Fixed-Effect Analysis of Baseline Serum Creatinine across sex and blood pressure levels

Azuka Atum

Background: Context

CKD, also known as chronic kidney disease, or colloquially known as "**kidney disease**" is a condition in which the kidneys are damaged and cannot filter blood[1]. As a results, fluid and waste from blood remain in the body for longer than they should and can lead to other illnesses such as heart disease.

According to the CDC, **every 24 hours 360 people begin dialysis treatment for kidney failure**[2], **or for every 1 hour, 15 people go on dialysis**. Additionally, it is a **leading cause of death in the United States,** costing Medicare beneficiaries **$87.2 billion dollars**[3]. More than 1 in 7 (37 million) US adults are estimated to have CKD[4].

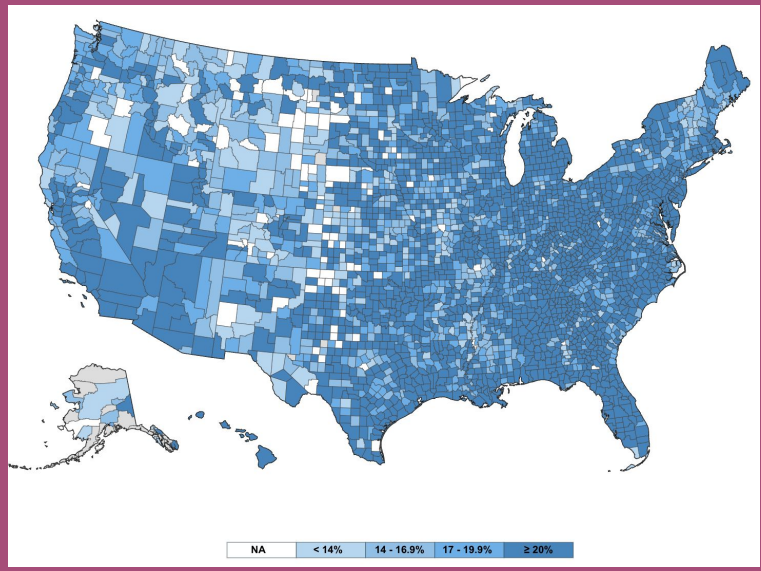

Figure 1. Prevalence of Diagnosed CKD Among Medicare Beneficiaries Aged ≥ 65 Years, by U.S. State and County.

Source:

1. CKD FACTS
2. Kidney Disease Surveillance System
3. Chronic Kidney Disease in the United States, 2023

Background: Context

There are 5 stages of CKD, each one being worse than the last[5]. Each stage is characterized by eGFR (estimated glomerular filtration rate) and an additional marker, such as proteinuria and/or hematuria.[6]

A nephrologist will typically order a series of blood and urine tests in order to assess kidney function. Common measures include but are not limited to: **creatinine, eGFR, and blood urea nitrogen (BUN)**[7].

There are many factors that can impact risk of CKD, such as creatinine levels, sex, and blood pressure level. Blood pressure especially, as kidney disease is closely associated with high blood pressure[8]. Also, creatinine is a waste product that is solely handled and filtered out by the kidneys[9]. A high creatinine concentration in the blood indicates that there is an issue in the kidneys[10].

There is a question of whether there is an observable difference in the effect between sex and baseline diastolic blood pressure in across sex between men and women.

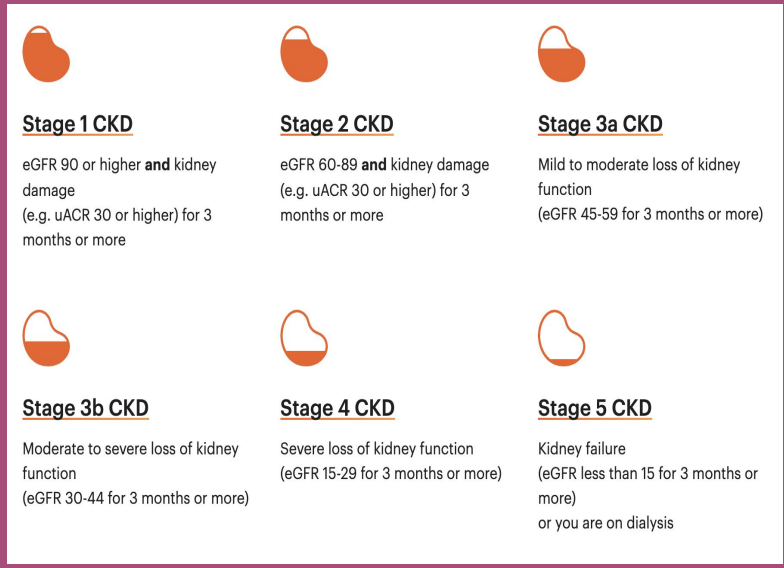

Figure 2. There are 5 stages of CKD.

Background: Dataset

- https://www.kaggle.com/datasets/davidechicco/chronic-kidney-disease-ehrs-abu-dhabi/
- Contains variables:
- Electronic medical records of 491 patients collected at the Tawam Hospital in Al-Ain city (Abu Dhabi, United Arab Emirates), between 1st January and 31st December 2008. The patients included 241 women and 250 men, with an average age of 53.2 years[11].
- .csv file
- 21 variables
- 22 by 498 dimension dataset
- 491 datapoints

feature	explanation	measurement unit	values
ACEIARB	if the patient has taken ACEI or ARB	boolean	[0, 1]
AgeBaseline	age of the patient	integer	[23, 24, …, 80, 89]
BMIBaseline	body–mass index of the patient	kg/m^2	[13, 16, 17, …, 53, 57]
CholesterolBaseline	level of cholesterol	mmol/L	[2.23, 2.40, …, 8.20, 9.30]
CreatinineBaseline	level of creatinine in the blood	mol/l	[6, 27, …, 113, 123]
dBPBaseline	diastolic blood pressure	mmHg	[41, 45, …, 110, 112[
DLDmeds	if the patient has taken dyslipidemia medications	boolean	[0, 1]
DMmeds	if the patient has taken diabetes medications	boolean	[0, 1]
eGFRBaseline	estimated glomerular filtration rate	$ml/min/1.73m^2$	[60, 60.4, …, 242.6]
HistoryCHD	patient history of coronary heart disease	boolean	[0, 1]
HistoryDiabetes	patient history of diabetes	boolean	[0, 1]
HistoryDLD	patient history of dyslipidemia	boolean	[0, 1]
HistoryHTN	patient history of hypertension	boolean	[0, 1]
HistoryObesity	patient history of obesity	boolean	[0, 1]
HistorySmoking	patient history of smoking	boolean	[0, 1]
HistoryVascular	patient history of vascular diseases	boolean	[0, 1]
HTNmeds	if the patient has taken hypertension medications	boolean	[0, 1]
sBPBaseline	systolic blood pressure	mmHg	[92, 95, …, 177, 180]
Sex	if the patient is a woman (0) or a man (1)	binary	[0, 1]
time year	year from follow-up start to severe CKD event or last visit	integer	[0, 1, ..., 9, 10]
[target] EventCKD35	if the patient had moderate-extreme CKD	boolean	[0, 1]

Figure 3. Variables list contained within dataset.

Proposal

Is there an observable difference in baseline creatinine levels and blood pressure between sexes ?

Proposal: Hypothesis Test

H_0: $\mu_0 = \mu_1$ Difference between sexes ->

- There is no difference in mean baseline creatinine levels between sexes.

H_1: $\mu_0 = \mu_1$

- There is a difference in mean baseline creatinine levels between sexes.

H_3: $\mu_0 = \mu_1$

- There is no difference in mean baseline creatinine levels across blood pressure levels.

H_4: $\mu_0 = \mu_1$

- There is a difference in mean baseline creatinine levels across blood pressure levels.

Analysis: Fixed Effect Model

- The model for the analysis is given on the right, where
  - Y = the response, the creatinine baseline
  - α= individual fixed effect
  - $\beta_{1,2}$= coefficients
  - Sex, Blood pressure = variables
  - ϵ= error term
- Blood pressure is assumed to be normally distributed, as well as the interaction effect and the error term, which is iid[13].

$$\gamma_{ik} = \alpha_i + \beta_1(Sex)_{ik} + \beta_2(BloodPressure)_{ik} + \epsilon_{ik}$$

Equation 1. ANOVA model.

Analysis: Baseline Characteristics

	Total (N = 491)	Males (N = 250)	Females (N = 241)
Age (years), Mean (SD)	53.20 ± 13.82	52.68 ± 15.30	53.75 ± 12.11
Male gender (%)	50.9	-	-
History of (%)			
CHD	9.2	13.2	5.0
DM	43.8	46.4	41.1
Vascular disease	5.9	8.8	2.9
HTN	68.2	68.8	67.6
Dyslipidemia	64.6	63.6	65.6
Smoking	15.3	29.2	0.8
Obesity	50.5	40.4	61.0
ACEI/ARB use	46.6	49.6	39.4
Anthropometric values			
BMI (kg/m^2), Mean (SD)	30.19 ± 6.21	28.73 ± 5.77	31.7 ± 6.30
SBP (mmHg), Mean (SD)	131.37 ± 15.69	132.03 ± 15.65	130.70 ± 15.74
DBP (mmHg), Mean (SD)	76.87 ± 10.71	77.64 ± 11.34	76.07 ± 9.98
Laboratory values			
TC (mmol/L), Mean (SD)	4.98 ± 1.10	4.83 ± 1.09	5.13 ± 1.09
TG (mmol/L), Mean (SD)[a]	1.32 ± 0.79	1.35 ± 0.80	1.28 ± 0.78
Cr (μmol/l), Mean (SD)	67.86 ± 17.92	78.97 ± 16.00	56.33 ± 11.30
eGFR (mL/min/1.73m^2), Mean (SD)	98.12 ± 18.50	95.56 ± 20.64	100.77 ± 15.60
HbA1c (%), Mean (SD)[b]	6.60 ± 1.71	6.77 ± 1.98	6.43 ± 1.39

CHD; Coronary heart disease, DM; diabetes mellitus, HTN; hypertension, ACEI; Angiotensin-converting enzyme inhibitors, ARB; Angiotensin II receptor blockers, BMI; Body mass index, eGFR; estimated glomerular filtration rate, SBP; Systolic blood pressure. DBP; Diastolic blood pressure, Cr; Creatinine, TC; Total cholesterol, TG; Triglycerides, SD; standard deviation, HbA1c; glycosylated Hemoglobin, Type A1C.

[a] N = 485.

[b] N = 476.

https://doi.org/10.1371/journal.pone.0199920.t001

Figure 4. Baseline characteristics.

Analysis: Definitions

- **sBPBaseline/dBPBaseline** -> Systolic/Diastolic baseline measures.
  - This allowed me to create a categorization of blood pressure readings according to official literature for ease of analysis[14].
  - Trade-off is you lose degrees of freedom for statistical tests[15].
  - Indicated in dataset as separate results, so I had to combine them. 5 categories:
    - 1 = "Normal"
    - 2 = "Elevated"
    - 3 = "High Blood Pressure (Hypertension STG 1)"
    - 4 = "High Blood Pressure (Hypertension STG 2)"
    - 5 = "Hypertensive Crisis" ;
- **Creatinine:**
  - A normal result for serum creatinine is 0.7 to 1.3 mg/dL (61.9 to 114.9 µmol/L) for men and 0.6 to 1.1 mg/dL (53 to 97.2 µmol/L) for women[16].
  - Left as continuous result in the dataset for analysis.
- **Sex:**
  - 2 categories:
    - 0 = female
    - 1 = male

Healthy and unhealthy blood pressure ranges

Learn what's considered normal, as recommended by the American Heart Association.

BLOOD PRESSURE CATEGORY	SYSTOLIC mm Hg (upper number)	and/or	DIASTOLIC mm Hg (lower number)
NORMAL	LESS THAN 120	and	LESS THAN 80
ELEVATED	120 – 129	and	LESS THAN 80
HIGH BLOOD PRESSURE (HYPERTENSION) STAGE 1	130 – 139	or	80 – 89
HIGH BLOOD PRESSURE (HYPERTENSION) STAGE 2	140 OR HIGHER	or	90 OR HIGHER
HYPERTENSIVE CRISIS (consult your doctor immediately)	HIGHER THAN 180	and/or	HIGHER THAN 120

Figure 5. Blood pressure ranges[16].

Analysis: Parameters

- “High blood pressure” is defined as SBP 130-139 or DBP 80-89[17].
- Gender is a binary predictor, where 0 is “female and 1 is “male”.
- Alpha-level is 0.05.

Analysis: Parameters (Code)

a

```
proc format ;
    value BP
    1 = "Normal"
    2 = "Elevated"
    3 = "High Blood Pressure (Hypertension STG 1)"
    4 = "High Blood Pressure (Hypertension STG 2)"
    5 = "Hypertensive Crisis"  ;

    VALUE CKD
    0 = "No"
    1 = "Yes" ;

    VALUE Sex
    0 = "Female"
    1 = "Male" ;

RUN;
```

b

```
data ckd2 ;
    set ckd ;
    format BP BP.;
    *format Sex Sex.;
    *format CKD CKD. ;
    IF (sBPBaseline < 120) AND (dBPBaseline < 80) THEN BP = 1;
    IF (120 =< sBPBaseline <= 129) AND (dBPBaseline < 80) THEN BP = 2;
    IF (130 <= sBPBaseline <= 139) OR (80 <= dBPBaseline <= 89) THEN BP = 3;
    IF (sBPBaseline => 140) OR (dBPBaseline => 90) THEN BP = 4;
    IF (sBPBaseline > 180) OR (dBPBaseline > 120) THEN BP = 5;
    run;
```

c

```
proc glm data=ckd2;
    title "This is the fixed effect model";
    class Sex BP;
    model CreatinineBaseline=Sex | BP/solution;
    *random temp temp*model / test;
    run;
quit;
```

d

```
proc glm data=ckd2;
    class Sex BP;
    model CreatinineBaseline = Sex BP Sex*BP;
    lsmeans Sex BP / pdiff adjust=tukey;
run;
```

Figure 6 a-d. The code here shows BP recode as well as a fixed effect proc step using proc glm.

Analysis: Results

- H_0 vs H_1 -> Difference between sexes
  - ➢ The p-value is less than the 0.05 alpha level.

    We reject the null hypothesis in favor of the alternative, which says there is a difference in baseline creatinine between the sexes.

- H_3 vs H_4 -> Difference across blood pressure levels
  - ➢ The p-value is 0.55, which is greater than the 0.05 alpha level.

    We fail to reject the null hypothesis in favor of the alternative, meaning that there is no significant difference in baseline creatinine across blood pressure levels.

This is the Fixed Effect model

The GLM Procedure

Dependent Variable: CreatinineBaseline

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	63350.4639	15837.6160	81.90	<.0001
Error	486	93977.3607	193.3691		
Corrected Total	490	157327.8246			

R-Square	Coeff Var	Root MSE	CreatinineBaseline Mean
0.402665	20.49274	13.90572	67.85682

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Sex	1	62945.14826	62945.14826	325.52	<.0001
BP	3	405.31569	135.10523	0.70	0.5532

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Sex	1	61990.69490	61990.69490	320.58	<.0001
BP	3	405.31569	135.10523	0.70	0.5532

Parameter	Estimate		Standard Error	t Value	Pr > |t|
Intercept	77.69318344	B	1.56204750	49.74	<.0001
Sex 0	-22.53859222	B	1.25880083	-17.90	<.0001
Sex 1	0.00000000	B	.	.	.
BP Elevated	0.27245893	B	1.97867960	0.14	0.8905
BP High Blood Pressure (Hypertension STG 1)	2.17626846	B	1.84761561	1.18	0.2394
BP High Blood Pressure (Hypertension STG 2)	1.75809562	B	1.81874979	0.97	0.3342
BP Normal	0.00000000	B	.	.	.

Figure 7. Output of ANOVA analysis using PROC GLM step.

Analysis: Results

Parallel lines suggest no interactions between sex and blood pressure.

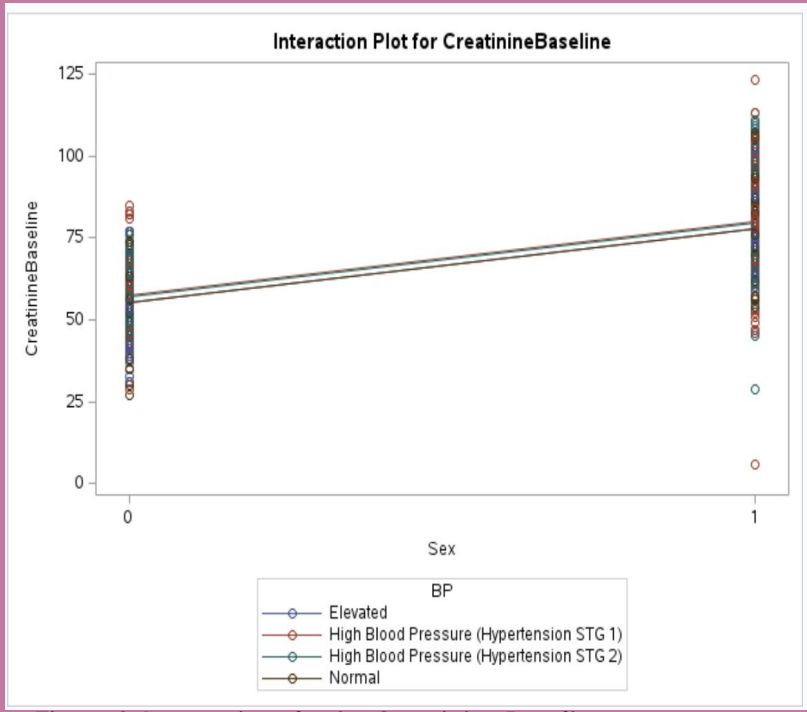

Figure 8. Interaction plot for Creatinine Baseline.

Analysis: Results

- The mean for female creatinine baseline is lower than that for males, and it is statistically significant. This means on average, males tend to have higher creatinine at baseline compared to females in this dataset.
- There is no statistical significance of the means within blood pressure categories.

a

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

Sex	CreatinineBaseline LSMEAN	H0:LSMean1=LSMean2 Pr > |t|
0	56.1168516	<.0001
1	78.9374822	

b

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

BP	CreatinineBaseline LSMEAN	LSMEAN Number
Elevated	66.7748480	1
High Blood Pressure (Hypertension STG 1)	68.7138016	2
High Blood Pressure (Hypertension STG 2)	68.1862857	3
Normal	66.4337323	4

Least Squares Means for effect BP
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: CreatinineBaseline

i/j	1	2	3	4
1		0.7089	0.8579	0.9982
2	0.7089		0.9884	0.6082
3	0.8579	0.9884		0.7710
4	0.9982	0.6082	0.7710	

Figure 9a-b. LSmeans post-hoc analysis.

Conclusion: Discussion

In conclusion, there is a difference in baseline creatinine across sexes, and it is statistically significant. On average, women were over 20 µmol/L lower compared to men. This is consistent with literature that says that creatinine tends to be higher in men than in women.

There was no effect of blood pressure on serum creatinine. This could suggest that there are other factors at play not accounted for in the model.

Conclusion: Limitations

- All the potential features that could impact serum creatinine are not listed in the model.
- The average age of the participants in the data was 53.20 and it was collected in a clinical setting. People who are hospitalized tend to have more serious disease, and older people tend to experience more frequent hospitalizations, so the results are not that generalizable, only for this specific group[19].

Questions

- Q & A

Appendix

Source:

1. CKD FACTS
2. Kidney Disease Surveillance System
3. Chronic Kidney Disease in the United States, 2023
4. https://latex-editor.freebusinessapps.net/tex-editor
5. https://instruction.bus.wisc.edu/jfrees/jfreesbooks/Longitudinal%20and%20Panel%20Data/Book/Chapters/Chapter2.9a.pdf
6. https://murraylax.org/rtutorials/fixedeffects.html#:~:text=The%20term%20%CE%B1i%20is,%2Dtime%2Dinvariant%20omitted%20variables.
7. https://www.statology.org/tukey-test-sas/
8. https://www.cdc.gov/nchs/data/nhsr/nhsr035.pdf
9. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0199920#sec002
10. https://ieeexplore.ieee.org/document/9641833/metrics#metrics
11. https://www.kaggle.com/datasets/davidechicco/chronic-kidney-disease-ehrs-abu-dhabi/
12. https://ieeexplore.ieee.org/mediastore_new/IEEE/content/media/6287639/9312710/9641833/chicc.t1-3133700-large.gif
13. https://stats.oarc.ucla.edu/sas/code/two-way-mixed-anova-using-sas-proc-glm-and-sas-proc-mixed/
14. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0199920
15. Atum, A. (2022) “Cardiovascular disease project”. CSUEB.
16. Atum, A. (2022) “Cardiovascular disease project”. CSUEB.
17. Atum, A. (2022) “Cardiovascular disease project”. CSUEB.
18. Atum, A. (2022) “Cardiovascular disease project”. CSUEB.
19. Atum, A. (2022) “Cardiovascular disease project”. CSUEB.

END