

Predictors of Hyperkalemia and Hypokalemia in Individuals with Diabetes: a Classification and Regression Tree Analysis



Emily B. Schroeder, MD, PhD^{1,2}, John L. Adams, PhD³, Michel Chonchol, MD⁴, Gregory A. Nichols, PhD⁵, Patrick J. O'Connor, MD, MA, MPH⁶, J. David Powers, MS¹, Julie A. Schmittiel, PhD, MA⁷, Susan M. Shetterly, MS¹, and John F. Steiner, MD, MPH¹

¹Institute for Health Research, Kaiser Permanente Colorado, 2550 S. Parker Road, Suite 200, Aurora, CO, USA; ²Parkview Health, 11109 Parkview Plaza Drive, Fort Wayne, IN, USA; ³Center for Effectiveness and Safety Research, Kaiser Permanente, Pasadena, CA, USA; ⁴Division of Renal Diseases and Hypertension, University of Colorado, Aurora, CO, USA; ⁵Center for Health Research, Kaiser Permanente Northwest, Portland, OR, USA; ⁶HealthPartners Institute and HealthPartners Center for Chronic Care Innovation, Minneapolis, MN, USA; ⁷Division of Research, Kaiser Permanente Northern California, Oakland, CA, USA.

BACKGROUND: Both hyperkalemia and hypokalemia can lead to cardiac arrhythmias and are associated with increased mortality. Information on the predictors of potassium in individuals with diabetes in routine clinical practice is lacking.

OBJECTIVE: To identify predictors of hyperkalemia and hypokalemia in adults with diabetes.

DESIGN: Retrospective cohort study, with classification and regression tree (CART) analysis.

PARTICIPANTS: 321,856 individuals with diabetes enrolled in four large integrated health care systems from 2012 to 2013.

MAIN MEASURES: We used a single serum potassium result collected in 2012 or 2013. Hyperkalemia was defined as a serum potassium ≥ 5.5 mEq/L and hypokalemia as < 3.5 mEq/L. Predictors included demographic factors, laboratory measurements, comorbidities, medication use, and health care utilization.

KEY RESULTS: There were 2556 hypokalemia events (0.8%) and 1517 hyperkalemia events (0.5%). In univariate analyses, we identified concordant predictors (associated with increased probability of both hyperkalemia and hypokalemia), discordant predictors, and predictors of only hyperkalemia or hypokalemia. In CART models, the hyperkalemia “tree” had 5 nodes and a *c*-statistic of 0.76. The nodes were defined by prior potassium results and eGFRs, and the 5 terminal “leaves” had hyperkalemia probabilities of 0.2 to 7.2%. The hypokalemia tree had 4 nodes and a *c*-statistic of 0.76. The hypokalemia tree included nodes defined by prior potassium results, and the 4 terminal leaves had hypokalemia probabilities of 0.3 to 17.6%. Individuals with a recent potassium between 4.0 and 5.0 mEq/L, eGFR ≥ 45 mL/min/1.73m², and no hypokalemia in the previous year had a $< 1\%$ rate of either hypokalemia or hyperkalemia.

CONCLUSIONS: The yield of routine serum potassium testing may be low in individuals with a recent serum

potassium between 4.0 and 5.0 mEq/L, eGFR ≥ 45 mL/min/1.73m², and no recent history of hypokalemia. We did not examine the effect of recent changes in clinical condition or medications on acute potassium changes.

KEY WORDS: potassium; diabetes mellitus; hypokalemia; hyperkalemia; prediction.

J Gen Intern Med 35(8):2321–8
DOI: 10.1007/s11606-020-05799-x
© Society of General Internal Medicine 2020

INTRODUCTION

Medical laboratory testing is common in the USA, with over 4 billion tests ordered a year.¹ Serum potassium is frequently monitored in individuals on specific medications and individuals with diabetes and other chronic diseases. Both hyperkalemia and hypokalemia can be dangerous, potentially leading to severe cardiac arrhythmias and increased mortality.^{2–6} Additionally, hyperkalemia can lead to muscle weakness, paralysis, cardiac arrhythmias, and sudden cardiac death,² while hypokalemia increases risk of arrhythmias due to alterations in membrane excitability, cardiac automaticity, and ventricular repolarization.³ The association between mortality and abnormalities of potassium homeostasis has been shown to be higher at both extremes of serum potassium in the general population⁷ and in individuals with chronic kidney disease,^{4–6} heart failure, and diabetes.⁵

Diabetes is associated with a higher risk of hyperkalemia,^{8, 9} which can be due to insulin deficiency, acidosis, hyporeninemic hypoaldosteronism, renal tubular acidosis type IV, or comorbid chronic kidney disease.^{8, 10} In addition, blood pressure medications commonly used in diabetes, such as renin-angiotensin-aldosterone system (RAAS) inhibitors, beta blockers, and potassium-sparing diuretics, also increase the risk of hyperkalemia.¹⁰ Chronic kidney disease and heart failure, which are also associated with potassium abnormalities, are common in individuals with diabetes.^{3, 8}

Information on the predictors of hyperkalemia and hypokalemia in individuals with diabetes in routine clinical practice is lacking, however. This information could guide decisions about

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11606-020-05799-x>) contains supplementary material, which is available to authorized users.

Prior Presentations None.

Received September 3, 2019

Revised February 3, 2020

Accepted March 11, 2020

Published online April 16, 2020

the frequency of surveillance for both conditions, especially in light of attempts to minimize unnecessary testing in a patient population at risk for abnormalities in serum potassium.¹ Using data from four large integrated health systems, we used classification and regression tree (CART) models to determine the prevalence and important predictors of hyperkalemia and hypokalemia among individuals with diabetes.

METHODS

Participants

This study was based in the SUPREME-DM (Surveillance, Prevention, and Effectiveness of Management in Diabetes Mellitus) cohort of individuals with diabetes at 4 integrated health systems (HealthPartners [Minnesota], Kaiser Permanente Colorado, Kaiser Permanente Northern California, and Kaiser Permanente Northwest [Oregon/Washington]) from 2005 to 2013. Research institutes embedded in these health systems have developed a distributed virtual data warehouse that contains information on demographics, outpatient pharmacy dispensing, laboratory tests and results, and diagnostic and procedure codes from outpatient and inpatient health care encounters from their electronic health record and administrative data systems.^{11–13}

In this study, diabetes was defined using diagnosis codes, medications, and lab results, as previously described (Appendix Table 2).^{11, 12} Individuals with diabetes were included in these analyses if they were (1) ages 21 or older, (2) enrolled in a participating health care system at any point in 2012–2013 with at least 1 year of prior enrollment, (3) not pregnant, (4) did not have end-stage renal disease, (5) had a serum potassium test in 2012–2013, and (6) had a serum creatinine test within the previous 18 months.

Measures

Outcomes. Hyperkalemia was defined as a serum potassium ≥ 5.5 mEq/L, and hypokalemia was defined as a serum potassium < 3.5 mEq/L. We collected outpatient serum potassium results between 2012 and 2013 and used the most recent one for the classification trees. The date of this serum potassium result was considered the index date.

Exposures. We identified potential predictors of hyperkalemia and hypokalemia using literature review and subject matter knowledge. Potential predictors included demographic factors (age, gender, race/ethnicity, body mass index (BMI) and diabetes type (type 1 vs type 2), the most recent laboratory measurements of eGFR, A1c, and albuminuria in the 18 months prior to but not including the date of the serum potassium lab (the index date)), the number of potassium results in the prior year, comorbidities defined as the presence of one or more diagnosis codes in the prior year (Quan-Elixhauser comorbidity score,¹⁶

congestive heart failure, hypertension, dementia, coronary artery disease, cerebrovascular disease, malignancy, and depression), the number of unique medications prescribed in the prior year, and acute care utilization (emergency department visit in the prior year, inpatient hospitalization in the prior year). We also created variables for current medications of interest (angiotensin-converting enzyme inhibitors and/or angiotensin receptor blockers (ACE/ARB), sulfamethoxazole/trimethoprim, beta blockers, digoxin, potassium-sparing diuretics, other diuretics, potassium chloride supplements, use of medications from 4 or more blood pressure medication classes) using prescriptions where the days' supply plus 10% would have included the date of the index serum potassium result.

We also included a variable for either the presence of prior hyperkalemia or hypokalemia in the 30–365 days (depending on the model), as well as the most recent potassium results within the prior 30–365 days. We did not consider serum potassium results in the 29 days prior to the index date, as these may have represented sequential tests in the same “episode” of hyperkalemia or hypokalemia. Predictors based on diagnostic codes that were considered and excluded due to very low prevalence ($< 1\%$) included the following: renal tubular acidosis, connective tissue disorders, HIV/AIDS, sickle cell disease, calcineurin inhibitors, amyloid disease, and adrenal insufficiency. Use of pentamidine or mineralocorticoids was also excluded due to low frequency. All covariates were assessed as the most recent value prior to but not including the day of the serum potassium result. Diagnostic codes for the different exposure variables are shown in Appendix Table 2.

Analysis

To conduct classification and regression tree (CART) analysis, we built classification trees using binary recursive partitioning for the hyperkalemia and hypokalemia categorical outcomes. Recursive partitioning uses a nonlinear and nonparametric approach to divide events into two subgroups based on continuous or categorical predictors. Unlike linear modeling methods, this approach is able to easily identify complex nested interactions and to identify nonlinear associations.^{18, 19} The goal was to find parsimonious trees that performed well using pruning to try to limit the number of terminal nodes.

When building the CART models, we built the models using 50% of the observations and used the other half to validate the models.^{20–23} Missing data were present for BMI (0.3%), A1c (4.0%), proteinuria (17.1%), race (2.5%), and most recent serum potassium result (20.3%) and were included as a separate category in the models. For the most recent serum potassium result and eGFR, we used conventional and standardized categories to define cutpoints (potassium results: missing, < 3.5 mEq/L, 3.5–3.9 mEq/L, 4.0–4.4 mEq/L, 4.5–4.9 mEq/L, 5.0–5.4 mEq/L, ≥ 5.5 mEq/L; eGFR values 15–29 mL/min/1.73m², 30–44 mL/min/1.73m², 45–59 mL/min/1.73m², 60–89 mL/min/1.73m², ≥ 90 mL/min/1.73m²).

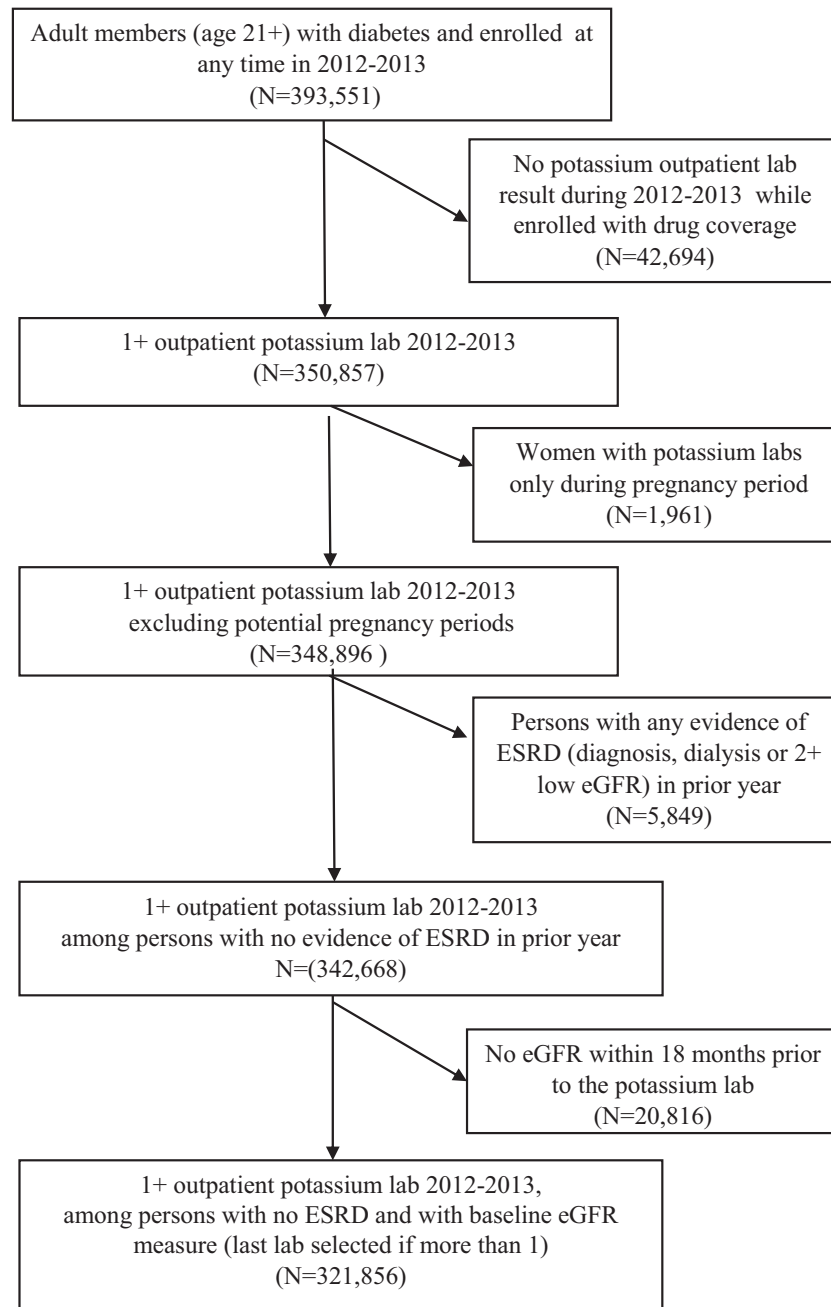


Fig. 1 Cohort construction.

During the tree training phase, if a numeric variable being considered for a candidate split has missing values (i.e., BMI, A1c), these are excluded from the assessment of that variable. If that variable is chosen for the split, observations with missing values on that variable are removed from the entire tree. During prediction, if there are missing values in a numeric variable used in the chosen tree, those records are dropped. Because most recent potassium value was specified using categories, missing values for most recent potassium value remained in the tree.

We used the *c*-statistic to assess performance on trees ranging from 2 to 20 terminal nodes. Tenfold cross-validation was used on each tree size to determine best performance, and the

c-statistic was calculated from predicting the best-sized tree developed in the training sample on the holdout sample. In this setting, the *c*-statistic can be conceptualized as the area under the ROC curve, or the proportion of all possible outcome discordant pairs where the predicted probability ranks the pair correctly. For the final model, we estimated a tree of the chosen size on the combined training and validation sample. The binary recursive partitioning was all conducted in R version 3.4.3 using the tree package.^{24, 25}

Using clinical judgment, we consider a risk of less than 1% to be low risk of hyperkalemia or hypokalemia, 1–5% to be intermediate risk, and 5% or greater to be high risk.^{26, 27}

Table 1 Covariate Associations with Hyperkalemia and Hypokalemia (N = 321,856)

N (%) unless otherwise noted	Hyperkalemia				p value*	Hypokalemia				
	No N = 320,340; 99.5%		Yes N = 1516; 0.5%			No N = 319,300; 99.2%		Yes N = 2556; 0.8%		p value*
	N/ mean	%/SD	N/ mean	%/SD		N/ mean	%/SD	N/ mean	%/SD	
Age, mean (SD)	64.9	13.1	69.1	12.8	<0.001	64.9	13.1	65.8	13.4	<0.001
Male gender	166,319	51.9	913	60.2	<0.001	166,216	52.1	1016	39.7	<0.001
Race										
Hispanic	54,562	17.0	230	15.2	<0.001	54,474	17.1	318	12.4	<0.001
Non-Hispanic White	173,797	54.3	920	60.7		173,340	54.3	1377	53.9	
Non-Hispanic Black	28,402	8.9	110	7.3		28,061	8.8	451	17.6	
Other	55,622	17.4	231	15.2		55,479	17.4	374	14.6	
Unknown	7957	2.5	25	1.7		7946	2.5	36	1.4	
BMI**	31.5	7.3	30.1	7.5	<0.001	31.5	7.3	31.2	7.9	0.048
GFR, mean (SD)	76.6	23.4	53.4	25.7	<0.001	76.6	23.4	72.0	26.4	<0.001
A1c**, mean (SD)	7.4	1.5	7.5	1.6	0.034	7.4	1.5	7.2	1.6	<0.001
Elixhauser comorbidity score, mean (SD)	4.0	2.7	6.1	3.5	<0.001	4.0	2.7	5.7	3.6	<0.001
Congestive HF	28,413	8.9	334	22.0	<0.001	28,239	8.8	508	19.9	<0.001
Malignancy	17,781	5.6	142	9.4	<0.001	17,616	5.5	307	12.0	<0.001
Hypertension	218,206	68.1	1245	82.1	<0.001	217,364	68.1	2087	81.7	<0.001
Dementia	9674	3.0	111	7.3	<0.001	9624	3.0	161	6.3	<0.001
CAD/CVD	52,680	16.4	457	30.1	<0.001	52,601	16.5	536	21.0	<0.001
Type 1 diabetes	6246	1.9	50	3.3	<0.001	6253	2.0	43	1.7	0.32
Depression	48,729	15.2	290	19.1	<0.001	48,397	15.2	622	24.3	<0.001
Proteinuria					<0.001					<0.001
Normal	198,393	61.9	620	40.9		197,690	61.9	1323	51.8	
Macroalbuminuria	19,255	6.0	343	22.6		19,351	6.1	247	9.7	
Microalbuminuria	47,871	14.9	306	20.2		47,691	14.9	486	19.0	
Missing	54,821	17.1	247	16.3		54,568	17.1	500	19.6	
ACE/ARB	198,777	62.1	905	59.7	0.06	198,462	62.2	1220	47.7	<0.001
Beta blocker	109,957	34.3	751	49.5	<0.001	109,634	34.3	1074	42.0	<0.001
Digoxin	5444	1.7	40	2.6	0.005	5434	1.7	50	2.0	0.32
Diuretic, K-sparing	5586	1.7	14	0.9	0.01	5449	1.7	151	5.9	<0.001
Diuretic, other	113,165	35.3	593	39.1	0.002	112,105	35.1	1653	64.7	<0.001
Potassium CL supplement	16,242	5.1	97	6.4	0.02	15,739	4.9	600	23.5	<0.001
Sulfamethoxazole/trimethoprim	3121	1.0	40	2.6	<0.001	3119	1.0	42	1.6	0.001
4+ classes BP medications#	28,858	9.0	208	13.7	<0.001	28,589	9.0	477	18.7	<0.001
Number of medications in prior year, mean (SD)	9.2	5.6	11.7	6.9	<0.001	9.2	5.6	12.3	7.2	<0.001
Any ED visit in prior year	88,530	27.6	709	46.8	<0.001	88,139	27.6	1100	43.0	<0.001
Any inpatient stay prior year	41,732	13.0	471	31.1	<0.001	41,478	13.0	725	28.4	<0.001
Count of K tests prior year, mean (SD)	1.7	2.4	3.5	4.9	<0.001	1.7	2.4	3.5	5.3	<0.001
Count of K tests 2012–2013, mean (SD)	3.8	3.7	6.6	7.8	<0.001	3.8	3.7	6.2	8.4	<0.001
1+ Hyper-K prior year***	6141	1.9	328	21.6	<0.001	6423	2.0	46	1.8	0.45
1+ Hypo-K prior year***	6903	2.2	42	2.8	0.10	6157	1.9	788	30.8	<0.001
Any hyper-K or hypo-K in prior 30–365 days	12,817	4.0	360	23.8	<0.001	12,366	3.9	811	31.7	<0.001
Prior K value, mean (SD)***	4.4	0.4	4.9	0.5	<0.001	4.4	0.4	3.8	0.4	<0.001
Missing prior K value	65,260	20.4	197	13.0	<0.001	65,078	20.4	379	14.8	<0.001

*p values from chi-square tests for categorical variables and t tests for continuous variables

#Blood pressure medication classes were as follows: ACE/ARBs, aldosterone antagonists, alpha blockers, alpha-1 adrenergic blockers, alpha-2 adrenergic agonists, beta blockers, CCB-DHP, CCB-nonDHP, thiazide diuretics, direct renin inhibitors, potassium-sparing diuretics, loop diuretics, CAI diuretics, peripheral vasodilators, and reserpine

**Missing data: for BMI n = 1104 (0.3%), A1c n = 12,848 (4.0%)

***Prior hyper-K, hypo-K, and prior K value all examined in the previous 30–365 days. No prior K value was available for n = 65,457 (20.3%)

This study was approved by the KPCO Institutional Review Board (IRB), with the other institutions' IRBs ceding oversight to the KPCO IRB. Informed consent was not required.

RESULTS

We identified 393,551 individuals with diabetes enrolled at any time during 2012–2013 (Fig. 1). Of those, 42,694 (10.8%)

did not have any qualifying serum potassium results. After exclusions for pregnancy, ESRD, and lack of baseline eGFR, the final cohort included 321,856 individuals. The mean number of serum potassium results over the 2-year period was 3.8 (STD 3.8).

For the 2-year period (2012–2013), the prevalence of ever having hyperkalemia was 3.6% (11,450/321,856) and of every having hypokalemia was 3.4% (11,075/321,856). Using only the most recent serum potassium results between 2012 and

2013, we included 2556 hypokalemia events (0.8%) and 1517 hyperkalemia events (0.5%) in subsequent analyses.

Table 1 describes the characteristics of the population and their univariate associations with hyperkalemia and hypokalemia. Many factors were concordant and associated with increased probability of both hyperkalemia and hypokalemia. These include older age, decreased eGFR, proteinuria, total number of comorbidities (Elixhauser comorbidity scale), congestive heart failure, malignancy, hypertension, dementia, coronary artery disease and cerebrovascular disease, and depression, emergency department or inpatient utilization, beta blocker use, other diuretic use, sulfamethoxazole/trimethoprim use, potassium supplement use, use of 4 or more classes of blood pressure medications, number of medications in the prior year, and number of serum potassium results in the previous year. There were also several discordant characteristics. Men were more likely to have hyperkalemia, while women were more likely to have hypokalemia. Hyperkalemia was more common among white individuals, while hypokalemia was more common among African Americans. Digoxin use was associated with an increased probability of hyperkalemia. ACE/ARB use was associated with a decreased probability of hypokalemia. Interesting, ACE/ARB use was not associated with an increased probability of hyperkalemia. Use of potassium-sparing diuretics was associated with a decreased probability of hyperkalemia and increased probability of hypokalemia. Type 1 diabetes was associated with an increased probability of hyperkalemia.

Appendix Table 3 shows similar associations for ever having hyperkalemia or hypokalemia from 2012 to 2013, with the following exceptions: digoxin use was also associated with an increased probability of hypokalemia, ACE/ARB use was associated with an increased probability of hyperkalemia, and type 1 diabetes was associated with a decreased probability of hypokalemia.

Figure 2 a and Appendix Table 4 illustrate that the probability of both hyperkalemia and hypokalemia increased with decreasing eGFR. The association was stronger for hyperkalemia, with a prevalence of hyperkalemia 16.5 times higher in individuals with an eGFR of 15–29 mL/min/1.73m² than in individuals with an eGFR \geq 90 mL/min/1.73m². For hypokalemia, the prevalence of hypokalemia was 2.9 times higher in individuals with an eGFR of 15–29 mL/min/1.73m² than in individuals with an eGFR of \geq 90 mL/min/1.73m².

Figure 2 b and Appendix Table 4 illustrate the strong relationship between the most recent serum potassium result and the risk of hyperkalemia and hypokalemia. In individuals with a prior serum potassium < 5.0 mEq/L, the risk of hyperkalemia was < 1.0%, while the risk in individuals with a prior serum potassium of 5.0–5.4 was 2.6%, and in individuals with a prior serum potassium \geq 5.5 was 8.4%. In individuals with a prior serum potassium > 4.0 mEq/L, the risk of hypokalemia was < 1.0%, while the risk in individuals with a prior serum potassium of 3.5–4.0 mEq/L was 3.2%, and in individuals with a prior serum potassium < 3.5 mEq/L was 17.0%.

The final hyperkalemia tree had 5 nodes (Fig. 3a), and the final hypokalemia tree had 4 nodes (Fig. 3b). The

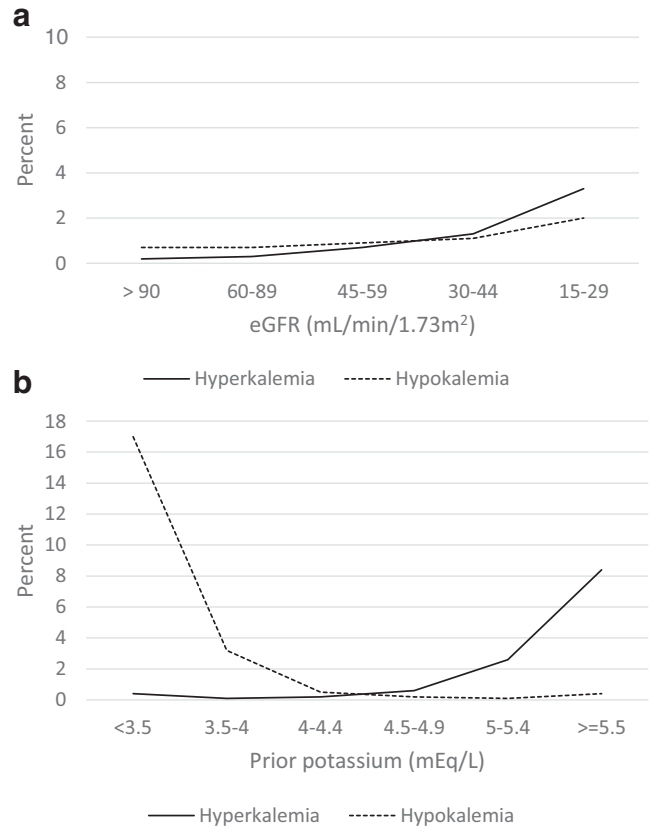


Fig. 2 a Association between eGFR and hyperkalemia and hypokalemia. b Association between most recent prior potassium value and hyperkalemia and hypokalemia.

hyperkalemia tree had a *c*-statistic of 0.76 in the training set, and the hold out *c*-statistic was 0.76. The nodes were limited to prior serum potassium results and eGFRs, and the 5 terminal leaves had probabilities of hyperkalemia ranging from 0.2 to 7.2%. The vast majority (84.4%) of the sample was in the lowest probability terminal leaf. Using cutpoints as defined above for low (< 1%), intermediate (1–5%), and high (\geq 5%) risk, then the population can be divided into risk groups using a few simple rules: (1) most recent potassium value > 5.0 and any hyperkalemia in the past year = high risk; (2) most recent potassium value > 5.0, no hyperkalemia in the past year, and eGFR < 30 mL/min/1.73m² = high risk; (3) most recent potassium result < 5.0, and eGFR \geq 45 = low risk; (4) all others = intermediate risk.

The hypokalemia tree had a *c*-statistic of 0.76 in the training set, and the hold out *c*-statistic was 0.77. The hypokalemia tree included nodes based on prior serum potassium results, and the 4 terminal leaves had probabilities of hypokalemia ranging from 0.3 to 17.6%. The vast majority (88.5%) of the sample was in the lowest probability terminal leaf. Using the same risk cut points as for hyperkalemia, the population can be divided into risk groups with the following rules: (1) any hypokalemia in the past year = high risk; (2) no hypokalemia in the past year and most recent K is between 3.5 and 4 = intermediate risk; (3) all others = low risk.

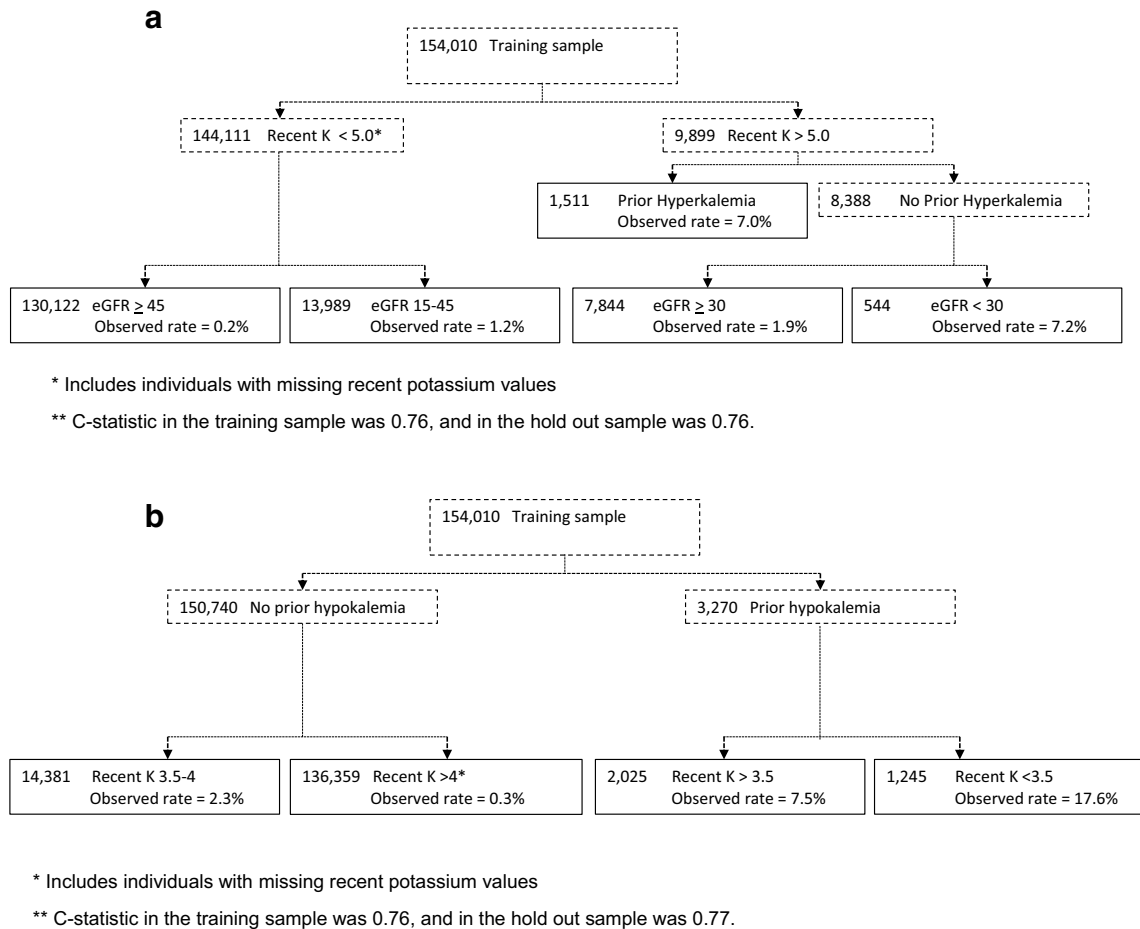


Fig. 3 a Hyperkalemia tree with 5 nodes, training sample. *Includes individuals with missing recent potassium values. *c-statistic in the training sample was 0.76, and in the hold out sample was 0.76. b Hypokalemia tree with 4 nodes, training sample. *Includes individuals with missing recent potassium values *c-statistic in the training sample was 0.76, and in the hold out sample was 0.77.

DISCUSSION

In this population of individuals with diabetes, the prevalence rates of hyperkalemia and hypokalemia were low (0.5% and 0.8%, respectively). Demographic factors, comorbid conditions, and medications were associated with an increased probability of both hyperkalemia and hypokalemia. There were also some discordant factors, including gender, race, certain medications, and prior serum potassium results. Low eGFR and prior serum potassium results, but not other demographic characteristics, clinical variables, or medications, defined the branch points in both the hyperkalemia trees, while the hypokalemia nodes were all based on prior serum potassium results. These risk factors are well established but emphasize the low overall rates of hyperkalemia and hypokalemia and the prognostic importance of prior serum potassium results.

Individuals with diabetes often undergo both routine serum potassium testing (often as part of a battery of annual or semi-annual labs) and serum potassium testing that are prompted by clinical factors (illness, medication changes, recent hospitalization). As a result, the true point prevalence of hyperkalemia and hypokalemia is likely lower than that observed in our study.

Strengths of our analysis include our evaluation of the frequency and likelihood of hyperkalemia and hypokalemia in real world setting, where the frequency likely differs from that observed in clinical trials. In addition, we had access to nearly complete outpatient laboratory values and a wide range of covariates. Our focus on prediction of hypokalemia and hyperkalemia using the information available to the clinician at the time the decision is made to test for serum potassium replicates the clinical decision-making process. Meanwhile, our use of CART regression divides the population into risk categories in a clear and transparent manner.

In the Stockholm Creatinine Measurements (SCREAM) project of 364,955 individuals receiving health care in Stockholm, Sweden in 2009 and followed for 3 years, 13.6% had at least one episode of hypokalemia and 2.5% had hyperkalemia. The median number of measurements was 2 per year.⁹ In the Geisinger Health System, over a 3-year time period, the proportion of patients with hyperkalemia was 2.3%.² Over a 2-year time period, we found a lower prevalence of hypokalemia (3.4%) and slightly higher prevalence of hyperkalemia (3.6%).

Most of the risk factors that we identified for hyperkalemia are similar to those reported by others,^{7,9} with the exception that we

did not find a higher risk of hypokalemia in younger individuals. In addition, the direction of associations for some blood pressure medications differed. For example, we did not find an association between ACE/ARBs and increased risk of hyperkalemia and did find associations between potassium-sparing diuretics and hypokalemia and between potassium supplements and hypokalemia. These likely reflect decisions that clinicians have made in the setting of prior hypokalemia, or due to co-prescription with other medications that may cause hypokalemia, and thus could reflect different patterns of confounding by indication.

Limitations of our study include the lack of standardized laboratory testing processes across the study sites and within study sites. While we attempted to identify erroneous values, spurious values (especially hyperkalemia due to hemolyzed samples) may have occurred. While we did not include potassium values in the 29 days preceding the index value, there is the possibility of recent clinical actions after the most recent potassium value affecting the index potassium value. Our study reflects patterns of care in 4 large integrated health systems and so may not be generalizable to other health care settings. Confounding by indication complicates the ability to interpret some of the medication-potassium associations. In addition, potassium values exist along a continuum, while we used cutpoints of < 3.5 mEq/L and ≥ 5.5 mEq/L to define hypokalemia and hyperkalemia, respectively. In doing so, we followed common clinical laboratory reference ranges, sought to identify values that would likely prompt clinical action, and followed examples in the literature.^{2, 9} However, different cutpoints could have been chosen (for example, < 4 mEq/L or ≥ 5.0 mEq/L) which might have resulted in different CART models and/or *c*-statistics.

No consensus exists about how often to monitor serum potassium. Our findings suggest that routine serum potassium testing may be low yield in individuals with a recent serum potassium between 4.0 and 5.0 mEq/L, eGFR ≥ 45 mL/min/1.73m², and no hypokalemia in the previous year ($< 1\%$ rate of hypokalemia and hyperkalemia). We did not examine the effect of recent changes in clinical condition or medications on acute potassium changes.

Corresponding Author: Emily B. Schroeder, MD, PhD; Institute for Health Research, Kaiser Permanente Colorado, 2550 S. Parker Road, Suite 200, Aurora, CO 80014, USA (e-mail: emily.schroeder@parkview.com).

Author Contributors The authors thank Ms. Andrea Paolino for her invaluable help with this manuscript and the SUPREME-DM study.

Funding Information This project was supported by grant 1K23DK099237 from the National Institute of Diabetes and Digestive and Kidney Disease (NIDDK) and grants 1R01HS022963 and R01HS019859 from the Agency for Healthcare Research and Quality. J.A.S. and P.J.O. received additional support from NIDDK-funded Health Delivery Systems Center for Diabetes Translational Research grant 1P30DK92924.

Compliance with Ethical Standards:

This study was approved by the KPCO Institutional Review Board (IRB), with the other institutions' IRBs ceding oversight to the KPCO IRB. Informed consent was not required.

Conflict of Interest: M.C. reports additional grant funding from the Department of Defense, Sanofi, Kadmon, AstraZeneca, and Corvidia. G.A.N. reports additional grant funding from Boehringer Ingelheim, Merck & Co, Bristol-Myers Squibb, and Amarin Pharma.

Disclaimer: The study sponsors had no role in the study design, collection, analysis, and interpretation of data; writing the report; and the decision to submit the report for publication. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Agency for Healthcare Research and Quality.

REFERENCES

1. Morgan DJ, Dhruva SS, Coon ER, Wright SM, Korenstein D. 2018 Update on medical overuse. *JAMA Intern Med.* 2019;179:240–46.
2. Chang AR, Sang Y, Leddy J, et al. Antihypertensive medications and the prevalence of hyperkalemia in a large health system. *Hypertension.* 2016;67:1181–8.
3. Bowling CB, Pitt B, Ahmed MI, et al. Hypokalemia and outcomes in patients with chronic heart failure and chronic kidney disease: findings from propensity-matched studies. *Circ Heart Fail.* 2010;3:253–60.
4. Hayes J, Kalantar-Zadeh K, Lu JL, Turban S, Anderson JE, Kovesdy CP. Association of hypo- and hyperkalemia with disease progression and mortality in males with chronic kidney disease: the role of race. *Nephron Clin Pract.* 2012;120:c8–16.
5. Collins AJ, Pitt B, Reaven N, et al. Association of serum potassium with all-cause mortality in patients with and without heart failure, chronic kidney disease, and/or diabetes. *Am J Nephrol.* 2017;46:213–21.
6. Nakhoul GN, Huang H, Arrigain S, et al. Serum potassium, end-stage renal disease and mortality in chronic kidney disease. *Am J Nephrol.* 2015;41:456–63.
7. Kovesdy CP, Matsushita K, Sang Y, et al. Serum potassium and adverse outcomes across the range of kidney function: a CKD Prognosis Consortium meta-analysis. *Eur Heart J.* 2018;39:1535–42.
8. Dunn JD, Benton WW, Orozco-Torrentera E, Adamson RT. The burden of hyperkalemia in patients with cardiovascular and renal disease. *Am J Manag Care.* 2015;21(15 Suppl):s307–15.
9. Nilsson E, Gasparini A, Arnlov J, et al. Incidence and determinants of hyperkalemia and hypokalemia in a large healthcare system. *Int J Cardiol.* 2017;245:277–84.
10. Liamis G, Liberopoulos E, Barkas F, Elisaf M. Diabetes mellitus and electrolyte disorders. *World J Clin Cases.* 2014;2:488–96.
11. Nichols GA, Desai J, Elston LJ, et al. Construction of a multisite DataLink using electronic health records for the identification, surveillance, prevention, and management of diabetes mellitus: the SUPREME-DM Project. *Prev Chronic Dis.* 2012;9:E110.
12. Nichols GA, Schroeder EB, Karter AJ, et al. Trends in diabetes incidence among 7 million insured adults, 2006–2011: The SUPREME-DM Project. *Am J Epidemiol.* 2015;181:32–9.
13. Ross TR, Ng D, Brown JS, et al. The HMO Research Network Virtual Data Warehouse: a public data model to support collaboration. *EGEMS (Wash DC).* 2014;2:1049.
14. American Diabetes Association. Classification and diagnosis of diabetes. *Standards of Medical Care in Diabetes – 2019.* *Diabetes Care* 2019; 42 (Supplement 1); S13–S28.
15. Schroeder EB, Donahoo WT, Goodrich GK, Raebel MA. Validation of an algorithm for identifying type 1 diabetes in adults based on electronic health record data. *Pharmacoepidemiol Drug Saf.* 2018;27:1053–9.
16. Guan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care.* 2005;43:1130–9.
17. Chronic Conditions Data Warehouse. Chronic Conditions Data Warehouse condition categories. Available at: <https://www.ccwdata.org/web/guest/condition-categories>. Accessed 6/1/2017.
18. Karter AJ, Warton EM, Lipska KJ, et al. Development and validation of a tool to identify patients with type 2 diabetes at high risk of hypoglycemia-related emergency department or hospital use. *JAMA Intern Med.* 2017;177:1461–70.
19. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and Regression Trees.* Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software; 1984.
20. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* 2015;162:W1–73.

21. **Collins GS, Reitsma JB, Altman DG, Moons KG.** Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med.* 2015;162:55–63.
22. **Altman DG, Vergouwe Y, Royston P, Moons KG.** Prognosis and prognostic research: validating a prognostic model. *BMJ.* 2009;338:b605.
23. **Royston P, Moons KG, Altman DG, Vergouwe Y.** Prognosis and prognostic research: Developing a prognostic model. *BMJ.* 2009;338:b604.
24. R Core Team. R: A language and environment for statistical computing. Available at: <https://www.R-project.org>. Accessed 9/19/2018.
25. **Ripley B.** tree: Classification and Regression Trees. R package version 1.0–39. Available at: <https://CRAN.R-project.org/package=tree>. Accessed 9/19/2018.
26. **Kim K, Thomsen RW, Nicolaisen SK, Hasvold LP, Palaka E, Sorensen HT.** Healthcare resources utilisation and cost associated with elevated potassium levels: a Danish population-based cohort study. *BMJ Open.* 2019;9:e026465.
27. **Rodenburg EM, Visser LE, Hoorn EJ, Ruiter R, Lous JJ, Hofman A, Uitterlinden AG, Stricker BH.** Thiazide and the risk of hypokalemia in the general population. *Journal of Hypertension.* 2014;32:2092–2097.

Publisher's Note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.