Using CUSUM in Real Time to Signal Clinically Relevant Decreases in Estimated Glomerular Filtration Rate

Reyhaneh Zafarnejad
Pennsylvania State University
Steven Dumbauld
Purdue University West Lafayette
Diane Dumbald
Good Samaritan Hospital
Mohammad Adibuzzaman
Oregon Health & Science University
Paul Griffin (pmg14@psu.edu)
Pennsylvania State University
Edwin Rutsky
University of Alabama at Birmingham

Research Article

Keywords: Chronic Kidney Disease (CKD), Early Detection, CUSUM chart, Electronic Health Record (EHR), End Stage Kidney Disease (ESKD)

Posted Date: April 18th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1527003/v1

Abstract

Background The electronic health record (EHR), utilized to apply statistical methodology, assists provider decision-making, including during the care of chronic kidney disease (CKD) patients. When estimated glomerular filtration (eGFR) decreases, the rate of that change adds meaning to a patient's single eGFR and may represent severity of renal injury. Since the cumulative sum chart technique (CUSUM), often used in quality control and surveillance, continuously checks for change in a series of measurements, we selected this statistical tool to detect clinically relevant eGFR decreases and developed $CUSUM_{GFR}$.

Methods In a retrospective analysis we applied $CUSUM_{GFR}$ to signal identification of eventual ESKD patients prior to diagnosis date. When the patient signaled by reaching a specified threshold $CUSUM_{GFR}$ value, days from $CUSUM_{GFR}$ signal date to ESKD diagnosis date were measured, along with the corresponding eGFR measurement at the signal.

Results Signaling occurred 790 days prior to ESKD diagnosis date with sensitivity of 0.830 and specificity of 0.910. Mean days prior to ESKD diagnosis were significantly greater in Black patients (875), and in patients with hypertension (849), diabetes (940), cardiovascular disease (1037), and hypercholesterolemia (971). Sensitivity and specificity did not vary by sociodemographic and clinical risk factors.

Conclusions $CUSUM_{GFR}$ correctly identified nearly 25% of CKD patients destined for ESKD when eGFR was > 60 ml/min/1.73 m^2. If utilized in an EHR, signaling patients could focus providers' efforts to slow or prevent progression to ESKD.

Background

Given the morbidity, mortality, and financial burden[1] of CKD, identifying eventual ESKD patients, when eGFR is $\geq$ 60 ml/min/1.73 m^2, might provide opportunity to prevent deterioration leading to ESKD. Because of the silent nature of early kidney disease, and lack of recommendation by the US Preventive Services Task Force (USPSTF) for measuring serum creatinine in routine health screening,[2] providers may not identify early CKD patients. The inverse relationship between serum creatinine (S_{cr}) and eGFR results in underappreciation of early small increases in serum creatinine.

Rosansky suggested renal function trajectory might be more important than CKD staging.[3] The trajectory model measured in ml/min/1.73m^2/year assumes a regression line fitted to data points over time. Determining trajectory is difficult as eGFR varies due to volume status, short term medication usage, underlying renal disease activity, age, and gender.[4] Time intervals between eGFR measurements in practice vary widely. Goodness-of-fit with regression analysis depends on observation number. Despite these limitations, Altman and Royston[4] emphasized the role time plays in a series of measurements. For the provider monitoring renal function, "one is specifically looking for the time when something changes."

Using this concept that eGFR change rate is meaningful, CUSUM can be used for monitoring and detecting statistically significant change points in sequential data.[5] Often used for industrial process control, CUSUM provided a useful tool to analyze clinical data in 1977.[6] Subsequent CUSUM reviews demonstrated its use in healthcare applications.[7–9] Related to serial laboratory measurements, Peeks et al[10] identified changes in glucose levels using CUSUM. In nephrology, CUSUM was also used to determine initial dialysis stability[1] and transplant center quality.[12]

By using a notification threshold value T, or signal, for a pre-determined magnitude from a given mean, a CUSUM chart can detect clinically relevant eGFR decreases in a patient's series of measurements. The CUSUM statistic allows the assignment of weights (w) to each calculation, which tunes the signal for optimal sensitivity and specificity for detection of a future clinical risk outcome. In this retrospective data analysis using the statistic $CUSUM_{GFR}$, ESKD diagnosis is the risk outcome, and tuned values of w and T optimize the performance of $CUSUM_{GFR}$. Once the $CUSUM_{GFR}$ value reaches threshold, the patient is likely to progress to ESKD.

Methods

We selected participants from Cerner Health Facts database (Fig. 1), containing EHR data of 1.3 million adult patients with multiple S_{cr} measures from 2010–2019. We calculated eGFRs using the 2009 CKD-EPI Eq. 1[3] for all patients. Patients with acute kidney injury (all eGFR's < 90 ml/min/1.73m^2 within 3 months) were excluded, and the remaining were divided into two mutually exclusive subgroups based on ICD9/10 diagnosis: a group diagnosed with ESKD (ICD9 585.6 or ICD10 N18.6) as the risk outcome, and a group without ESKD. To determine intrinsic, non-pathologic variation in eGFR in the non-ESKD patients, we excluded patients with any CKD Diagnoses (see Appendix Table 1), and those with any eGFR measurement < 60. This Normal Group totaled 85,699 patients and were used to calculate the eGFR mean, $\hat{\mu}$, and standard deviation, $\hat{\sigma}$, for use in the $CUSUM_{GFR}$ statistic and were included in $CUSUM_{GFR}$ calculations. To signal ESKD patients as early as possible, we excluded patients in the ESKD Group with initial eGFR < 60 ml/min/1.73m^2 (5,410 patients). LOINC codes (Appendix Table 2) were used to collect laboratory data including S_{cr} in all patients.

We use the following cumulative statistic:

$$\mathrm{CUSUMGFRi} = min\left[0,\ \left(\frac{eGFRi - \hat{\mu}}{\hat{\sigma}}\right) + w + CUSUMGFRi - 1\right]$$

where $CUSUM_{GFR0} = 0$, $\hat{\mu}$ is the mean of eGFR and $\hat{\sigma}$ is the standard deviation for patients in the Normal Group, and $eGFR_i$ is the ith measurement of eGFR for each patient in both groups. The tuning parameter, w, is determined as noted below. Note $CUSUM_{GFR}$ will always be less than or equal to zero due to the use of minimum operator, which ensures that $CUSUM_{GFR}$ only detects significant decline in eGFR. If the

$CUSUM_{GFR}$ calculation falls below the threshold signal value T, the patient signals likelihood of progressing to ESKD. To determine the optimal value for T, we analyzed the Normal and ESKD Groups using k-fold cross validation (k = 10) for several w and T values. A receiver operator characteristics (ROC) curve (Fig. 2) revealed the best sensitivity, specificity, and accuracy for T, the threshold signal value. When signaled, the patient's eGFR and days prior to ESKD diagnosis were recorded. We determined total population performance measures and when stratified by the sociodemographic variables of age, sex, and race and the clinical factors of hypertension, diabetes, cardiovascular disease, and hypercholesterolemia.

Results

Baseline data on demographics, diagnoses, laboratory results, and medications for the Normal and ESKD Groups are provided in Table 1. The ESKD Group had a significantly higher proportion that were male, Black, Native American, Asian/Pacific Islander, and Hispanic, and higher rates of smoking, hypertension, diabetes, cardiovascular disease, and history of cancer, hypercholesterolemia, and urinary tract abnormalities. All measured laboratory results were significantly different between the Normal Group and the ESKD Group. The ESKD Group had higher rates of non-steroidal anti-inflammatory drug, proton pump inhibitor, and lithium use.

The mean eGFR value for the Normal Group was 80.65 mL/min/1.73 m^2 (SD 7.78). Using Kolmogorov Smirnov goodness of fit test, we could not reject the hypothesis that the mean eGFR for the Normal Group was normally distributed (α=0.05). $CUSUM_{GFR}$ values $CUSUM_{GFR0}$ = 0, w = 0.75, and T = -4.0, gave best mean accuracy (0.83), mean sensitivity (0.91), and mean specificity (0.83) to signal a patient likely to progress to ESKD. Those patients who signaled in the Normal Group were considered false positives, and those in the ESKD Group who failed to signal false negatives.

Figure 3 shows the distribution of eGFR at time of risk signal, and the distribution of signal earliness to actual diagnosis date. Of those in the ESKD Group who signaled as likely to progress, over 83% did so when eGFR was $\geq$ 30, 61.93% when $\geq$ 45, and 24.15% when $\geq$ 60 ml/min/1.73m^2 and signaled 790 days (mean earliness) prior to ESKD diagnosis date (median earliness 361 days). $CUSUM_{GFR}$ signal in two ESKD patients is illustrated in Appendix Fig. 2. Note that the first patient had no eGFR measure less than 60 ml/min/1.73m^2 and yet still signaled correctly.

$CUSUM_{GFR}$ performance is shown in Appendix Table 3 for population subgroups based on sociodemographic factors and clinical risk conditions. Accuracy, sensitivity, and specificity did not vary significantly by subgroup compared to the total values, except in two subgroups. Sensitivity dropped for the non-hypertension subgroup and specificity dropped for the adults over 65 years of age. Black patients signaled earlier than non-Black patients, though their proportion in the Normal Group was relatively low. Mean and median earliness was greater for those with cardiovascular disease, hypercholesterolemia, diabetes, and hypertension.

Discussion

Global prevalence of CKD was 9.1% in 2017 and has increased by over 29% since 1990.[14] CKD progression to ESKD affected over 746,557 individuals in the US in 2017 and is projected at 1.2M by 2030.[15] ESKD is a leading cost in healthcare with Medicare spending for ESKD totaling $35.9B in 2017, 7.2% of Medicare paid claims[2]. Earlier identification of CKD patients likely to progress might reduce the incidence of ESKD.

Despite previous studies using various models[16–21] to predict CKD progression, identification of these at-risk patients is challenging. In early renal injury, S_{cr} increases are subtle, with small increments representing substantial reductions in eGFR, and may be unrecognized. While normal individuals show a fairly constant rate change over a lifetime,[22] CKD patients do not have predictable patterns of progression.[23] In the absence of parametric patterns, linear regression analysis does not perform reliably, and any non-pathologic eGFR change measurement must be differentiated from pathologic causes. No widely accepted method for computing eGFR changes for individual patients is available and $CUSUM_{GFR}$ provides a useful computed statistical application easily incorporated within any healthcare system's EHR.

In our retrospective data analysis using $CUSUM_{GFR}$, it is possible to signal CKD patients likely to progress early in the course of their renal disease. We emphasize that this statistic provides continual monitoring, looking for significant change in eGFR for every serum creatinine measurement for every patient enrolled in a healthcare system's EHR. With the current eGFR indication for nephrology CKD consultation commonly accepted at < 30 ml/min/1.73m^2, opportunity for best intervention at higher eGFR levels may be lost. Since almost a quarter of ESKD Group patients signaled likeliness to progress when eGFR ≥ 60, this indication should be reconsidered. Inclusion of $CUSUM_{GFR}$ within the EHR fits directly into provider workflow since the signal alert is to the provider only when T exceeds the threshold value and would lead the provider to evaluation and treatment algorithms. Early recognition of the CKD patients who signal early might reduce ESKD incidence, and decrease the high morbidity and mortality associated with late nephrology referral. [24,25]

Retrospective analysis of $CUSUM_{GFR}$ in other medical databases is needed to validate these findings, but ultimately the benefit of $CUSUM_{GFR}$ can only be truly estimated through randomized prospective studies. Beyond signaling providers of CKD patients likely to progress to ESKD, other $CUSUM_{GFR}$ applications include timing referral for transplantation and placement of arteriovenous fistulae, correlating $CUSUM_{GFR}$ signaling with renal biopsy activity staging, and has potential use as an endpoint in randomized controlled trials. Not intended as a stand-alone statistic in the care of CKD patients, $CUSUM_{GFR}$ can serve as an important new tool for primary care provider and nephrologist alike.

Abbreviations

CKD – chronic kidney disease

CUSUM – cumulative sum statistic

eGFR – estimated glomerular filtration rate

HER – electronic health record

ESKD – end stage kidney disease

m - meters

ml – milliliters

S_{cr} – serum creatinine

SD – standard deviation

Declarations

Ethics Approval and Consent to Participate

The study was based on deidentified retrospective data from Cerner Health Facts, a HIPAA-compliant database collected from participating clinical facilities. All methods were carried out in accordance with relevant guidelines and regulations. The study was approved by the Purdue University Institutional Review Board (2019–118), including a waiver for informed consent for this study.

Consent for Publication

Not applicable

Availability of Data and Materials

The data that support the findings of this study are not publicly available. They were made available the research team through a data use agreement with Cerner.
https://www.cerner.com/ap/en/solutions/data-research

The source code for the analysis may be found at: https://github.com/Rey-Zafarnejad/Identifying_Clinically_Relevant_Decrease_in_Estimated_Glomerular_Filtration_Rate

Competing Interests

The authors have no competing interests or conflicts of interest to declare.

Funding

This work was supported in part by funding from the Regenstrief Center for Healthcare Engineering, Purdue University.

Author Contributions

Reyhaneh Zafarnejad: Conceptualization; Formal analysis; Writing - original draft. Steve Dumbauld: Conceptualization; Formal analysis; Writing - original draft. Diane Dumbauld: Conceptualization; Formal analysis; Writing - original draft; Mohammad Adibuzzaman: Formal analysis; Writing – review and editing. Paul Griffin: Conceptualization; Formal analysis; Writing - original draft. Edwin Rutsky: Formal analysis; Writing – review and editing.

Acknowledgements

Not applicable

References

1. U.S. Renal Data Systems. 2019 Annual Data Report: Epidemiology in the United States. 2019. Available at: https://www.usrds.org/annual-data-report/ Accessed December 9, 2020.
2. Moyer VA, U.S. Preventive Services Task Force. Screening for chronic kidney disease: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2012;157(8):567.
3. Rosansky SJ. Renal function trajectory is more important than chronic kidney disease stage for managing patients with chronic kidney disease. Am J Nephrol 2012;36:1-10.
4. Altman DG and Royston JP. The hidden effect of time. Statistics in Medicine. 1988;7:629-637.
5. Grigg, O. A., Farewell, V. T., & Spiegelhalter, D. J. (2003). Use of risk adjusted CUSUM and RSPRTcharts for monitoring in medical contexts. Statistical Methods in Medical Research, 12(2), 147–170. doi:10.1177/096228020301200205
6. Wohl H. The CUSUM plot: its utility in the analysis of clinical data. N Engl J Med 1977;296:1044-1045.
7. Chang WR and McLean IP. CUSUM: a tool for early feedback about performance? BMC Medical Research Methodology. 2006;doi:10.1186/1471-2288-6-8.
8. Noyez L. Control charts, Cusum techniques and funnel plots. A review of methods for monitoring performance in healthcare. Interactive CardioVascular and Thoracic Surgery. 2009;9:494-499.
9. Sasikumar R and Devi SB. Cumulative sum charts and its healthcare applications; a systematic review. Sri Lankan Journal of Applied Statistics. 2014;15-1:47-56.
10. Peeks F, Hoogeveen IJ, Feldbrugge RL et al. A retrospective in-depth analysis of continuous glucose monitoring datasets for patients with hepatic glycogen storage disease: recommended outcome parameters for glucose management. J Inherit Metab Dis 2021;44:1136-1150.
11. Rosa AA, Fryd DS, Kjellstrand CM. Dialysis symptoms and stabilization in long-term dialysis: practical application of the CUSUM plot. Arch Intern Med. 1980;140:804-807.

12. Axelrod DA, Guidinger MK, Metzger RA, Wiesner RH, Webb RL, Merion RM. Transplant center quality assessment using a continuously updatable, risk-adjusted technique (CUSUM). Am J Transplant. 2006;6:313-323.
13. Cockwell P, Fisher LA. The global burden of chronic kidney disease. Lancet. 2020;395:662-664.
14. Tonelli M, Dickinson JA. Early detection of CKD: implications for low-income, middle-income, and high-income countries," J Am Soc Nephrol. 2020;31:1931-1940.
15. McCullough KP, Morgenstern H, Saran R, Herman WH, Robinson BM. Projecting ESRD incidence and prevalence in the United States through 2030. J Am Soc Nephrol. 2019;30:127-135.
16. Echouffo-Tcheugui JB, Kengne AP. Risk models to predict chronic kidney disease and its progression: a systematic review. PLoS Med, 2012;9:e1001344.
17. Perotte A, Ranganath R, Hirsch JS, Blei D, Elhadad N. Risk prediction for chronic kidney disease progression using heterogeneous electronic health record data and time series analysis. J Am Med Inform Assoc. 2015;22:872-880.
18. Brosnahan GM, Abebe KZ, Moore CG, et al. Patterns of kidney function decline in autosomal dominant polycystic kidney disease: a post hoc analysis from the HALT-PKD trials. Am J Kidney Dis. 2018;71:666-676.
19. Inker LA, Heerspink HJ, Tighiouart H, et al., GFR slope as a surrogate end point for kidney disease progression in clinical trials: A meta-analysis of treatment effects of randomized controlled trials. J AM Soc Nephrol. 2019;30:1735-1745.
20. Grams ME, Sang Y, Ballew SH, et al. Evaluating glomerular filtration rate slope as a surrogate end point for ESKD in clinical trials: an individual participant meta-analysis of observational data. J Am Soc Nephrol. 2019;30:1746-1755.
21. Greene T, Ying J, Vonesh EF, et al. Performance of GFR slope as a surrogate end point for kidney disease progression in clinical trials: a statistical simulation. J Am Soc Nephrol. 2019;30:1756-1769.
22. Pottel H, Delanaye P, Weekers L, et al. Age-dependent reference intervals for estimated and measured glomerular filtration rate. Clin Kidney J 2017; 10:545.
23. Caravaca-Fontan F, Azeedo L, Luna E, Caravaca F. Patterns of progression of chronic kidney disease at later stages. *Clin Kidney J*. 2020;11:246-253.
24. Navaneethan SD, Aloudat S, Singh S. A systematic review of patient and health system characteristics associated with late referral in chronic kidney disease. *BMC Nephrol.* 2008;9:3.
25. Smart NA, Titus TT. Outcomes of early versus late nephrology referral in chronic kidney disease: a systematic review. *Am J Med.* 2011;124:1073-1080.

Table

Table 1: Baseline demographics, diagnoses, laboratory results, and medications data for normal and ESKD patient groups.

	Normal Group (n = 85,699)	ESKD Group (n = 5,410)
DEMOGRAPHICS		
Mean Age in years**	64.5	57.9
Sex*		
Number Female (percent)	46,456 (54%)	2,354 (44%)
Number Male (percent)	39,182 (46%)	3,056 (56%)
Race/Ethnicity*		
Number Black (percent)	5,826 (7%)	1,147 (21%)
Number Native American (percent)	181 (0%)	110 (2%)
Number Asian/Pacific Islander (percent)	1,062 (1%)	114 (2%)
Number Hispanic (percent)	26 (0%)	55 (1%)
Number Middle Eastern/Indian (percent)	490 (1%)	7 (0%)
Number White (percent)	69,294 (81%)	3,589 (67%)
Number Biracial (percent)	45 (0 %)	7 (0 %)
Number Unknown (percent)	8,754 (10%)	381 (7%)
Number with History of Smoking (percent)*	15,063 (18%)	2,423 (45%)
DIAGNOSES		
Number with Hypertension (percent)*	46,502 (54%)	4,816 (89%)
Number with Diabetes Mellitus (percent)*	22,215 (26%)	3,403 (63%)
Number with Cardiovascular Disease (percent)*		
Coronary Artery Disease	12,812 (15%)	2,346 (43%)
Cerebrovascular Disease (CVA, Stroke)	5,041 (6%)	764 (14%)
Peripheral Vascular Disease	4,338 (5%)	1,168 (22%)

Number with History of Cancer (percent)*	10,294 (12%)	767 (14%)
Number with Hypercholesterolemia (percent)*	48,716 (57%)	3,404 (63%)
Number with History of Urinary Tract Abnormalities (percent)*	4633 (5%)	1512 (28%)
LABORATORY RESULTS		
Urine Microalbumin/Creatinine (mg/g)*		
Number patients < 30 (percent)	12,755 (81.7%)	31 (25.6%)
Number patients between 30 and 300 (percent)	2,593 (16.6%)	40 (33.1%)
Number of patients >= 300 (percent)	255 (1.7%)	50 (41.3%)
Urine Protein/Creatinine (g/g) (se)*	0.11 (0.010)	3.91 (1.247)
Hemoglobin A1c (g/dL) (se)*	5.3 (0.008)	7.2 (0.062)
Hemoglobin (g/dL) (se)*	13.4 (0.002)	10.9 (0.037)
Serum Calcium (mg/dL) (se)*	9.4 (0.001)	8.8 (0.014)
Serum Cholesterol (mg/dL) (se)*	182 (0.051)	159 (1.814)
Serum Albumin (g/dL) (se)*	4.1 (0.001)	3.2 (0.013)
Serum Phosphorus (mg/dL) (se)*	3.4 (0.001)	4.2 (0.034)
Number of patients Hepatitis C positive (percent)*	945 (1%)	237 (4%)
MEDICATION		
Number with any NSAID Use (ibuprofen, naproxen, etc.) (percent)*	16,459 (19%)	2,518 (47%)
Number with any Proton Pump Inhibitor Use (omeprazole, etc.) (percent)*	9,689 (11%)	4,031 (75%)
Number with Bipolar Drug Use (Lithium) (percent)*	233 (0%)	28 (1%)

se = standard error

*Significant difference in means between normal and ESKD Groups based on chi-squared test ($p < 0.05$).

** Significant difference in means between normal and ESKD Groups based on *t*-test ($p < 0.05$).

Appendix

Appendix Figure 2 is not available with this version

Figures

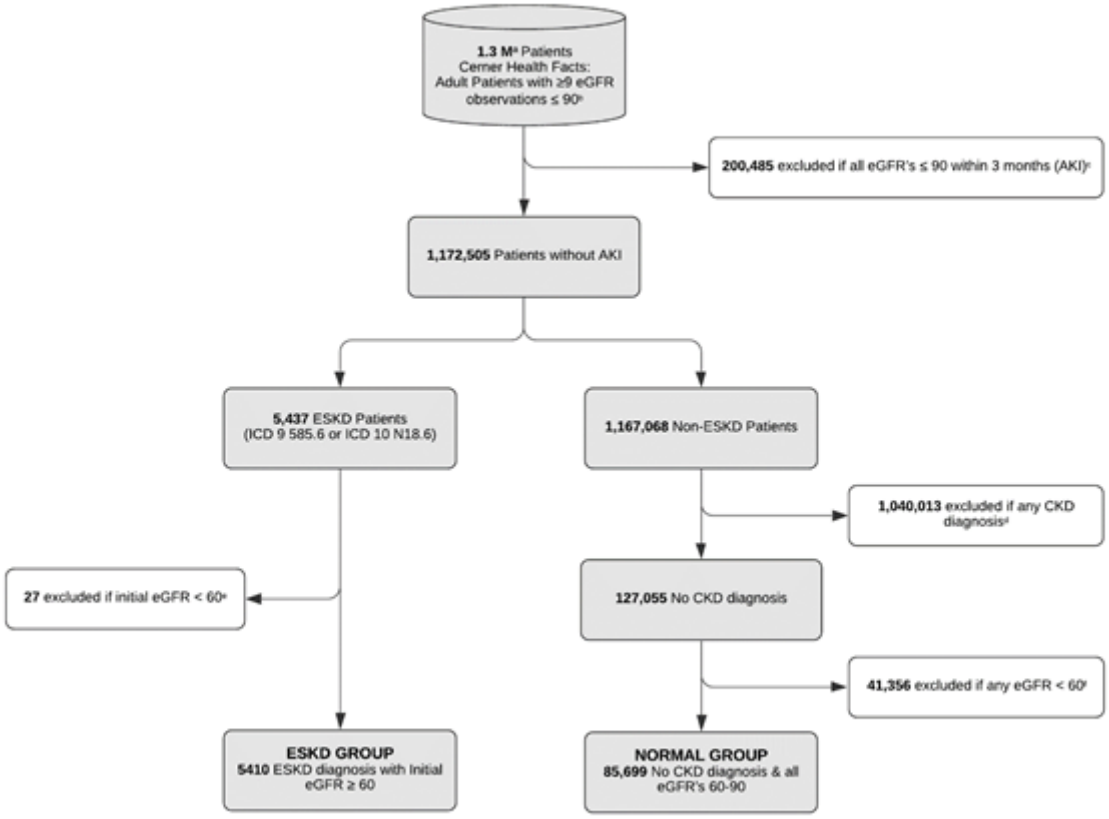

Figure 1

Selection criteria.

[a] Million

[b] All eGFRs in min/ml/1.73m^2

[c] Acute kidney injury

[d] Patients excluded for any ICD9/10 CKD diagnosis (see Appendix Table 1)

[e] Excluded any ESKD Group patient with initial eGFR measurement < 60 min/ml/1.73m^2

[f] Excluded any Normal Group patient with any eGFR < 60 min/ml/1.73m^2

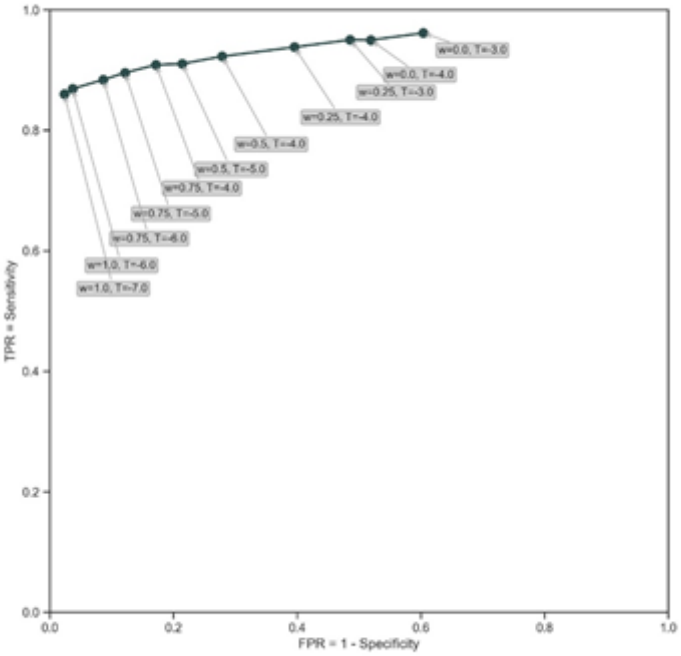

Figure 2

Receiver Operating Characteristic (ROC) curve with sample values for w (tuning parameter) and T (signal value) demonstrating the effect on performance measures (sensitivity and specificity).

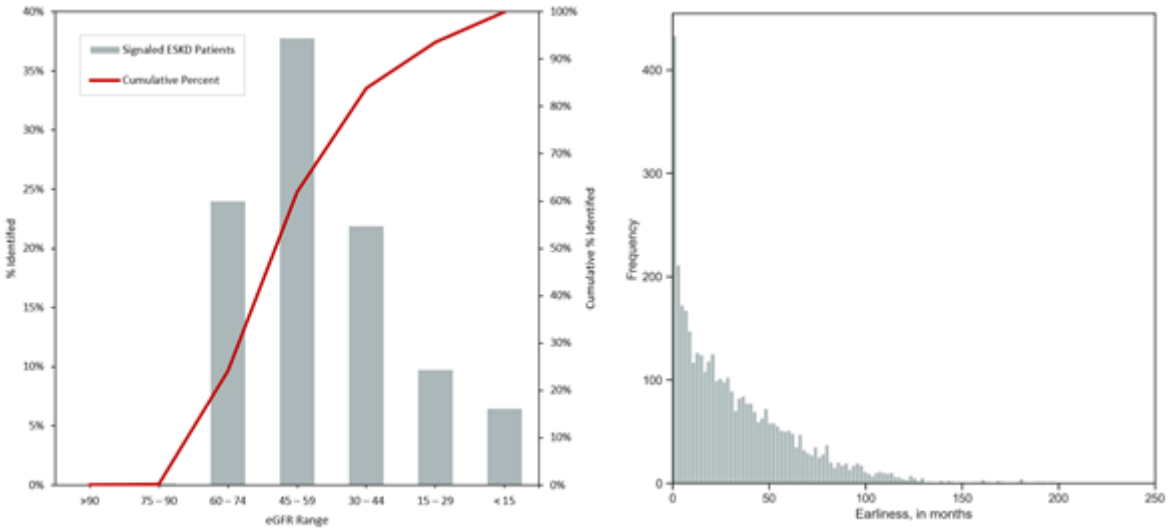

Figure 3

eGFR at $CUSUM_{GFR}$ Signal, in ml/min/1.73m2/year (Left); earliness (in months) from $CUSUM_{GFR}$ Signal ($CUSUM_{GFRi}$ <= -4.0) to ESKD diagnosis. Mean earliness is 26.3 months and median earliness is 12 months. Only those patients correctly identified prior to their diagnosis were included (Right).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- BMCAppendix.docx