Developing Prediction Models for Kidney Stone Disease

Joseph Palko*, Advisor: Jue Wang*

*Department of Mathematics, Union College, Schenectady, NY

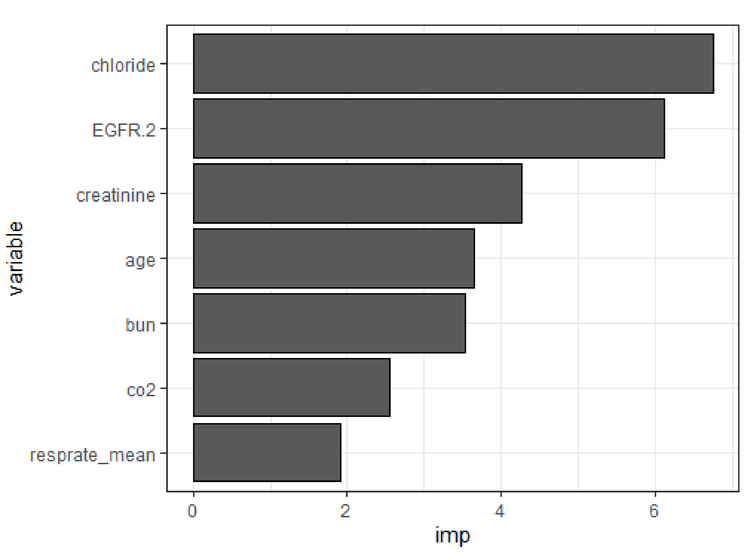
Abstract

Kidney stone disease has become more prevalent through the years, leading to high treatment cost and associated health risks. In this study, we explore a large medical database and machine learning methods to extract features and construct models for diagnosing kidney stone disease. Data of 46,250 patients and 58,976 hospital admissions were extracted and analyzed, including patients' demographic information, diagnoses, vital signs, and laboratory measurements of the blood and urine. We compared the kidney stone (KDS) patients to patients with abdominal and back pain (ABP), patients diagnosed with nephritis, nephrosis, renal sclerosis, chronic kidney disease, or acute and unspecified renal failure (NCA), patients diagnosed with urinary tract infections and other diseases of the kidneys and the uterus (OKU), and patients with other conditions (OTH). We built logistic regression models and random forest models to determine the best prediction outcome.

Methods

We constructed machine learning models to find the best way to categorize patients and predict outcomes, including decision trees, random forests, and logistic regression. We extracted the most important features from 81 total to determine model variables. The performance was quantified using sensitivity (recall or true positive rate), specificity (true negative rate), accuracy, and AUROC (area under the receiver operating characteristics). Sensitivity indicates what percentage of patients with a disease are correctly identified and specificity indicates what percentage of patients without a disease are correctly identified. A way to see how accurate a logistic regression model is, is to use AUROC, or area under the receiver operator characteristic (ROC). This graph gives us a threshold independent measurement of how well this model is compared to a random model. The closer these values are to 1, the better the outcome.

Feature Importance



Results

After exploring many different models including decision trees, random forests, and logistic regression models, we found that the logistic regression models produced the best overall results out of all the models. For KDS VS ABP the most correlated variables were EGFR, CO2, BUN, age, and creatinine. For KDS VS NCA the most correlated variables were Elixhauser comorbidity score and BUN. For KDS VS OKU the most correlated variables were BUN, bands, and creatinine. For KDS VS OTH, the most correlated variables were BUN and creatinine. We have included some of the logistic regression results tables that we created. Also shown below is an ethnicity and gender distribution table.

Table 1. Logistic regression using most correlated variables summary

	KDS VS ABP			KDS VS NCA			KDS VS OKU			KDS VS OTH		
	Avg.	1st Lab Res.	1st Lab Res. and 1st Adm.	Avg.	1st Lab Res.	1st Lab Res. and 1st Adm.	Avg.	1st Lab Res.	1st Lab Res. and 1st Adm.	Avg.	1st Lab Res.	1st Lab Res. and 1st Adm.
AUC	.753	.744	.718	.744	.742	.754	.742	.700	.722	.671	.669	.637
SPECIFICITY	.757	.670	.591	.829	.835	.898	.717	.507	.793	.794	.782	.782
SENSITIVITY	.609	.713	.746	.557	.540	.508	.667	.789	.588	.487	.483	.443
ACCURACY	.670	.695	.683	.821	.825	.883	.708	.560	.751	.787	.775	.776

Table 2. Regression summaries when using EGFR, age, creatinine, CO2, and BUN

	KDS VS ABP			KDS VS NCA			KDS VS OKU			KDS VS OTH		
	Avg.	1st Lab Res.	1st Lab Res. and 1st Adm.	Avg.	1st Lab Res.	1st Lab Res. and 1st Adm.	Avg.	1st Lab Res.	1st Lab Res. and 1st Adm.	Avg.	1st Lab Res.	1st Lab Res. and 1st Adm.
AUC	.753	.744	.718	.748	.741	.752	.658	.648	.632	.679	.671	.649
SPECIFICITY	.757	.670	.591	.788	.744	.758	.677	.724	.754	.845	.881	.903
SENSITIVITY	.609	.713	.746	.627	.657	.657	.554	.495	.444	.450	.394	.323
ACCURACY	.670	.695	.683	.783	.742	.755	.664	.700	.720	.837	.871	.894

Table 3. Ethnicity and gender distribution

	KDS		ABP			NCA			OKU			OTH		
Total	534		451			16701			4620			24295		
	N	%	N	%	p-value	N	%	p-value	N	%	p-value	N	%	p-value
Gender														
Male	311	58.24	245	54.32	.2418	9727	58.27	1	1798	38.92	<.0001	14189	58.40	.9821
Female	223	41.76	206	45.68		6974	41.73		2822	61.08		10106	41.60	
Ethnicity														
Asian	11	2.06	6	1.33		415	2.48		101	2.19		565	2.33	
Black	40	7.49	61	13.53	.0027	2490	14.91	<.0001	357	7.73	.9821	1498	6.17	<.0001
Hispanic	17	3.18	19	4.21		567	3.40		144	3.12		838	3.45	
Native	0	0	0	0		11	0.66		5	1.08		11	0.45	
Other	18	3.37	8	1.77		238	1.43		82	1.77		677	2.79	
Unknown	27	5.06	39	8.65		1303	7.80		482	10.43		3405	14.02	
White	421	77.53	318	70.51	.0016	11677	69.92	<.0001	3449	74.65	.03899	17301	71.21	<.0001

Conclusion

KDS VS ABP Model

- Age, mean respiratory rate, blood chloride, blood creatinine, and blood CO2 levels using the patients' first lab results.
- accuracy of 0.699 and maximized sensitivity with a value of 0.726

KDS VS NCA Model

- Elixhauser score and blood urea nitrogen (BUN) values using the first lab results for patients with first admittance
- Accuracy of 0.883 and maximized specificity of 0.898.

KDS VS OKU Model

- Estimated glomerular filtration rate (EGFR) calculated from the average lab values
- Accuracy of 0.852 and maximized specificity of 0.922

KDS VS OTH Model

- Age, EGFR, BUN, blood creatinine, and blood CO2 using the first lab results for patients with first admittance
- an accuracy of 0.894 and maximized specificity of 0.903.

Possible Model Uses

1. Diagnose patients that come into critical care hospitals
2. Provide a steppingstone for researchers to build off if they want to build kidney stone models for a different population of patients.

References

Gilbert, S. J., & Weiner, D. E. (2014). *NATIONAL KIDNEY FOUNDATION'S PRIMER ON KIDNEY DISEASES* (6th ed.) (D. S. Gipson, M. A. Perazella, & M. Tonelli, Eds.). Philadelphia, PA: National Kidney Foundation.

Starmer, J. StatQuest statistics and machine learning, https://statquest.org