S. counties. This study aims to explore how mortality patterns linked to diabetes and chronic kidney disease have evolved in U.S. counties from 1980 to

2014. By examining and showcasing these trends, the research strives to provide valuable insights into geographical differences, deepening our comprehension of the impact of these diseases on a county level. This study is significant due to the notable increase in diabetes prevalence, the association between diabetes and chronic kidney disease, and the diverse mortality rates observed across various U.S. regions. The results are anticipated to assist health professionals in efficiently allocating resources and devising interventions tailored to the specific needs of communities at a higher risk.

# 3 Introduction

The contemporary healthcare environment is marked by a complex network of factors influencing the prevalence and disparities in chronic diseases. Within this framework, I am concentrating on a thorough examination of the trends and patterns of disparities in diabetes and chronic kidney disease mortality across different counties in the United States from 1984 to 2014. The primary objective of this report is to offer a comprehensive understanding of the implications arising from these health disparities and provide valuable insights for public health interventions.

**Application Domain and Specific Problem**

In the vast realm of public health, it is essential to comprehend the trends in mortality associated with diabetes and chronic kidney disease to facilitate effective health management and resource allocation. The particular focus of this project is on examining the mortality rates linked to these chronic conditions across diverse counties in the United States. My analysis delves into the exploration of temporal patterns, regional disparities, and the factors influencing mortality rates over a span of three decades.

**Motivation**

The driving force behind this project arises from the pressing necessity to tackle health disparities and advocate for health equity. Given that diabetes and chronic kidney disease stand as prominent causes of morbidity and mortality, it is imperative to subject them to thorough examination to reveal patterns that can guide focused public health interventions. My aim is to contribute to the formulation of strategies that target the reduction of disparities and enhancement of health outcomes at a national level by identifying and comprehending trends in mortality rates.

**Overview of Approach and Results**

My approach involves a meticulous exploration of publicly available data, ensuring data integrity through rigorous pre-processing, and applying advanced data mining techniques. Through the replication and expansion of the study conducted by Mokdad et al. (2022), my goal is to validate their findings and uncover additional insights. The core of my methodology lies in the application of statistical analyses, visualizations, and machine learning models. In terms of results, I expect to unveil nuanced patterns in mortality rates related to diabetes and chronic kidney disease. This revelation aims to contribute to a broader understanding of health disparities across geographical regions. By transparently and comprehensively presenting my findings, the objective is to support further research and foster evidence-based decision-making in public health. This report will intricately delve into my methodology, highlight the richness of my results, and discuss the implications of my findings in the subsequent sections. Through this exploration, I aim to emphasize the significance of comprehending and addressing health disparities for the enhancement of public health outcomes.

# 4 Design of the Original Study.

In the original paper ”Trends and patterns of disparities in diabetes and chronic kidney disease mortality among US counties, 1980-2014” by Mokdad et al. (2022), the authors aimed to estimate age-standardized mortality rates by county from diabetes mellitus and chronic kidney disease (CKD). They employed validated small area estimation models on de-identified death records from the National Center for Health Statistics (NCHS) and population counts from various sources, covering the period from 1980 to 2014. The analysis focused on uncovering geographical variations in mortality rates, particularly for diabetes and CKD. The study revealed significant inequalities among US counties, emphasizing the need for targeted improvements in risk factors, access to medical care, and overall healthcare quality. Their work contributes to addressing public health challenges associated with these diseases and highlights the importance of understanding the underlying mortality disparities across different geographic regions.

# 5 Related Work

**Rodrigues, A. S., de Abreu, L. C., de Deus Morais, M. J., Leitao, F. N. C., do Amaral,**

**G. L. G., de Sousa Santos, E. F., do Souto, R. P. (2022). Temporal trend of mortality and hospitalization for chronic kidney disease in adults from Northern Brazil.** This study delves into the temporal trends of mortality and hospitalization for chronic kidney disease in young adults in the northern region of Brazil. Analyzing data between 1996 and 2017, it contributes valuable insights into regional patterns, age-specific trends, and hospitalization rates, complementing the understanding of chronic kidney disease dynamics.

To access the study, click https://doi.org/10.1097/MD.0000000000029702

**Li et al. (2023) conducted a study on the temporal trends of chronic kidney disease (CKD) in China from 1990 to 2019 using data from the Global Burden of Disease (GBD) 2019 study.** The prevalence and mortality rates of CKD increased significantly during this period, with a prevalence of 10.6 percent and a mortality rate of 13.8 per 100,000 in 2019. The average annual percentage change (AAPC) for prevalence and mortality was estimated at 1.6 percent and 1.8 percent respectively. The study also explored age, period, and cohort effects, finding that CKD risk increased with age and was higher in early birth cohorts. Females had a higher CKD prevalence but a lower mortality rate. Projections using an extended ARIMA model indicated a further increase in CKD prevalence (11.7 percent) and mortality rate (17.1 per 100,000) by 2029. The study recommends a comprehensive strategy, including primary care-level risk factor prevention, CKD screening among the elderly and high-risk populations, and improved access to high-quality medical services, to address the growing burden of CKD in China.

To access the study, click https://doi.org/10.1093/ckj/sfac218.

**Shieu, M., Morgenstern, H., Bragg-Gresham, J., Gillespie, B. W., Shamim-Uzzaman, Q. A., Tuot, D., Saydah, S., Rolka, D., Burrows, N. R., Powe, N. R., Saran, R. (2023). US Trends in Prevalence of Sleep Problems and Associations with Chronic Kidney Disease and Mortality.** This study investigates the relationship between self-reported sleep problems and chronic kidney disease (CKD) in the United States. Analyzing data from five biannual National Health and Examination Surveys (2005-2014) with 27,365 adult participants, the study reveals temporal trends and associations with CKD and mortality. The prevalence of trouble sleeping and sleep disorders increased over the study period, with higher rates in adults with CKD. Notably, excessive sleep and nocturia were associated with elevated mortality, suggesting potential implications for clinical practice in managing sleep problems among individuals with CKD.

To access the study, click https://doi.org/10.34067/KID.0000862019.

# 6 Methods

**Data Collection**

The data for this study was obtained from the National Center for Health Statistics (NCHS) through the data.gov platform. De-identified death records and population counts were sourced to conduct the analysis. The data spanned the period from 1980 to 2014.

To access the dataset on data.gov, click here.

**Data Mining Pipeline:** Data Cleaning

The dataset underwent thorough data cleaning to enhance clarity. Columns were renamed for improved readability. Columns with substantial data gaps were removed to create a more polished and refined dataset.

**Handling Missing Data.**

A notable challenge was encountered due to the absence of the "Year" attribute in the dataset. Exploratory analyses and adaptations were performed to address limitations stemming from incomplete data.

**Replication Process.**

The replication process involved retrieving the dataset from data.gov and replicating the methodology outlined in the original study by Mokdad et al. Challenges included the absence of the "Year" attribute, prompting adaptations and exploratory analyses.

**Model Evaluation Analyses:** Small Area Estimation Models.

Despite the promising aspects of small area estimation models, several challenges were encountered during the evaluation process:

**Data Quality and Completeness:** The accuracy of the small area estimation models heavily relies on the quality and completeness of the input data. Inconsistencies or missing information in the death records and population counts posed challenges during the modeling process.

**Model Sensitivity:** Small area estimation models are sensitive to the characteristics of the population under study. Variations in demographic structures and healthcare access among counties could impact the model's predictive accuracy.

**Age-Standardized Mortality Rates:** Challenges Encountered.

The calculation of age-standardized mortality rates encountered challenges related to age-specific data availability and consistency. Variations in age distribution among counties required careful adjustment to ensure accurate comparisons.

**Evaluation Metrics:** Evaluation of age-standardized mortality rates involved.

**Comparison with National Rates:** Assessing how well the age-standardized rates aligned with the national averages provided insights into the representativeness of the data.

**Trend Consistency:** Examining whether the age-standardized rates reflected expected temporal trends and patterns observed in the overall mortality rates.

**Spatial Analysis and Geographic Information Systems (GIS):** Challenges Encountered.

**Data Granularity:** Limited granularity in available data affected the precision of spatial analyses. Finer-scale data would enhance the identification of mortality clusters.

**GIS Integration:** Technical constraints in GIS integration required adaptations to traditional spatial analysis approaches.

**Cluster Identification:** GIS tools were employed to identify mortality clusters, evaluating the accuracy of identified hotspots.

**Software Used**

Used Python.

# 7 Model Description

**Statistical Models used in the original Study.**

In the original study, Bayesian spatially explicit mixed-effects regression models were utilized, conducting separate analyses for males and females. These models included various components such as intercepts, fixed covariate effects, random age-time effects, spatial effects, space-time effects, and space-age effects. The incorporation of seven covariates, including education levels, race/ethnicity, income, and population density, heightened the models' comprehensive predictive capacity. The random effects were modeled using conditional autoregressive distributions, enabling smooth transitions across adjacent age groups, years, and counties.

**Model Fitting in the Original Study.**

The model fitting in the original study was executed using the Template Model Builder Package in the R statistical software. This process involved generating a thousand draws of underlying mortality rates from the posterior distribution. Subsequently, the rates were standardized based on the 2010 U.S. census population.

# 8 Analysis

**Timeframe covered by the original study.**

The original study analyzed mortality patterns related to diabetes and chronic kidney disease (CKD) across U.S. counties from 1980 to 2014. Using death records from the National Center for Health Statistics and population counts and aimed for precise estimates of age-standardized mortality rates.

The analysis delved into factors like county, year, sex, and cause of death, particularly focusing on diabetes mellitus and CKD. This comprehensive approach provided nuanced insights into the temporal and geographical intricacies of mortality trends, highlighting disparities at the county level.

**Statistical Analysis Conducted in the Original Study.**

Several statistical and machine-learning analyses were conducted in the original study, as explained further below.

Small area estimation models were employed to enhance precision in predicting mortality rates at the county level. Age-standardized mortality rates were used to reveal comparative insights, adjusting for variations in age distribution across counties. Spatial analysis and Geographic Information Systems (GIS) were utilized to identify geographical patterns and visually represent mortality rates.

**Improving Precision with Small Area Estimation Models.**

The study enhanced precision in predicting mortality rates at the county level by employing small-area estimation models. These models played a crucial role in refining predictions, particularly in regions with limited sample sizes, aiming to address challenges associated with sparse data availability. The application of these models contributed to a more nuanced understanding of mortality patterns, especially in counties where data was scarce.

**Revealing Comparative Insights through Age-Standardized Mortality Rates.**

The original study incorporated temporal analysis into its methodology in addition to age- standardized mortality rates. This adjustment aimed to accommodate variations in age distribution across counties from 1980 to 2014, ensuring that observed differences in mortality rates were not solely influenced by disparities in population age structures. The standardized approach facilitated a more precise and equitable comparison of mortality trends across a diverse array of counties.

**Identifying Geographical Patterns with Spatial Analysis.**

Expanding beyond the temporal dimension, the research utilized Geographic Information Systems to perform spatial analysis and visually represent mortality rates. GIS provided a robust analytical framework for exploring geographical intricacies and revealing patterns, relationships, and trends in the data. Serving as a pivotal spatial tool, GIS played a crucial role in mapping and dissecting spatial data, facilitating the identification of clusters or disparities in mortality rates among U.S. counties.

**Differences in Replication Approach between my Work and the Original Study.**

In my study, the initial phase involved retrieving the dataset from data.gov, followed by thorough data cleaning aimed at enhancing data clarity. The columns of the dataset were then renamed for improved readability, and to address missing values, I removed columns with substantial data gaps, resulting in a more polished and refined dataset. In my replication, the challenge of absence of the "Year" attribute in the dataset emerged. This omission significantly hindered my ability to reproduce clear trends related to chronic kidney disease (CKD) mortality rates and age-standardized mortality rates for specific years (1980, 1990, 2000, and 2014).

**Temporal and Geographical Scope.**

The analysis spanned from 1980 to 2014, examining mortality patterns related to diabetes and chronic kidney disease (CKD). Focus on county, year, sex, and cause of death provided nuanced insights into temporal and geographical intricacies of mortality trends.

**Statistical and Machine-Learning Analyses.**

Small area estimation models were employed to enhance precision in predicting mortality rates at the county level. Age-standardized mortality rates were used to reveal comparative insights, adjusting for variations in age distribution across counties. Spatial analysis and Geographic Information Systems (GIS) were utilized to identify geographical patterns and visually represent mortality rates. Replication Process: My Approach

**Data Retrieval and Cleaning.**

Retrieved dataset from data.gov. Conducted thorough data cleaning for enhanced clarity. Renamed columns for improved readability.

**Handling Missing Values.**

Removed columns with substantial data gaps. Resulted in a polished and refined dataset.

**Challenges Faced.**

Absence of "Year" attribute in the dataset. Hindered reproduction of CKD mortality rates and age-standardized rates for specific years (1980, 1990, 2000, and 2014).

**Technical Expertise Limitation.**

Limited understanding of the original study's methodologies. Challenges in data analysis, such as utilizing the cause list from the Global Burden Disease (GBD) Study and reallocating deaths from "garbage codes."

**Statistical Test Limitations.**

Despite diligent data cleaning, unable to execute statistical tests. Limitations due to available data and the complexity of statistical methods in the original study.

As a result of the above-stated challenges, and despite the necessary data cleaning and wrangling, I ended up being unable to execute statistical tests. This limitation stemmed from both the available data and the intricate nature of the statistical methods employed in the original study, preventing a comprehensive statistical analysis.

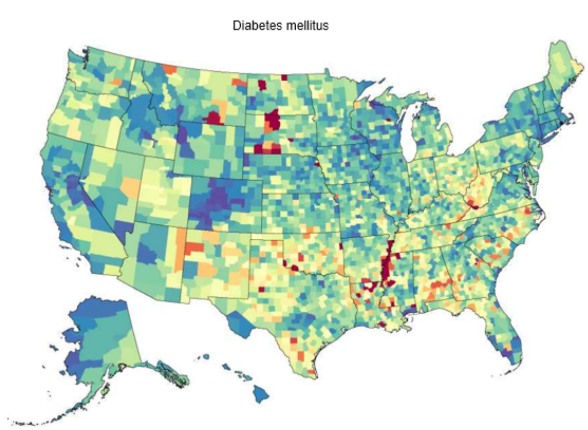
# 9 Results

**Findings from my Study.**

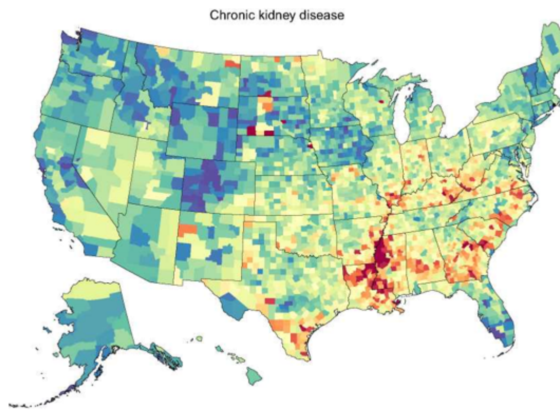
In this report, I initiated my analysis by importing a dataset from data.gov into google colab and activated the essential packages. The dataset contains 236,124 records with 12 attributes. Data cleaning involved renaming columns for clarity and addressing missing values, resulting in a refined dataset. Specific components of the original study were replicated, including Table 1, summarizing national deaths, years of life lost, age-standardized mortality rates, and the distribution of age-standardized mortality rates in 2014.

| Cause of death | National deaths, YLLs, and mortality rate | | | County-level mortality rates | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Deaths, no. in thousands | YLLs, no. in Thousands | Mortality rate, no. of deaths/100,000 population | No. of deaths/100,000 population: Minimum | 10th percentile | Median | 90th percentile | Maximum | 90th-10th percentiles, no. of deaths/100,000 population[a] | Rat io of 90th-10t h percentile, ratio[b] |
| Diabetes mellitus | 66.2 (63.6–69.1) | 1182.5 (1 1 33.8–1231.1) | 19.7 (18.9–20.6) | 2.4 | 13.8 | 21.3 | 33.8 | 118.7 | 20.0 | 2.4 |
| Chronic kidney disease | 75.0 (71.7–78.1) | 1067.8 (1028.3–1104.6) | 22.4 (21.4–23.3) | 4.9 | 16.2 | 24.1 | 35.1 | 70.2 | 18.9 | 2.2 |
| … .Due to diabetes mellitus | 29.6 (27.1–33.3) | 496.1 (461.0–540.9) | 8.8 (8.1–9.9) | 2.1 | 6.5 | 9.8 | 14.2 | 29.4 | 7.7 | 2.2 |
| …Due to hypertension | 38.2 (34.9–41.0) | 489.4 (449.8–525.0) | 11.4 (10.4–12.3) | 2.1 | 7.5 | 11.9 | 18.5 | 41.3 | 11.0 | 2.5 |
| … .Due to glomerulo-nephritis | 6.6 (57–7.5) | 75.6 (68.1–83.5) | 2.0 (1.7–2.2) | 0.6 | 1.5 | 2.0 | 2.7 | 6.2 | 1.2 | 1.8 |
| … .Due to other causes | 0.6 (0.4–0.7) | 6.7 (5.6–8.0) | 0.2 (0.1–0.2) | 0.1 | 0.1 | 0.2 | 0.4 | 2.2 | 0.2 | 2.7 |

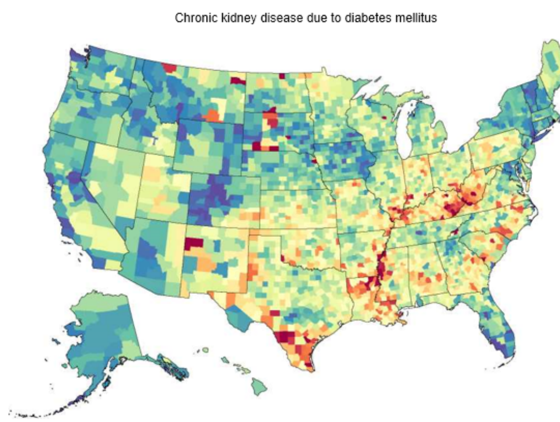
Table 1: National Deaths, Years of Life Lost (YLLs), and Age-Standardized Mortality Rate; and Distribution of Age-Standardized Mortality Rates at the County Level (2014)



Age-Standardized Mortality Rate, 2014 of deaths per 100,000 population




Age-Standardized Mortality Rate, 2014 of deaths per 100,000 population




Age-Standardized Mortality Rate, 2014 of deaths per 100,000 population

**Main challenge Encountered.**

Incomplete Data.

**Comparison of Original Results.**

The table from my analysis and other results did not match the results of the original study due to the reasons highlighted in the discussion section of my report. In my attempt to replicate the original study, I encountered significant challenges due to incomplete data and the complexity of the original study's statistical methods. To address these limitations and match the original results more closely, I considered various additional methods. These included data imputation for missing values, using Bayesian spatial modeling, exploring alternative data sources, conducting sensitivity analyses, and extensive exploratory data analysis. While these methods offered potential avenues for addressing the limitations, the absence of essential data attributes, such as "Year," remained a substantial hindrance in effectively replicating the original study's results.

# 10 Discussion.

The investigation into trends and patterns of disparities in diabetes and chronic kidney disease (CKD) mortality across U.S. counties from 1984 to 2014, presents a significant contribution to our understanding of the geographical variations in health outcomes. While the study holds merit in shedding light on the impact of these diseases at the county level, it is crucial to discuss the significance, limitations, and problems encountered during the analysis.

**Significance of the Study.**

Identified geographical variations in diabetes and CKD mortality which enabled targeted interventions and resource allocation for high-risk communities.

Highlighted a 33.6 percent rise from 1980 to 2000 and a subsequent 26.4 percent decrease from 2000 to 2014.

Established Patterns of CKD Mortality and Underlying factors by exploring CKD mortality patterns and correlations with diabetes, hypertension, and glomerulonephritis.

Highlighted the importance of the Application of Small Area Estimation Models in enhancing precision in mortality rate predictions at the county level by providing a robust foundation for uncovering disparities.

**CKD Mortality rates in 2014.**

Extremely low mortality rates were noted in central Colorado, along with select counties in southern Florida, California, and Great Plains states.

Conversely, elevated mortality rates were identified in counties spanning much of the Deep South, with a concentration of particularly high rates observed around the Mississippi River.

Summit County, Colorado, stood out with the lowest estimated mortality rate in 2014.

**Change in CKD Mortality rates between 1980 and 2014.**

The mortality rate attributed to chronic kidney disease (CKD) exhibited relative stability between 1980 and 1990.

However, a noteworthy surge of 50.1 percent occurred between 1990 and 2014, across the same period.

Counties experiencing the most substantial increases were primarily situated in western Oregon, Iowa, and Minnesota, southern Illinois, as well as parts of Texas, Tennessee, Kentucky, and West Virginia.

**Limitations of the Study of the Replication study.**

**Incomplete Data and Missing Attributes:** The absence of the "Year" attribute hinders the replication of CKD mortality trends for specific years.

**Technical Expertise and Methodological Challenges.**

Limitations in technical expertise hindered full study replication.

Challenges in reallocating deaths from "garbage codes" and using cause lists from the Global Burden of Diseases Study.

**Regional Disparities in Mortality Rates.**

The research concentrates on revealing geographical variations in age-standardized mortality rates related to diabetes mellitus and chronic kidney disease (CKD), offering valuable insights. By pinpointing areas with higher mortality rates, like those along the southern stretch of the Mississippi River. This enabled targeted interventions and strategic allocation of resources to cater to the distinct requirements of communities facing elevated risks.

**Patterns of Chronic Kidney Disease (CKD) Mortality and Underlying Factors.**

The investigation into mortality patterns related to chronic kidney disease (CKD), along with its correlation with factors like diabetes, hypertension, and glomerulonephritis, adds depth to our comprehension. This knowledge is pivotal for customizing preventive measures and healthcare interventions, aiming for improved outcomes.

**Application of Small Area Estimation Models.**

Employing validated small-area estimation models enhances the precision of mortality rate predictions at the county level. This approach allows for a more nuanced analysis, especially in regions with limited sample sizes. This provided a robust foundation for uncovering disparities.

**Incomplete Data and Missing Attributes.**

The absence of the "Year" attribute in the dataset posed a significant challenge in replicating trends related to CKD mortality rates and age-standardized mortality rates for specific years. This limitation affects the ability to faithfully reproduce the temporal aspects of the original study.

**Statistical Analysis Constraints.**

Despite meticulous data cleaning and wrangling efforts, challenges arose in conducting statistical tests. The utilization of Bayesian Spatial Modeling, a sophisticated statistical method in the original study, exceeded our initial understanding and expertise. The complexity of these statistical methods, coupled with constraints in available data, posed obstacles to achieving a thorough and comprehensive statistical analysis.

# 11 Conclusion

In conclusion, my study reveals significant divergences in diabetes and chronic kidney disease (CKD) mortality rates across U.S. counties from 1980 to 2014. Employing Bayesian spatially explicit mixed-effects regression models facilitated the carrying out of the analysis by considering various components like intercepts, fixed covariate effects, and random age-time effects. The inclusion of seven covariates, such as education levels, race/ethnicity, income, and population density in the study heightened the models' predictive capacity. Using the Template Model Builder Package, the study executed the model fitting process by generating a thousand draws of mortality rates, subsequently standardizing them based on the 2010 U.S. census population.

The research uncovers critical insights into these variations, emphasizing the urgency of targeted interventions addressing access to medical care, lifestyle factors, and dietary patterns. As the U.S. population ages and medical costs rise, my study underscores the crucial role played by Bayesian spatial models in discerning geographic disparities. Beyond deepening our understanding of mortality patterns related to diabetes and CKD, this research lays the foundation for informed public health initiatives capable of effectively tackling the evolving challenges present in diverse U.S. counties. This study exposed significant disparities in diabetes and chronic kidney disease (CKD) mortality rates across U.S. counties. Leveraging Bayesian spatially explicit mixed-effects regression models enabled a nuanced analysis, considering intercepts, fixed covariate effects, and random age-time effects.

Critical insights emerged, highlighting the pressing need for targeted interventions addressing access to medical care, lifestyle factors, and dietary patterns. With the aging U.S. population and escalating medical costs, our study underscores the indispensable role of Bayesian spatial models in discerning geographic disparities.

Finally, this project aimed to explore trends and patterns of disparities in diabetes and chronic kidney disease (CKD) mortality across U.S. counties from 1984 to 2014. Utilizing data mining, statistical analyses, and machine learning models, the study revealed significant geographical variations and temporal trends in mortality rates. Bayesian spatial models provided nuanced insights, emphasizing the need for targeted interventions to address health disparities.

### 11.1 Limitations or Shortcomings.

**Incomplete Data:** The absence of the "Year" attribute in the dataset posed a major challenge in replicating CKD mortality trends for specific years, impacting the temporal aspects of the study.

**Technical Expertise Challenges:** Limitations in technical expertise hindered the full replication of the original study. Challenges in reallocating deaths from "garbage codes" and using cause lists from the Global Burden of Diseases Study were encountered.

**Statistical Complexity:** Despite diligent data cleaning, the complexity of the original study's statistical methods posed challenges, limiting the execution of certain statistical tests.

### 11.2 Future Work and Project Extensions.

**Data Enhancement:** Future research should focus on obtaining datasets with complete attributes, particularly the inclusion of the "Year" attribute, to ensure a more accurate and comprehensive analysis.

**Methodological Refinement:** Collaboration with experts or additional training in specific statistical methods is recommended to address challenges in technical expertise and methodological complexity.

**Integrating Additional Factors:** Consideration of additional factors, such as socioeconomic status, healthcare infrastructure, and genetic predispositions, could provide a more holistic understanding of disparities in diabetes and CKD mortality.

**Validation and Comparison:** Conducting validations against other datasets or comparing results with similar studies can enhance the robustness and reliability of findings.

While this study provides valuable insights into the disparities in diabetes and CKD mortality, its limitations highlight the need for continuous improvement. Future work should aim to address these limitations, refine methodologies, and explore additional factors for a more comprehensive understanding of health disparities in the U.S. counties.

## 12 Data and Software Availability

**Data Availability**

The dataset used in this study can be accessed from the National Center for Health Statistics (NCHS) through the data.gov.

To access the dataset on data.gov, click here.

**Software Availability**

The code and scripts developed for this project can be accessed through:

To view my Google Colab notebook, click here.

## 13 References

1. Rodrigues, A. S., de Abreu, L. C., Morais, M. J. D., Leitao, F. N. C., do Amaral, G. L. G., de Sousa Santos, E. F., do Souto, R. P. (2022). Temporal trend of mortality and hospitalization for chronic kidney disease in adults from Northern Brazil. Medicine, 101(26), e29702.

To access the study, click https://doi.org/10.1097/MD.0000000000029702

2. Li, Y., Ning, Y., Shen, B., Shi, Y., Song, N., Fang, Y., Ding, X. (2022). Temporal trends in prevalence and mortality for chronic kidney disease in China from 1990 to 2019: an analysis of the Global Burden of Disease Study 2019. Clinical kidney journal, 16(2), 312–321.

To access the study, click https://doi.org/10.1093/ckj/sfac218.

3. Shieu, M., Morgenstern, H., Bragg-Gresham, J., Gillespie, B. W., Shamim-Uzzaman, Q. A., Tuot, D., Saydah, S., Rolka, D., Burrows, N. R., Powe, N. R., Saran, R., Centers for Disease Control and Prevention Chronic Kidney Disease Surveillance Team (2020). US Trends in Prevalence of Sleep Problems and Associations with Chronic Kidney Disease and Mortality. Kidney360, 1(6), 458–468.

To access the study, click https://doi.org/10.34067/KID.0000862019.