Effect of Dietary Patterns on Chronic Kidney Disease (CKD) Measures (ACR), and on the Mortality of CKD Patients

by

Sayed Ahmed
BSc Eng. in Computer Science and Engineering, BUET, Bangladesh, 2001
MSc in Computer Science, University of Manitoba, Canada, 2006

A Major Research Paper
Presented to Ryerson University
in partial fulfillment of the requirements for the degree of

Master of Science
in the Program of
Data Science and Analytics

Toronto, Ontario, Canada, 2019

Supervisor
Dr. Youcef Derbal
Information Technology Department
Ted Rogers School of ITM
Ryerson University

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A MAJOR RESEARCH PROJECT (MRP)

I hereby declare that I am the sole author of this Major Research Paper. This is a true copy of the MRP, including any required final revisions.

I authorize Ryerson University to lend this MRP to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this MRP by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my MRP may be made electronically available to the public.

Sayed Ahmed

Effect of Dietary Patterns on Chronic Kidney Disease (CKD) Measures (ACR), and on the Mortality of CKD Patients

Sayed Ahmed

Master of Science 2019

Data Science and Analytics

Ryerson University

Abstract

Chronic Kidney Disease (CKD) leading to End-Stage Renal Disease (ESRD) is very prevalent today. Over 37 millions of Americans have CKD. CKD/ESRD and interrelated diseases cause a majority of the early deaths. Many researchers studied the effect of drugs on CKD. Little research studied the effect of whole dietary patterns. This research has identified the effect of dietary patterns on CKD mortality; also on a CKD measure named Albumin Creatinine Ratio (ACR). Dietary surveys from NHANES, and CKD Mortality dataset from USRDS were utilized. Principal Component Analysis and Regression were utilized to find the effect. Machine Learning Approaches including Regression, and Bayesian were applied to predict ACR values. Grains, Other Vegetables showed positive correlations with Mortality where Alcohol, Sugar, and Nuts showed negative correlations. ACR values were not found strongly correlated with dietary patterns. For ACR value prediction, 10 Fold Cross Validations with Polynomial Regression showed 95% accuracy.

Keywords:

CKD, ESRD, Dietary Patterns, Mortality, ACR, Polynomial Regression, Bayesian

Acknowledgements

I am very grateful to Professor Dr. Youcef Derbal to supervise this research project towards my MSc. Dr. Derbal has been a great support throughout the work. Dr. Derbal has helped to formalize the problem, referred to valuable ideas, articles, and research papers, and provided valuable feedback. I am grateful for the guide and direction that I have received from Dr. Derbal.

Thank you, Professor Dr. Derbal.

Contents

List of Figures

List of Tables

1 Introduction

Chronic kidney disease (CKD) is very prevalent in today's world and CKD incidents are continually increasing such as 10 to 13% of the US population get affected by Chronic Kidney Disease. CKD/ESRD and other interrelated diseases such as Hypertension, Heart Diseases, and Diabetes cause a majority of the early deaths [31]. In addition to kidney failure, CKD is also a major cause of death from stroke, and heart diseases. On the other hand, hypertension and diabetes are also major causes of CKD. CKD is not reversible, and is progressive that gradually reduces kidney function.

Primary causes of CKD include High Blood Pressure, and Diabetes. Other causes include infection, kidney stones, genetics, genetical polycystic kidney disease, certain foods and food habits, pain killers, and drug usage/abuse. As CKDs are not curable and reversible controlling diabetes and blood pressure with or without medication can slow the progress of CKDs in most cases. As Kidneys filter waste products and our diet produces those waste products controlling diet have an effect on how much work kidney has to perform, and how well the kidney will function. Studies show that drugs as well as lifestyle choices (food, diet, exercise) can prevent CKD, slow the progression of CKD [29], delay dialysis and kidney transplantation; consequently can prevent early deaths. Though there are many studies on the effect of drugs to control CKD and related complications, there are few studies on the effect of diets, dietary patterns, and lifestyles [29]. There are studies in how controlling nutrients/chemicals in food items can help to prevent or to slow the progression of CKD. CKD patients are also provided recommendations on certain chemicals or food items. However, adhering to the recommended amount of nutrients and/or food item is challenging. Hence, there is an emerging trend where the effect is studied utilizing dietary patterns with food groups and food subgroups rather than nutrients/chemicals in food or individual food item. This research analyzes the effect of dietary patterns using food groups and food subgroups on the mortality and survival of CKD patients. This research has also identified how ACR values are associated with dietary patterns. Additionally, ACR values are predicted based on a dataset utilizing machine learning approaches.

1.1 How CKD is Identified and Measured

CKD is identified with one of two measures such as a blood test named Glomerular Filtration Rate (GFR) or a urine test named Albumin Creatinine Ratio (ACR). ACR values less than 30 indicates no CKD or mild CKD. ACR values between 30 and 300 indicate moderate CKD. ACR values >300 indicates severe CKD. ACR values when are consistent for three months and fall within the range as above, the patients are diagnosed as CKD patients. GFR is measured in ml/min/1.73 m2. CKDs as measured with GFRs are described in stages such as Stage 1 with normal or high GFR (GFR greater than or equal to 90 mL/min), Stage 2 with Mild CKD (GFR = 60-89 mL/min), Stage 3A with Moderate CKD (GFR = 45-59 mL/min), Stage 3B with Moderate CKD (GFR = 30-44 mL/min), Stage 4 with Severe CKD (GFR = 15-29 mL/min), Stage 5 with End Stage CKD (GFR <15 mL/min) [5]. At stage 5, patients loss complete kidney function then either require dialysis or transplantation to survive.

2 Literature Review

Kidney patients commonly are given dietary advice based on individual nutrients or chemicals primarily or sometimes on food items instead of whole eating patterns. However, that advice is challenging to adhere to for the majority of the patients [2]. Also, there is limited evidence that adherence to such advice prevents clinical complications [23]. Hence, studying the whole dietary patterns rather than single nutrient or food group restrictions is an emerging trend for CKD/ESRD patient diets [2] [24-26]. This is also easier to adhere to. There are several studies on analyzing the relation between dietary patterns and clinical outcomes for CKD patients [3, 4, 5, 6, 7, 8, 9, 26].

Chen at al [3] studied the association of plant protein intake for all cause mortality in CKD. In the study higher plant protein ratio was found to cause lower mortality for CKD patients in stage 3 or higher (eGFR <60 ml/min/1.73 m2) though not for others (stage 1 and 2) [3]. This study primarily used statistical methods and Regression Models such as Cox regression models to find the association [3]. Hao-Wen et al [26] studied the association between vegetarian diets and CKD. The study found that vegetarian diets including vegan and ovo-lacto vegetarian diets were possible protective factors. The study utilized The multivariable logistic regression analysis [26].

Gutiérrez et al [4] studied 5 empirically derived dietary patterns such as ”convenience” (Chinese and Mexican foods, pizza, and other mixed dishes), ”plant-based” (fruits and vegetables), ”sweets/fats” (sugary foods), ”Southern” (fried foods, organ meats, and sweetened beverages), and ”alcohol/salads” (alcohol, green-leafy vegetables, and salad dressing) [4]. The study found that dietary pattern rich in processed and fried foods was associated with higher mortality in persons with CKD. On the other hand, a diet rich in fruits and vegetables was found to be protective [4].

Huang et al. [5] studied whether Mediterranean diet can preserve kidney function along with maintaining favorable cardiometabolic profile with reduced mortality risk for individuals with CKD. The study found that adhering to Mediterranean diet has a lower likelihood of having CKD in elderly men. The study also found that a greater adherence to this diet can improve survival for CKD patients [5]. Huang et al [5] in the above study, used unpaired t test, nonparametric Mann–Whitney test, or chi-square test as appropriate for Comparisons between CKD and non-CKD men. To evaluate the association of Mediterranean diet with the presence of CKD, Crude and multiple adjusted logistic regression models were fitted. All tests were two-tailed, and $P < 0.05$ was considered significant [5]. One aspect of Muntner et al [6] study was to find out how Life's Simple 7 factors (Smoke, Activity, BMI, Diet, Blood Pressure, Cholesterol, and Glucose) affect in getting ESRD. The study shows that people who have high/ideal scores in more of these factors have lower likelihood of getting ESRD. This study utilized Cox proportional hazards models. Adjustment were made for age, race, sex, stroke-based geographic region of residence, income, education, and history of stroke or coronary heart disease [6].

Ricardo et al [7] studied the association of death to healthy lifestyles esp. in relation to CKD. The study found that adherence to healthy lifestyles was associated with lower risk of all cause mortality in CKD patients. In this study, to determine the association between a healthy lifestyle and survival among individuals with CKD, Cox proportional hazards models were used while also adjusting for important covariates. Stratified survival analyses by eGFR and UACR was performed for Sensitivity analyses [7]. Suruya et al. [8] studied dietary patterns in hemodialysis patients in Japan and researched associations between dietary patterns and clinical outcomes. The study found that patients with unbalanced diet were more likely to have adverse clinical outcomes. Hence, such patients when in addition to portion control, maintains a well-balanced diet esp. for the food groups meat, fish, and vegetables will have less adverse clinical outcomes [8]. Suruya et al [8] utilized a principal components analysis (PCA) with Promax rotation to reduce to a smaller set of food groups for analysis. PCA was used to find food groups eaten with equal frequencies [8]. Cox regression model was used for the analysis with multiple models where each model had a different combination of covariants [8].

Another study by Ricardo et al [9] estimated the degree of adherence to a healthy lifestyle that decreases the risk of renal and cardiovascular events among adults with chronic kidney disease (CKD). The study found that adherence to a healthy lifestyle was associated with lower all-cause mortality risk in CKD. The greatest reduction in all-cause mortality was related to nonsmoking [9]. This study by Ricardo et al [9], to compare categorical and continuous variables used Chi-squared and analysis of variance tests respectively. To examine the association between healthy lifestyle and outcomes, Cox proportional hazards models were used. Death was treated as a censoring event. Three nested Cox proportional hazards models were fitted and were adjusted sequentially for potential explanatory variables [9].

G. Asghari et al studied the association of population-based dietary pattern with the risk of incident CKD The study concluded that high fat and high sugar diet pattern is associated with significantly increased (46%) odds of incident CKD where a lacto-vegetarian diet can be protective of CKD by 43%. The study utilized multivariable logistic regression to calculate odds ratio for the association.

final: Most of the studies primarily used direct clinical data of patients for several years and applied statistical analysis primarily. This research primarily utilized public datasets from CDC, USRDS. One of the studies above utilized the dietary pattern data from CDC and NHANES like this study. However, this study will differ in the methodology, exploration, and analysis. This study is finding relations between datasets from multiple sources and is focused on finding patterns and relations in general population than specific/selected individuals. Most of the studies above utilized primarily statistical methods and sensitivity analysis where primarily regression models esp. Cox regression models were used. In a couple of cases, Principal Component Analysis (PCA) was used. Primarily direct clinical data of patients for several years were studied in majority of the projects. This research primarily utilized public datasets from CDC, USRDS, Health.gov. For association and prediction, PCA, Regression, and several machine learning approaches are heavily utilized. For food groups and subgroups utilized the USDA categorization. Recommended amounts for food groups provided by CDC/Health.gov is used. Additionally, ACR values are predicted based on a dataset utilizing machine learning approaches. The machine learning approaches used for ACR value prediction are: Regression and Bayesian with or without 10 fold crosss validations.

Table 1: **Regression Output from Exploratory Analysis using Excel (Data Analysis Module)**

Food Groups: Multiple R: 0.880156954, R Square: 0.774676264, Adjusted R Square: 0.616949648, Standard Error: 6.223447127	**Food Subgroups:** Multiple R: 0.999378722, R Square: 0.99875783, Adjusted R Square: 0.978883104, Standard Error: 1.461167293, Observations: 18, R Square: 99%

3 Exploratory Study

Initial exploratory analysis aspects are provided in this section. Such studies and analysis resulted the methodologies utilized and experiment design. Initially Datasets from USRDS on CKD and ESRD patients were explored for Patient characteristics, prevelance of CKD, dietary intake, dietician care, mortality, and survival. Dietary intake data were missing though dietician care data, mortality and survival data by age groups were there; however, No linking data was there for dietician care data and target mortality and survival data. Hence, datasets from an study on dietary shift recommendation by CDC was explored where recommended intake amounts for each age groups including average intake amount (of food groups and food subgroups) by age groups were provided. However, Age groups between Dietary shift recommendation dataset and USRDS mortality i.e. target variable data were not aligned. Hence, dietary intake data from NHANES survey was explored. The same survey data was used by Dietary shift recommendation dataset. Hence, regrouped NHANES survey data to reflect the USRDS age groups. USDA codes wer used for Food groups/subgroups/intake food for NHANES survey. Hence, survey food intake was mapped using USDA food codes. Shift recommendation dataset along with additional data from CDC and other sources were explored to create food group recommended amounts for matching age groups. Initial exploratory analysis such as univariate analysis, bivariate analysis, Regression, and Factor analysis were done. Representative exploratory analysis output and plots from the initial analysis are given in Figure 1. However, the output and plots are primarily exploratory that may or may not reflect the final outcome.

4 Methodology

In this section, the methodology utilized for the research will be provided. A diagram showing all the steps in the methodology is provided in Figure 2. The primary purpose of this research is to assess the effect of dietary patterns on the cause (ACR - Albumin Creatinine Ratio) as well as the mortality and survival in relation to chronic kidney disease (CKD)/End Stage Renal Disease (ESRD).

4.1 Methodology Overview

A dataset released by CDC/Health.gov with the dietary habits and a CKD measure named Albumin Creatinine Ratio (ACR) of around 10,000 individuals are studied. Also, age group based mortality and survival of CKD/ESRD patients provided by USRDS are studied. Afterwards, utilized Principal Component Analysis to identify the most important food groups and subgroups affecting the ACR value

		Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
	Intercept	44.479	95.829	0.464	0.652	-169.041	257.999	-169.041	257.999
Actual Vegetable Intake	**X Variable 1**	**0.191**	**0.165**	**1.155**	**0.275**	**-0.177**	**0.559**	**-0.177**	**0.559**
Actual Fruit intakes	X Variable 5	0.031	0.169	0.183	0.858	-0.345	0.407	-0.345	0.407
Avg Fats oils and salad dressings take	X Variable 7	0.012	0.978	0.013	0.990	-2.168	2.192	-2.168	2.192
Actual Taken Sugars sweets and beve	**X Variable 6**	**-0.015**	**0.013**	**-1.118**	**0.290**	**-0.045**	**0.015**	**-0.045**	**0.015**
Actual Protein Intake	X Variable 2	-0.023	0.154	-0.152	0.882	-0.367	0.320	-0.367	0.320
Actual Dairy Intake	X Variable 4	-0.085	0.099	-0.852	0.414	-0.306	0.137	-0.306	0.137
Actual Grain Intake	**X Variable 3**	**-0.086**	**0.066**	**-1.298**	**0.223**	**-0.233**	**0.062**	**-0.233**	**0.062**

Regression on Actual Food Group Intake

		Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
	Intercept	1.372	54.715	0.025	0.984	-693.852	696.595	-693.852	696.595
Actual Starchy vegetables Intake	**X Variable 3**	**0.310**	**0.135**	**2.299**	**0.261**	**-1.404**	**2.024**	**-1.404**	**2.024**
Actual Other vegetables Intake	**X Variable 4**	**0.159**	**0.069**	**2.287**	**0.262**	**-0.722**	**1.039**	**-0.722**	**1.039**
Avg Meat, Poultry and Eggs subgroup taken	X Variable 7	0.144	0.111	1.304	0.416	-1.263	1.552	-1.263	1.552
Actual Dark-green vegetables Intake	X Variable 1	0.098	0.121	0.807	0.568	-1.443	1.639	-1.443	1.639
Avg Added Sugars/Sugars and sweets taken	X Variable 10	0.047	0.091	0.520	0.695	-1.110	1.205	-1.110	1.205
Actual Whole grains intakes	X Variable 5	0.036	0.114	0.312	0.807	-1.416	1.488	-1.416	1.488
Actual Taken Refined grains amount	X Variable 6	0.027	0.108	0.255	0.841	-1.339	1.394	-1.339	1.394
Avg Nonalcoholic beverages taken	X Variable 16	0.014	0.010	1.414	0.392	-0.115	0.143	-0.115	0.143
Avg Milks and milk drinks taken	X Variable 13	-0.001	0.079	-0.008	0.995	-1.009	1.008	-1.009	1.008
Avg Alcoholic beverages intake	**X Variable 15**	**-0.016**	**0.008**	**-1.949**	**0.302**	**-0.118**	**0.087**	**-0.118**	**0.087**
Avg Water, noncarbonated intake	**X Variable 14**	**-0.032**	**0.014**	**-2.288**	**0.262**	**-0.211**	**0.147**	**-0.211**	**0.147**
Avg Seafood taken	X Variable 8	-0.052	0.071	-0.742	0.594	-0.951	0.846	-0.951	0.846
Avg Nuts, Seeds, and Soy Products taken	**X Variable 9**	**-0.153**	**0.068**	**-2.245**	**0.267**	**-1.016**	**0.711**	**-1.016**	**0.711**
Actual Red and orange vegetables Intake	**X Variable 2**	**-0.245**	**0.075**	**-3.248**	**0.190**	**-1.202**	**0.712**	**-1.202**	**0.712**
Avg Solid Fats taken	**X Variable 12**	**-0.583**	**0.389**	**-1.497**	**0.375**	**-5.528**	**4.363**	**-5.528**	**4.363**
Avg Oils taken	X Variable 11	-0.594	0.668	-0.889	0.538	-9.087	7.899	-9.087	7.899

Regression on Actual Food Subgroup Intake

Line Plot for Regression outcome

Heatmap for exploration

Figure 1: **Exploratory Analysis and Representative Output**

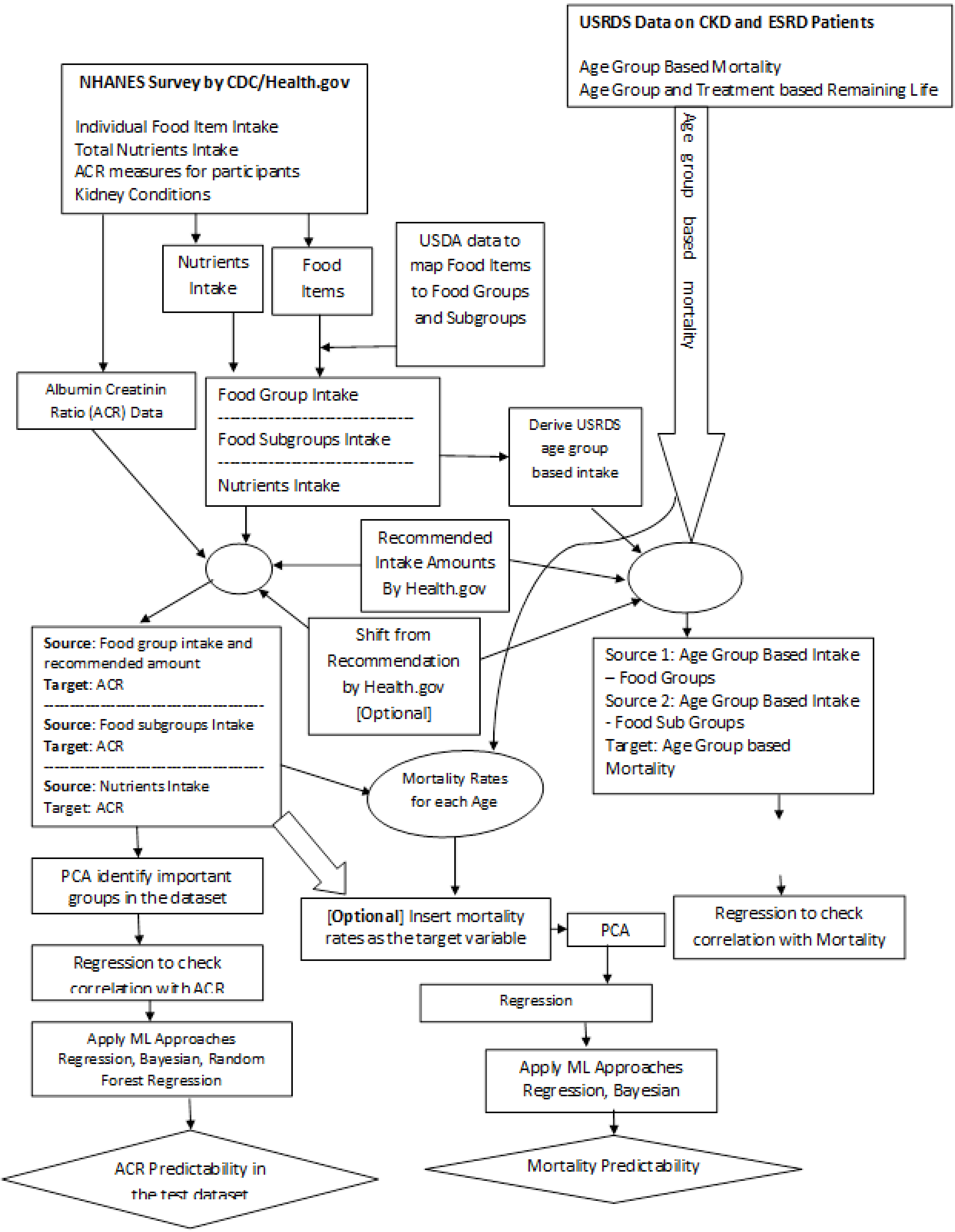


Figure 2: **Methodology in a Diagram**

and the Mortality/Survival then utilized Statistical Regression and Factor analysis to understand the correlation of ACR, and Mortality/Survival to dietary patterns. Machine learning approaches such as Regression, Polynomial Regression, and Bayesian with or without 10 fold cross validations are applied on the datasets to understand if dietary patterns can be used to predict ACR values and mortality.

Additionally, dietary recommendations as provided by health.gov are utilized for food groups and subgroups. Association of CKD mortality and ACR values with ratios from the recommended high amount is studied. Another study [11] on shifting from current recommendations conducted by health.gov is explored. Will the recommended shift [11] from current diet style [15] can have an improved outcome or not is also explored.

4.2 Study Selection

For dietary patterns, CKD measures (such as Albumin Creatinine Ratio - ACR) , and Kidney condition measures, a dataset from the National Health and Nutrition Examination Survey on dietary habits conducted by the Centers for Disease Control and Prevention (CDC) [10] was used. The survey has data from 1996 to 2016 [10]. This study primarily utilized data for 2015-2016. The survey recorded 24 hours individual food item intake amount. Two surveys were taken within 3 to 10 days apart. Each survey provided food item intake amount in a day, also mentioned the diet style, and diet-restrictions. Individual food items are represented using USDA food code. The survey also provided total nutrients data. CDC also released examination, laboratory, demographics, and other related data for those participants. This study explored and utilized examination data such as Kidney Condition data, laboratory data such as ACR data & Blood Pressure data, and demographics data for age.

For mortality and survival information, dataset from the United States Renal Data System (USRDS) on CKD and ESRD [16, 17] was utilized. “USRDS investigates the transition of care from CKD to ESRD and end-of-life care for those with advanced kidney disease” [19]. USRDS also releases data on the Incidence, Prevalence, Patient Characteristics, and Treatment Modalities on CKD, and ESRD patients. USRDS reports the survival and mortality using metrics such as Mortality rates: ESRD patients, Mortality rates: Dialysis patients, Total Mortality Count, 90 day survival for dialysis and/or transplantation patients, 10 year survival for dialysis and/or transplantation patients, Avg. Expected remaining lifetime with or without pre-condition and treatment options used. The data are either aggregated or patient specific detail data. However, only aggregated data are public where patient specific data access requires special request and permission. This research utilized only the public dataset i.e. age-group based aggregated data. In couple of experiments, aggregated age-group based data from USRDS was mapped to specific age for NHANES data for each ages and participants.

The dietary survey data (NHANES) represented the food items taken by the participants using USDA food codes [14, 12, 13]. Hence, USDA food codes [14, 12, 13] are used to assign food groups and subgroups to the NHANES [10] survey data to properly group/subgroup the dietary intake of the participants with some customizations.

Table 2: **Set 1: Mortality and CKD: Food Groups**

Primary Input Dataset:	NHANES survey aggregated to calculate average food item intake by USRDS like age groups
Target Variable:	ESRD: Avg. Annual Mortality rates
Experiment 1.1:	Identify contributing and important food groups in the dataset using PCA a) Using Actual Intake Amount b) Using ratios of intake and recommended high
Experiment 1.2:	Find out correlation (using Pearson's correlation and regression) between CKD mortality and important food groups as found using PCA in experiment 1. a) Using Actual Intake Amount b) Using ratios of intake amount and recommended high

4.3 Data Synthesis

NHANES survey data as provided for two days are averaged to get the intake amount for one day. Both individual food item data and nutrients intake data are averaged. USDA food codes are used to map food items to food groups and subgroups.

ACR and Kidney condition data for each individual are merged with the averaged food groups, subgroups, and nutrients data. ACR values are used as the target variable for couple of experiments and study. With this data food group recommendations from health.gov were also merged. This dataset was used for ACR association study. For one mortality study mortality rates from USRDS for each age were merged with the above data. Mortality is used as the target variable.

For both of the above cases, Principal Component Analysis (PCA) was applied to find out important food groups and subgroups. Afterwards, Regression was applied to find association with ACR and Mortality. Afterwards, Machine Learning approaches such as Linear Regression, Polynomial Regression, Random Forest Regression, Bayesian prediction with or without 10 fold cross validations were applied to study the predictability of ACR Values and Mortality in the test dataset. Test dataset was also part of the above datasets.

For another mortality/survival study, the above synthesized datasets were aggregated for USRDS age groups to calculate average food group/subgroup intake by age groups. With that aggregation mortality/survival data were merged. Mortality/survival is used as the target variable. For this, PCA and Regression were used to find association between food groups and CKD/ESRD mortality.

5 Experiment Design

Experiments as provided in Table 2 to 10 are designed to find associations between dietary patterns and CKD mortality as well as to predict ACR values. All of these experiments except set 9 are conducted.

Table 3: **Set 2: Mortality and CKD: Food Sub Groups**

Primary Input Dataset:	NHANES survey aggregated to calculate average food item intake by USRDS like age groups
Target Variable:	ESRD: Avg. Annual Mortality rates
Experiment 2.1:	Identify important food sub groups in the dataset using PCA a) Actual Intake Amount
Experiment 2.2:	Similar to experiment 1.2 (Regression); however, used food subgroups and actual intake only

Table 4: **Set 3: ACR and Food Groups**

Primary Input Dataset:	NHANES survey data for each participant averaged for two surveys
Target Variable:	Albumin Creatinine Ratio (ACR)
Experiment 3.1:	Identify contributing food groups in the input dataset using PCA. This is different than Experiment 1 because entire survey is being used here; not the aggregated data by age groups
Experiment 3.2:	Find out correlation (using Pearson's correlation, and regression) between ACR Values and important food groups as found using PCA in experiment 3.1.

Table 5: **Set 4: ACR Values and Nutrients**

Utilize the same experiments as done for ACR and Food Groups. However, nutrients intake with or without combining with food groups	
Experiment 4.1:	PCA to identify contributing factors
Experiment 4.2:	Regression to find correlations among factors found in experiment 4.1

Table 6: **Set 5: ACR Values and Food Subgroups**

Primary Input Dataset:	NHANES survey data for each participant averaged for two surveys
Target Variable:	Albumin Creatinine Ratio (ACR)
Similar experiments like set 3 and set 4. However, use food subgroups as the input/source variables	

Table 7: **Set 6: Experiments using Regression: ACR Values and Food Subgroups**

And then utilize Machine Learning Approaches for Mortality Prediction on Test Dataset. Input dataset from Set 5 can be used here as the Input dataset	
Experiment 6.1:	ACR value prediction using linear regression. (ACR class can also be an option)
Experiment 6.2:	Use 10 folds cross validations where possible.
Goal:	Check the % of predictability in the test dataset. Precision, recall, or similar might be calculated
Experiment 6.3:	Conduct experiment 6.1; however use Polynomial Regression
Experiment 6.4:	With 10 Folds Cross Validations
Experiment 6.5:	Conduct experiment 6.1; however, use Random Forest Regression with or without 10 Folds Cross Validations. Utilize Polynomial Regression in the process.
Experiment 6.6:	Conduct experiment 6.1; however, use Bayesian prediction with or without 10 Folds Cross Validations. b) Use Polynomial Fit

Table 8: **Set 7: CKD Mortality using Survey data i.e. No aggregation on Age Groups**

Input Data:	Bring CKD mortality data to each participant using the corresponding age
Use PCA (to find contributing food groups and subgroups) and then Regression to find correlation between mortality and Food Groups/Subgroups/ACR values	

Table 9: **Set 8: Experiments using Regression: No aggregated (on age groups) survey data**

Input Data:	Bring CKD mortality data to each participant using the corresponding age
And then utilize Machine Learning Approaches for Mortality Prediction on Test Dataset. Input dataset from Set 7 can be used here as the Input dataset	

Table 10: **Set 9: Remaining Life for CKD Patients and Food Groups/Subgroups: No aggregated survey data**

Input Data:	Bring remaining life data to each participant using the corresponding age and CKD status
Use PCA (to find contributing food groups and subgroups) and then Regression to find correlation between remaining life and Food Groups/ Subgroups/ (ACR values optional)	

6 Results

Associations of Food Groups, Food Subgroups, Food Nutrients with CKD mortality as discovered by the experiments above are provided in this section. Outcome of ACR value prediction in the test data are also provided.

6.1 Food Groups, Food Subgroups, Food Nutrients and CKD Mortality

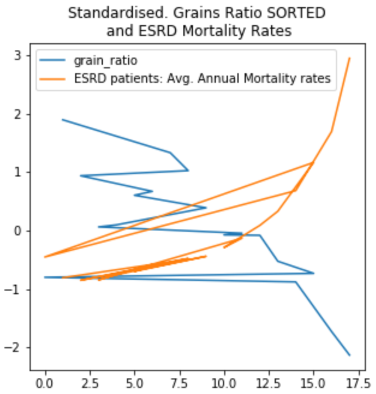
Food Groups and Mortality

Experiments with aggregated (both NHANES and USRDS data) data to find associations between food groups and CKD mortality using Principal Component Analysis (PCA) and Regression show that Grains (-0.84) and Fruits (-0.43) have negative correlations with CKD mortality i.e. mortality is high for the patients who took significantly lower amount of Grains and Fruits than recommended amounts. Data exploration (plots below) also reflects the negative relation. As the correlation for fruits is -0.43 i.e. not very high, hence, Fruits can be thought of mildly/moderately associated.

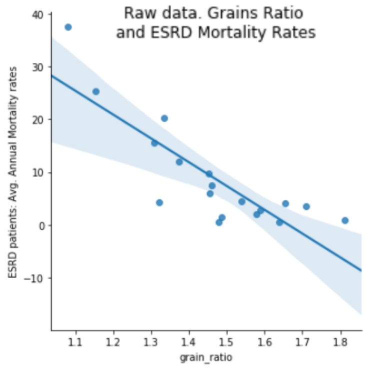
Vegetables show positive (0.58) correlation i.e. mortality is high for the patients who took more vegetables. As the correlation is 0.58, hence, this is not a very strong conclusion. Data shows this correlation in older adults. Even though ratios of food intake amount to high end of recommended amounts (actual intake amounts also show positive relation) were used; age might have biased the correlation. This does not show conformity with the general recommendation to take more vegetables for CKD patients. However, as experiments with food subgroups show that vegetable subgroups such as Other vegetables (0.68), Red and Orange vegetables (0.55), and Starchy vegetables (0.44) show positive correlations that have an impact on the Vegetable group correlation. Though more vegetable intake is a general recommendation for CKD patients; however, certain vegetables with more carbohydrates (and sugars) such as Starchy Vegetables as well as Vegetables with more Potassium and Calcium such as Tomatoes, and Spinach are recommended to take at a lower amount. For diabetes induced CKD, starchy vegetables are not highly recommended in general. Considering the subgroups, the moderate positive correlation as this study found for vegetables food group is in conformity with the general recommendations for CKD patients.

Experiments with aggregated (both NHANES and USRDS data) data to find associations between food subgroups and CKD mortality using PCA and Regression show that Other vegetables (0.68), Red and orange vegetables (0.55), and Starchy vegetables (0.44) have positive correlations with mortality i.e. mortality is low when the intake amounts are low, and mortality is high when intake amounts are high. Data Exploration plots as given below also show the positive relations.

Food subgroups such as Alcoholic beverages (-0.79), Added Sugars/Sugars and sweets (-0.64), Whole grains (-0.61), and Nuts, Seeds, and Soy Products' (-0.55) show the most negative correlations with CKD mortality. Data exploration also shows negative correlations as shown in the charts below. These outcomes can also be seen consistent with current knowledge except for Sugars. Prevalence of Stage 3 CKD is lower in Alcohol Drinkers than non-drinkers [2, 49], Nuts being Phosphorous rich and Whole Grains being Potassium rich are detrimental to CKD patients and can cause higher mortality when taken

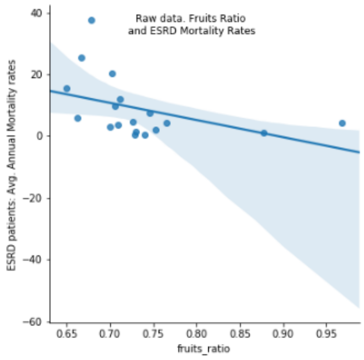


Grains

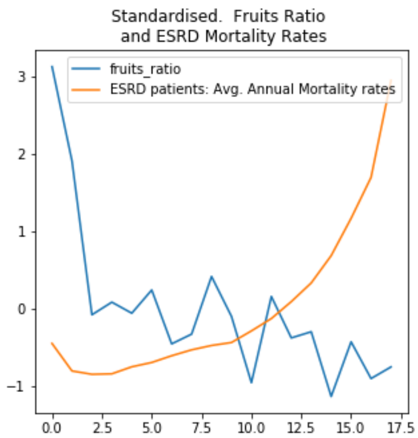


Grains

Figure 3: **Food Groups (Grains) and Mortality - Negative Correlations**

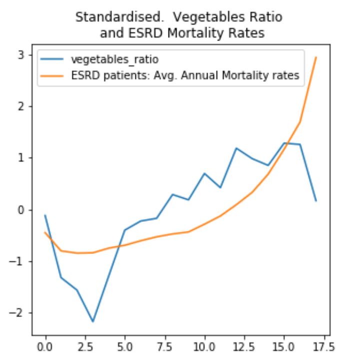


Fruits

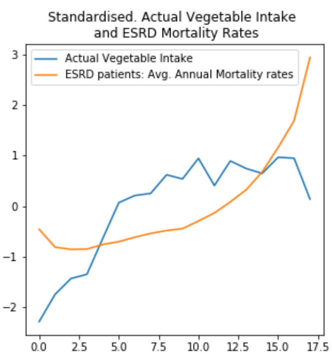


Fruits

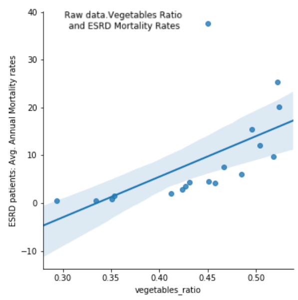
Figure 4: **Food Groups (Fruits) and Mortality - Negative Correlations**



Vegetables Ratio

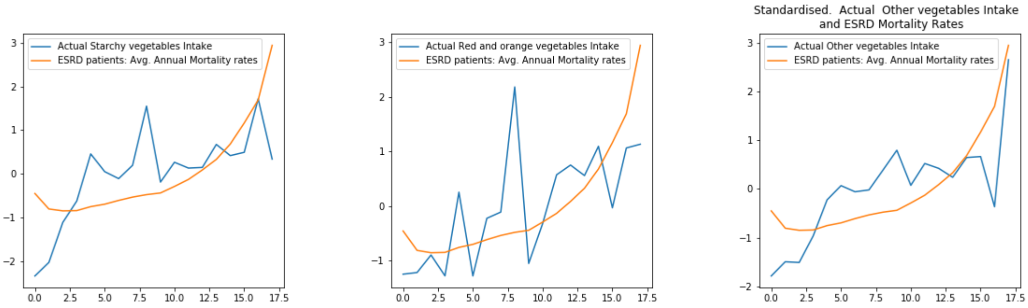


Vegetables Ratio



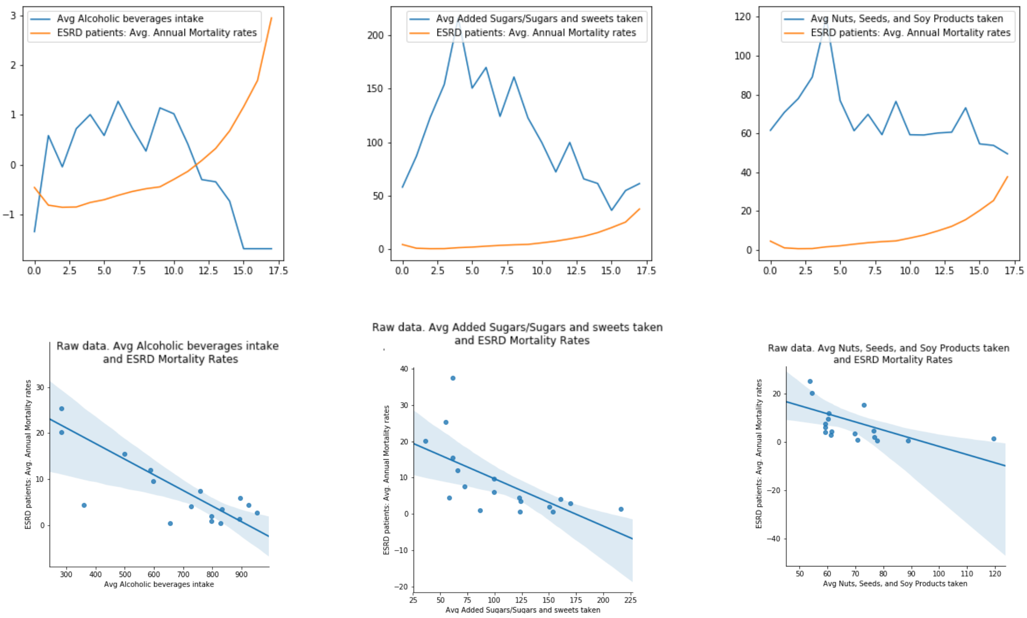
Vegetables Ratio

Figure 5: **Food Groups and Mortality - Positive Correlations (Vegetables)**



Starchy Vegetables Red and Orange Vegetables Other Vegetables

Figure 6: **Food Subgroups and Mortality - Positive Correlations**



Alcohol Sugar Nuts

Figure 7: **Food Subgroups and Mortality - Negative Correlations**

more.

Mortality Study with Non-aggregated Data

Experiments where mortality rates based on ages were brought for each NHANES survey row (i.e. not aggregated), the following food subgroups show more positive correlations than others: Fats, Eggs, Other vegetables, Nonalcoholic beverages, White potatoes and Puerto Rican starchy vegetables, Tomatoes and tomato mixtures, Oils, Deep-yellow vegetables respectively. In the study the following subgroups showed more negative correlations than others: Grain mixtures, frozen plate meals, soups, Crackers and salty snacks from grain products, Milks and milk drinks, Sandwiches with Meat, Poultry, and fish, Poultry. Though the correlation numbers in this study are very low and negligible, however, the positive and negative correlations are consistent with current knowledge [48, 49, 50, 51, 52, 53, 54, 55, 56, 57].

Food Groups, Food Nutrients, and Albumin Creatinine Ratio (ACR) Association

The experiments using PCA and Regression showed negligible effect on ACR with food groups and subgroups. However, some food groups and/or nutrients such as Dairy, and Sugars, Sweets, and Beverages' have higher and positive though negligible (0.02) effect than the others where Fruits (-0.01) showed negative effect. For nutrients, Poly unsaturated fatty acids (-0.02), and Iron (-0.02) have negative correlation where Choline (0.02) showed better positive correlation than others. Findings for Choline matches with medical knowledge [48]. As the correlations are not significant further analysis can be done on the data especially for food groups and nutrients that are found important (using PCA) in the data as provided below:

Dairy, Fats, oils, and salad dressings', Fruits, Grains, Protein, Sugars, sweets, and beverages', Vegetables, Avg energy kcal, avg protein gm, avg carbohydrate gm, avg total fat gm, avg total saturated fatty acids gm, avg total monounsaturated fatty acids gm, avg total polyunsaturated fatty acids gm, avg lutein zeaxanthin mcg, avg thiamin vitamin B1 mg, avg riboflavin Vitamin B2 mg, avg Niacin mg, avg Calcium mg, avg Phosphorus mg, avg Magnesium mg, avg Iron mg, avg Zinc mg, avg Copper mg, avg Sodium mg, avg Potassium mg, avg Selenium mcg, Hexadecenoic gm, Octadecenoic gm

6.2 Food Subgroups and Albumin Creatinine Ratio (ACR) association

The experiments showed Milk desserts, Sauces, Gravies' (0.22), and Alcoholic Beverages (0.087) have more positive correlations with ACR than the other food subgroups i.e. taking more of these food subgroups results higher ACR values. Research by Uehara et al. [55] also shows that excessive Alcohol consumption can cause Proteinuria/Albuminuria (high ACR). Nettleton et al. [56] found that high fat dairy can be linked to high ACR values where low-fat dairy is not strongly linked to high ACR values. However, the correlation as this research found is very low. Low values might still explain a correlation where ACR values might depend on other factors in together than only these food subgroups. Fruits and juicy baby foods show negative correlation (-0.04) though not significant i.e. taking high amount does not increase ACR values that matched with current knowledge [57].

6.3 Predictability of ACR values based on Dietary patterns

Experiments were conducted to predict ACR values from the dietary intake patterns data. Machine Learning (ML) Approaches such as Regression, Polynomial Regression, Random Forest Regression, Bayesian prediction with or without 10 fold cross validations were applied on Food Subgroups intake dataset. Only the food subgroups that were found to be important using PCA were used for the ML approaches.

Target Variables

Absolute ACR values and ACR Categories were used as the target variables. For ACR category, ACR <30 is assigned to class 0 (i.e. no ckd), and ACR >30 is assigned to class 1 (CKD). ACR values less than 30 indicate no CKD, where ACR values between 30 and 300 indicate moderate CKD. ACR >300 is considered severe CKD.

Table 11: **Outcome when ACR Values are used as the Target**

Data Normalized	**Target**	**Approach**	**MSE Train**	**MSE Test**	**RMSE Train**	**RMSE Test**	**R2 Score Train**	**Accuracy: Test R2 Score if not mentioned**
No	ACR Value	10 Fold Cross Validation Polynomial Regression			1	2	3	-0.957 cross val score
No	ACR Value	Polynomial Bayesian with Cross Validation			1	2	3	-0.682 cross val score
No	ACR Value	Polynomial Regression	90965	52946	301	301	0.359	-0.579 r2 score on test data
No	ACR Value	Bayesian on Polynomial fit	93047	47431	305	305	0.344	-0.414 r2 score on test data

Outcome when ACR Values are used as the Target

The best test set accuracies were found using the approaches such as: 10 Fold Cross Validations with Polynomial Regression (95%), Polynomial Bayesian with 10 fold Cross Validations (68%), Polynomial Regression (57%), Bayesian on Polynomial fit (41%), Cross Validation with Polynomial Random Forest Regression (21%). A list of the best performing approaches and the outcome are provided in Table 11 below.

Outcome when ACR Category is Used as the Target

After regression, y >0.5 is assigned to category 1 (CKD), others were assigned to category 0. Test accuracies are 88%. Format for confusion matrix used in the Table 12 below is: (Total, % Correct : [TP, FN, FP, TN]). The high prediction accuracies might relate to the fact that only 10 to 13% population

Table 12: **Outcome when ACR Category is used as the Target**

Data Normalized	**Target**	**Approach**	**Train Confusion Matrix**	**Test Confusion Matrix**
No	Category	Linear Regression	6927, 87%: [6032,0, 895,0]	770, 88% (692, 0, 78, 0)
Yes	Category	Linear Regression	6927, 87%: [6032,0, 895,0]	770, 88% (692, 0, 78, 0)

has ACR >30. However, as cross validations also show high accuracies, it can be concluded that ACR values can be well predicted using Machine Learning approaches especially with 10 Fold Cross Validation Polynomial Regression having accuracy 95%.

7 Conclusions

Chronic Kidney Disease (CKD) leading to End Stage Renal Disease (ESRD) is very prevalent today; treatment facilities for dialysis, and donors for transplantation are limited. Consequently, many die waiting for proper treatment (recent news on USA, Bloomberg). Majority of the studies focused on drugs to control CKD progression where some studies focused on diets, Nutrients, and Food Items. Controlling CKD using changes to dietary patterns can be beneficial to both the patients and the economy. Hence, this study focused on the effect of dietary patterns on CKD mortality as well as on a CKD measure named Albumin Creatinine Ratio (ACR). PCA and Regression are used for the association. Additionally, regression models are trained to predict ACR values from dietary patterns. Grains, Other Vegetables showed positive correlations with Mortality where Alcohol, Sugar, and Nuts showed negative correlations. ACR values were not found strongly correlated with dietary patterns. The outcome for association in general matched with other studies and current knowledge where in couple of cases contradicted with other studies and current knowledge. The association as this study found requires further study and investigation. Machine Learning approaches showed that ACR values could be predicted in the test dataset with high accuracy; the best performing approach was 10 Fold Cross Validation for Polynomial Regression (95%).

8 Appendix

Github Link Root

Github link for Python and SQL Code

9 References

[1] The National Institute of Diabetes and Digestive and Kidney Diseases. What Is Chronic Kidney Disease? https://www.niddk.nih.gov/health-information/kidney-disease/chronic-kidney-disease-ckd/what-is-chronic-kidney-disease

[2] Jaimon T. K., Suetonia C. P., Shu N. W., Marinella R., Juan-Jesus C., Katrina L. C., and Giovanni F. M. S. Healthy Dietary Patterns and Risk of Mortality and ESRD in CKD: A Meta-Analysis of Cohort Studies

[3] Chen X, Wei G, Jalili T, Metos J, Giri A, Cho ME, Boucher R, Greene T, Beddhu S: The associations of plant protein intake with all-cause mortality in CKD. Am J Kidney Dis 67: 423–430, 2016 (26)

[4] Gutierrez O.M., Muntner P, Rizk D.V., McClellan W.M., Warnock D.G., Newby P.K., Judd S.E.:

Dietary patterns and risk of death and progression to ESRD in individuals with CKD: A cohort study. Am J Kidney Dis 64: 204–213, 2014 (27)

[5] Huang X., Jime nez-Moleo J. J., Lindholm B., Cederholm T., Arnoldv J., Rise rus U., Sjogren P., Carrero J. J.: Mediterranean diet, kidney function, and mortality in men with CKD. Clin J Am Soc Nephrol 8: 1548–1555, 2013 (28)

[6] Muntner P, Judd S.E., Gao L, Gutierrez O.M., Rizk D.V., McClellanW., Cushman M., Warnock D.G.: Cardiovascular risk factors in CKD associated with both ESRD and mortality. J Am Soc Nephrol 24: 1159–1165, 2013 (29)

[7] Ricardo A.C., Madero M., Yang W., Anderson C., Menezes M., Fischer M.J., Turyk M., Daviglus M.L., Lash J.P.: Adherence to a healthy lifestyle and all-cause mortality in CKD. Clin J Am Soc Nephrol 8: 602–609, 2013 (30)

[8] Tsuruya K., Fukuma S., Wakita T., Ninomiya T., Nagata M., Yoshida H., Fujimi S., Kiyohara Y., Kitazono T., Uchida K., Shirota T., Akizawa T., Akiba T., Saito A., Fukuhara S.: Dietary patterns and clinical outcomes in hemodialysis patients in Japan: A cohort study. PLoS One 10: e0116677, 2015 (31)

[9] Ricardo A.C., Anderson C.A., Yang W., Zhang X., Fischer M. J., Dember L. M., Fink J. C., Frydrych A., Jensvold N. G., Lustigova E., Nessel L. C., Porter A. C., Rahman M., Wright Nunes J. A., Daviglus M. L., Lash J. P.; CRIC Study Investigators: Healthy lifestyle and risk of kidney disease progression, atherosclerotic events, and death in CKD: Findings from the Chronic Renal Insufficiency Cohort (CRIC) Study. Am J Kidney Dis 65: 412–424, 2015 (17)

[10] Centers for Disease Control and Prevention. National Health and Nutrition Examination Survey. https://wwwn.cdc.gov/nchs/nhanes/search/datapage.aspx?Component=Dietary.

[11] Health.gov. Shifts Needed to Align With Healthy Eating Patterns. https://health.gov/dietaryguidelines/2015/guidelines/chapter-2/a-closer-look-at-current-intakes-and-recommended-shifts/

[12] Agricultural Research Service (ARS), USDA. Food Code Numbers and the Food Coding Scheme. https://reedir.arsnet.usda.gov/codesearchwebapp/$gcp3kq55ssdyc445ry2k2rus$/coding_scheme.pdf

[13] The U.S. Department of Agriculture's (USDA). Vegetable Subgroups. http://www.cn.nysed.gov/common/cn/files/Vegetable%20Subgroups.pdf

[14] Centers for Disease Control and Prevention. Key Concepts About the USDA Food Coding Scheme. https://www.cdc.gov/nchs/tutorials/Dietary/SurveyOrientation/ResourceDietaryAnalysis/Info2.htm

[15] Health.gov. Appendix 3. USDA Food Patterns: Healthy U.S.-Style Eating Pattern. https://health.gov/dietaryguidelines/2015/guidelines/appendix-3/

[16] United states Renal Data System (USRDS). The 2018 USRDS Annual Data Report Reference Tables : https://www.usrds.org/reference.aspx

[17] United states Renal Data System (USRDS). 2018 ADR Chapters. https://www.usrds.org/2018/view/Default.aspx

[18] ARS, USDA. Food Surveys Research Group: Beltsville, MD . Documentation and Dataset: https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/wweia-documentation-and-data-setsY

[19] United states Renal Data System (USRDS). 2018 USRDS Annual Data Report: Executive

Summary. https://www.usrds.org/2018/download/v1_00_ExecSummary_18.pdf
[20] Health.gov, CDC. Dietary Guidelines for Americans 2015-2020. https://health.gov/dietaryguidelines/2015/resources/2015-2020_dietary_guidelines.pdf
[21] Health.gov, CDC. Dietary Guidelines for Americans. https://health.gov/dietaryguidelines/dga95/9DIETGUI.HTM
[22] Mortality and Causes of Death. https://www.usrds.org/2018/ref/ESRD_Ref_H_Mortality_2018.xlsx
[23] Tong A, Chando S, Crowe S, Manns B, Winkelmayer WC, Hemmelgarn B, Craig JC: Research priority setting in kidney disease: A systematic review. Am J Kidney Dis 65: 674–683, 2015
[24] Lin J, Fung TT, Hu FB, Curhan GC: Association of dietary patterns with albuminuria and kidney function decline in older white women: A subgroup analysis from the Nurses' Health Study. Am J Kidney Dis 57: 245–254, 2011
[25] Taylor EN, Fung TT, Curhan GC: DASH-style diet associates with reduced risk for kidney stones. J Am Soc Nephrol 20: 2253–2259, 2009
[26] Liu, Hao-Wen; Tsai, Wen-Hsin; Liu, Jia-Sin; Kuo, Ko-Lin. 2019. "Association of Vegetarian Diet with Chronic Kidney Disease." Nutrients 11, no. 2: 279.
[27] Golaleh Asghari, Mehrnaz Momenan, Emad Yuzbashian, Parvin MirmiranEmail author and Fereidoun Azizi. Dietary pattern and incidence of chronic kidney disease among adults: a population-based study
[28] Tanushree Banerjee1 , Deidra C. Crews2 , Delphine S. Tuot3 , Meda E. Pavkov4 , Nilka Rios Burrows4 , Austin G. Stack5 , Rajiv Saran6,7 , Jennifer Bragg-Gresham6 and Neil R. Powe1,8 ; for the Centers for Disease Control and Prevention Chronic Kidney Disease Surveillance Team9 Poor accordance to a DASH dietary pattern is associated with higher risk of ESRD among adults with moderate chronic kidney disease and hypertension
[29] Jacek R., Beata F., Aleksandra C., Anna G.The Effect of Diet on the Survival of Patients with Chronic Kidney Disease. Nutrients 2017, 9(5), 495; https://doi.org/10.3390/nu9050495
[30] National Kidney Foundation. One in Seven American Adults Estimated to Have Chronic Kidney Disease. https://www.kidney.org/news/one-seven-american-adults-estimated-to-have-chronic-kidney-disease
[31] Hannah N. What are the leading causes of death in the US?. https://www.medicalnewstoday.com/articles/282929.php
[32] Davita. Five Stages of CKD. https://www.davita.com/education/kidney-disease/stages
[33] Goal: How to Identify the Most Important Predictor Variables in Regression Models https ://blog.minitab.com/blog/adventures-in-statistics-2/how-to-identify-the-most-important-predictor-variables-in-regression-models
[34] How to Interpret Regression Analysis Results: P-values and Coefficients https://blog.minitab.com/blog/adventures-in-statistics-2/how-to-interpret-regression-analysis-results-p-values-and-coefficients
[35] Regression Analysis: How to Interpret the Constant (Y Intercept) https://blog.minitab.com/blog

/adventures-in-statistics-2/regression-analysis-how-to-interpret-the-constant-y-intercept
[36] How to Compare Regression Slopes: How to statistically test the difference between regression slopes and constants https://blog.minitab.com/blog/adventures-in-statistics-2/how-to-compare-regression-lines-between-different-models
[37] How Do I Interpret R-squared and Assess the Goodness-of-Fit? https://blog.minitab.com/blog/adventures-in-statistics-2/regression-analysis-how-do-i-interpret-r-squared-and-assess-the-goodness-of-fit
[38] How High Should R-squared Be in Regression Analysis? https://blog.minitab.com/blog/adventures-in-statistics-2/how-high-should-r-squared-be-in-regression-analysis
[39] How to Interpret a Regression Model with Low R-squared and Low P values https://blog.minitab.com/blog/adventures-in-statistics-2/how-to-interpret-a-regression-model-with-low-r-squared-and-low-p-values
[40] Use Adjusted R-Squared and Predicted R-Squared to Include the Correct Number of Variables https://blog.minitab.com/blog/adventures-in-statistics-2/multiple-regession-analysis-use-adjusted-r-squared-and-predicted-r-squared-to-include-the-correct-number-of-variables
[41] How to Interpret S, the Standard Error of the Regression https://blog.minitab.com/blog/adventures-in-statistics-2/regression-analysis-how-to-interpret-s-the-standard-error-of-the- regression
[42] What Is the F-test of Overall Significance in Regression Analysis?. https://blog.minitab.com/blog/adventures-in-statistics-2/what-is-the-f-test-of-overall-significance-in-regression- analysis
[43] Understanding Analysis of Variance (ANOVA) and the F-test. https://blog.minitab.com/blog/adventures-in-statistics-2/understanding-analysis-of-variance-anova-and-the-f-test
[44] How to Compare Regression Slopes. https://blog.minitab.com/blog/adventures-in-statistics-2/how-to-compare-regression-lines-between-different-models
[45] How to Present and Use the Results to Avoid Costly Mistakes, part 1. https://blog.minitab.com/blog/adventures-in-statistics-2/applied-regression-analysis-how-to-present-and- use-the-results-to-avoid-costly-mistakes-part-1
[46] How to Identify the Most Important Predictor Variables in Regression Models. https://blog.minitab.com/blog/adventures-in-statistics-2/how-to-identify-the-most-important-predictor- variables-in-regression-models
[47] How to Interpret your Regression Results http://sitestree.com/how-to-interpret-your-regression-results
[48] Fernandez-Prado R., Esteras R., Perez-Gomez M.V, Gracia-Iguacel C., Gonzalez-Parra E., Sanz A.B., Ortiz A., Sanchez-Nino M.D. Nutrients Turned into Toxins: Microbiota Modulation of Nutrient Properties in Chronic Kidney Disease. Nutrients. 2017 May; 9(5): 489. Published online 2017 May 12. doi: 10.3390/nu9050489 PMCID: PMC5452219 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5452219/
[49] Hsu, Y. H., Pai, H. C., Chang, Y. M., Liu, W. H., Hsu, C. C. (2013). Alcohol consumption is inversely associated with stage 3 chronic kidney disease in middle-aged Taiwanese men. BMC nephrology, 14, 254. doi:10.1186/1471-2369-14-254

[50] Rippe, J. M., Angelopoulos, T. J. (2016). Relationship between Added Sugars Consumption and Chronic Disease Risk Factors: Current Understanding. Nutrients, 8(11), 697. doi:10.3390/nu8110697
[51] Karalius V.P., Shoham D.A. Dietary sugar and artificial sweetener intake and chronic kidney disease: a review. Adv Chronic Kidney Dis. 2013; 20:157–164. doi: 10.1053/j.ackd.2012.12.005
[52] Jaimon T. Kelly, Suetonia C. Palmer, Shu Ning Wai, Marinella Ruospo, Juan-Jesus Carrero, Katrina L. Campbell and Giovanni F. M. Strippoli Healthy Dietary Patterns and Risk of Mortality and ESRD in CKD: A Meta-Analysis of Cohort Studies. CJASN February 2017, 12 (2) 272-279; DOI: https://doi.org/10.2215/CJN.06190616
[53] Jilian k. The 20 Best Foods for People With Kidney Problems. https://www.healthline.com/nutrition/best-foods-for-kidneys
[54] National Kidney Foundation (NKF), USA. Drinking Alcohol Affects Your Kidneys https://www.kidney.org/news/kidneyCare/winter10/AlcoholAffects
[55] Uehara, S., Hayashi, T., Kogawa Sato, K., Kinuhata, S., Shibata, M., Oue, K., Hashimoto, K. (2016). Relationship Between Alcohol Drinking Pattern and Risk of Proteinuria: The Kansai Healthcare Study. Journal of epidemiology, 26(9), 464–470. doi:10.2188/jea.JE20150158
[56] Nettleton, J. A., Steffen, L. M., Palmas, W., Burke, G. L., & Jacobs, D. R., Jr (2008). Associations between micro albuminuria and animal foods, plant foods, and dietary patterns in the Multiethnic Study of Atherosclerosis. The American journal of clinical nutrition, 87(6), 1825–1836. doi:10.1093/ajcn/87.6.1825
[57] Jacobs, D. R., Jr, Gross, M. D., Steffen, L., Steffes, M. W., Yu, X., Svetkey, L. P., Sacks, F. (2009). The effects of dietary patterns on urinary albumin excretion: results of the Dietary Approaches to Stop Hypertension (DASH) Trial. American journal of kidney diseases : the official journal of the National Kidney Foundation, 53(4), 638–646. doi:10.1053/j.ajkd.2008.10. 048. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2676223/
[58] National Kidney Foundation. 37 Million American Adults Now Estimated to Have Chronic Kidney Disease. https://www.kidney.org/news/37-million-american-adults-now-estimated-to-have-chronic-kidney-disease