



# Feasibility of stone recurrence risk stratification using the recurrence of kidney stone (ROKS) nomogram

Nicholas L. Kavoussi<sup>1,2,3,4</sup> · Alexandre Da Silva<sup>1,2,3,4</sup> · Chase Floyd<sup>1,2,3,4</sup> · Allison McCoy<sup>1,2,3,4</sup> · Tatsuki Koyama<sup>1,2,3,4</sup> · Ryan S. Hsi<sup>1,2,3,4</sup>

Received: 8 February 2023 / Accepted: 3 April 2023 / Published online: 17 April 2023  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

## Abstract

This study seeks to evaluate the recurrence of kidney stones (ROKS) nomogram for risk stratification of recurrence in a retrospective study. To do this, we analyzed the performance of the 2018 ROKS nomogram in a case–control study of 200 patients (100 with and 100 without subsequent recurrence). All patients underwent kidney stone surgery between 2013 and 2015 and had at least 5 years of follow-up. We evaluated ROKS performance for prediction of recurrence at 2- and 5-year via area under the receiver operating curve (ROC-AUC). Specifically, we assessed the nomogram’s potential for stratifying patients based on low or high risk of recurrence at: a) an optimized cutoff threshold (i.e., optimized for both sensitivity and specificity), and b) a sensitive cutoff threshold (i.e., high sensitivity (0.80) and low specificity). We found fair performance of the nomogram for recurrence prediction at 2 and 5 years (ROC-AUC of 0.67 and 0.63, respectively). At the optimized cutoff threshold, recurrence rates for the low and high-risk groups were 20 and 45% at 2 years, and 50 and 70% at 5 years, respectively. At the sensitive cutoff threshold, the corresponding recurrence rates for the low and high-risk groups were of 16 and 38% at 2 years, and 42 and 66% at 5 years, respectively. Kaplan–Meier analysis revealed a recurrence-free advantage between the groups for both cutoff thresholds ( $p < 0.01$ , Fig. 2). Therefore, we believe that the ROKS nomogram could facilitate risk stratification for stone recurrence and adherence to risk-based surveillance protocols.

**Keywords** Kidney stone · Kidney stone recurrence · Risk stratification

✉ Nicholas L. Kavoussi  
Nicholas.l.kavoussi@vumc.org

Alexandre Da Silva  
alexandre.v.da.silva@vanderbilt.edu

Chase Floyd  
chase.floyd@uscmed.sc.edu

Allison McCoy  
Allison.b.mccoy@vumc.org

Tatsuki Koyama  
tatsuki.koyama@vumc.org

Ryan S. Hsi  
ryan.hsi@vumc.org

## Introduction

The clinical management of kidney stone disease relies on the individual risk of future stone events. Such risk stratification, based on clinical characteristics, is recommended by clinical guidelines to appropriately tailor the intensity and burden of preventative interventions, often prescribed over the long term [1]. However, what constitutes high versus low-risk criteria is not clearly established, and as a result there is poor agreement among clinicians when predicting individual recurrence risk [2].

Within this context, the development of the recurrence of kidney stone (ROKS) nomogram sought to provide individual-level probability of future symptomatic stone events. The nomogram uses 13 features from the clinical history,

<sup>1</sup> Department of Urology, Vanderbilt University Medical Center, 1211 Medical Center Drive, Nashville, TN 37212, USA

<sup>2</sup> School of Medicine, Columbia-University of South Carolina, Columbia, SC, USA

<sup>3</sup> Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA

<sup>4</sup> Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN, USA

including stone characteristics on imaging [3, 4]. However, the nomogram has limited adoption in clinical practice primarily due to its limited predictive performance, including when evaluated in external cohorts [2, 5, 6].

Whether ROKS can be used to discriminate high and low risk populations among kidney stone patients, rather than individual-level risk prediction, is unknown. Therefore, our goal was to assess the performance of the ROKS nomogram to risk-stratify kidney stone patients based on future risk of stone events, including both 2- and 5-year recurrence. To do this, we performed an external evaluation of the nomogram, identified the optimal risk thresholds for high-low risk stratification, and report their respective predictive potential.

## Materials and methods

### Patient cohort

After local institutional review board approval, a retrospective review was performed of all patients who underwent kidney stone surgery from 1997 to 2015 at our institution. Patients were identified using an institutionally maintained and deidentified database of the electronic health record [7]. Patients were identified via Current Procedural Terminology (CPT) codes (see Appendix 1) for having received percutaneous nephrolithotomy (PCNL), ureteroscopy (URS), or shockwave lithotripsy (SWL). Patient demographics and comorbidities based on ICD codes were extracted from the Synthetic Derivative (see Appendix 2). We selected a contemporary cohort of 200 patients (100 without recurrence and 100 with recurrence) with at least annual follow-up for 5 years for their stone disease. All patients had at least yearly imaging that could assess for kidney stone disease. Manual chart review was performed to adjudicate follow-up for all patients.

### Stone recurrence

Stone recurrence episodes after index surgeries were identified by the first occurrence of 1) kidney stone surgeries (CPT codes, appendix (1), due to stone growth new stone on imaging or asymptomatic hydronephrosis, or 2) emergency room visits for symptomatic kidney stones (ICD codes, appendix (2). Then, manual chart review was performed to adjudicate all recurrence episodes. Patients who had secondary planned surgeries or recurrence events within 3 months of index

surgery were excluded to ensure that recurrence episodes were not immediately attributable to a staged procedure or a complication from the index surgery.

### Statistical methods

We identified the performance of each predictor in the 2018 ROKS nomogram for our patient population via Cox proportional hazard regression analysis [4]. Estimates of recurrence at 2- and 5-year follow-up were calculated for each patient. These predictions were used to assess the discriminative performance of the nomogram for our patient population using the area under the receiver operating curve (AUC-ROC).

Then, we evaluated the ability of the ROKS nomogram to stratify patients based on low or high risk of recurrence at two thresholds along the ROC obtained from: a) an optimized cutoff point (i.e., a point optimized for both sensitivity and specificity along the ROC), and b) a sensitive cutoff point (i.e., a point with high sensitivity (0.80) and low specificity along the ROC). The points were chosen in line with current statistical guidelines in urology [8]. This sensitive cutoff point was selected to determine the clinical utility of the ROKS nomogram as a tool for identifying high risk of stone recurrence. We performed a Cox proportional hazards regression analysis to evaluate the performance of each individual predictor included in the ROKS nomogram in our patient population. Time to recurrence was compared between the risk groups via survival analysis with Kaplan–Meier (KM) estimation and log rank testing. All analyses were performed with  $p < 0.05$  as significant and conducted in ‘R’ (R Foundation for Statistical Computing, Vienna, Austria)[9].

## Results

Patients followed up over a mean  $\pm$  SD time of  $96 \pm 38$  months. The mean  $\pm$  SD time to recurrence was  $29 \pm 32$  months. Table 1 summarizes patient demographics and clinical characteristics.

ROKS nomogram demonstrated fair discriminative ability in predicting risk of recurrence with an AUC-ROCs of 0.67 and 0.63 at 2 and 5 years, respectively (Fig. 1a, b). At 2 years recurrence prediction, sensitivity and specificity were 0.60 and 0.68 at the optimized cutoff point, and 0.80 and 0.45 at the sensitive cutoff point, respectively. At 5 years

**Table 1** Patient characteristics

| Demographics                              | Overall (N=200) | Recurrence (N=100) | No recurrence (N=100) | P value |
|---|-----------------|--------------------|-----------------------|---------|
| Age (mean $\pm$ SD)                       | 49 $\pm$ 15     | 47 $\pm$ 15        | 52 $\pm$ 6            | 0.03    |
| Body mass index (mean $\pm$ SD)           | 30.4 $\pm$ 8.0  | 30 $\pm$ 7         | 31 $\pm$ 9            | 0.47    |
| Gender, male (N, %)                       | 101 (50)        | 45 (45)            | 56 (56)               | 0.15    |
| Race, Caucasian (N, %)                    | 174 (87)        | 88 (88)            | 86 (86)               | 0.67    |
| Race, African American (N, %)             | 12 (6)          | 6 (6)              | 6 (6)                 | 0.99    |
| Ethnicity, Hispanic (N, %)                | 6 (3)           | 2 (2)              | 4 (4)                 | 0.41    |
| Past medical history                      |                 |                    |                       |         |
| Family history of nephrolithiasis (N, %)  | 86 (43)         | 50 (50)            | 36 (36)               | 0.05    |
| Hypertension (N, %)                       | 78 (39)         | 36 (36)            | 42 (42)               | 0.38    |
| Diabetes mellitus (N, %)                  | 38 (19)         | 17 (17)            | 21 (21)               | 0.47    |
| Bowel disease* (N, %)                     | 38 (19)         | 25 (25)            | 13 (13)               | 0.03    |
| Gout (N, %)                               | 7 (3.5)         | 4 (4)              | 3 (3)                 | 0.70    |
| Hyperparathyroidism (N, %)                | 2 (1)           | 1 (1)              | 1 (1)                 | 0.99    |
| Cystinuria (N, %)                         | 2 (1)           | 1 (1)              | 1 (1)                 | 0.99    |
| Index surgery                             |                 |                    |                       |         |
| Ureteroscopy and laser lithotripsy (N, %) | 160(80)         | 76 (76)            | 84 (84)               | 0.16    |
| Percutaneous Nephrolithotomy (N, %)       | 38(19)          | 23 (23)            | 15 (15)               | 0.15    |
| Shock-wave lithotripsy (N, %)             | 2 (1)           | 1 (1)              | 1 (1)                 | 0.99    |

SD standard deviation; absolute numbers (N) and percentages reported

\*Bowel disease included patients with history of bowel resection, chronic diarrhea, inflammatory bowel disease, intestinal malabsorption, or gastroesophageal reflux disease

recurrence risk prediction, sensitivity and specificity were 0.54 and 0.72 at the optimized cutoff point, and 0.80 and 0.26 at the sensitive cutoff point, respectively (Fig. 1c).

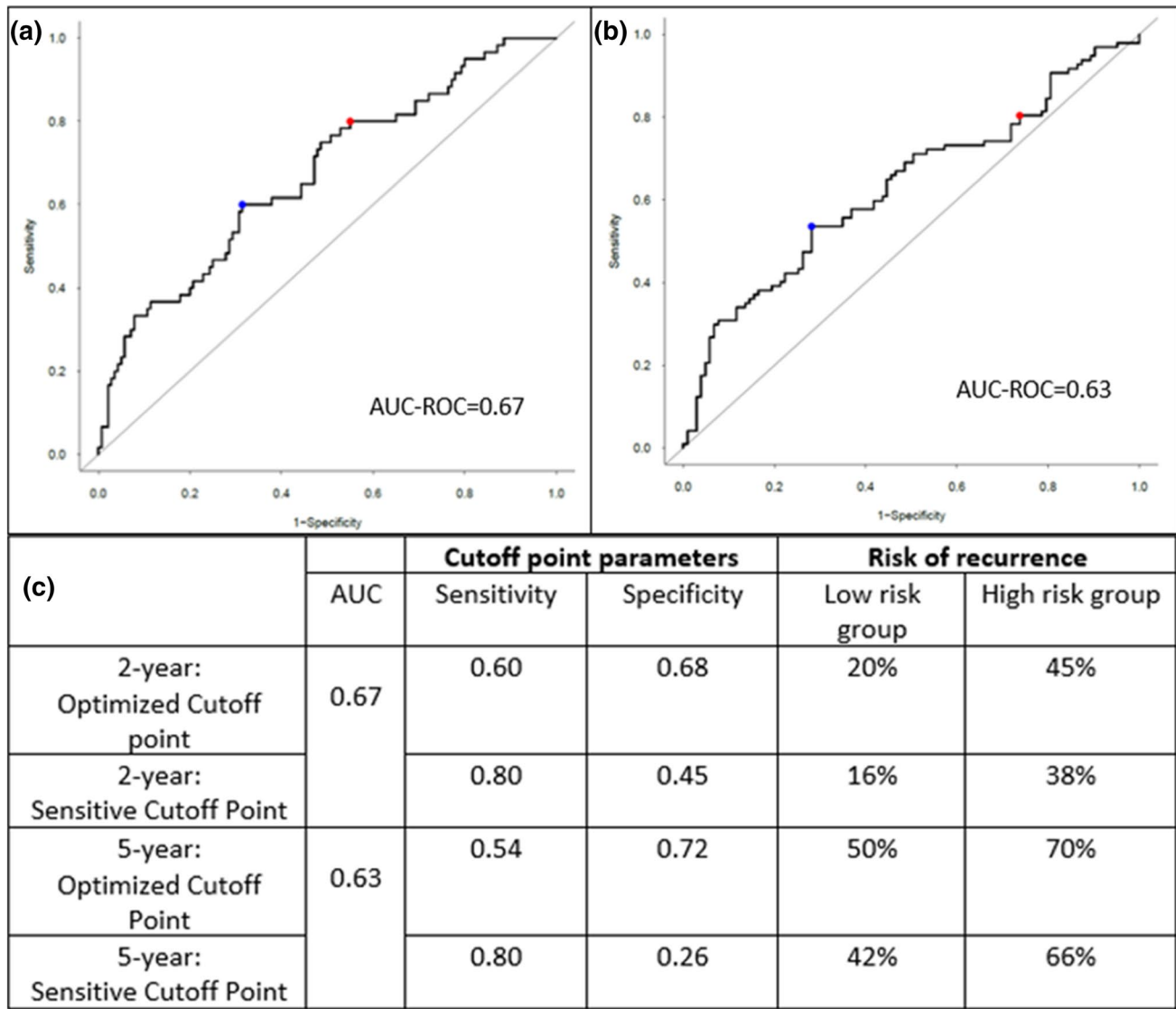
For the 2 year recurrence prediction, the AUC-ROC values of 0.623 for the optimized cutoff point and 0.505 for the sensitive cutoff point were examined. At the optimized cutoff threshold, 120 patients were classified as low, while 80 were classified as high risk. Recurrence rates for the low and high-risk groups were 20 and 45% at 2 years, and 50 and 70% at 5 years, respectively. Then, for the sensitive cutoff threshold, 75 patients were classified as low risk and 125 patients were classified as high risk. The corresponding recurrence rates for the low and high-risk groups was of 16 and 38% at 2 years, and 42 and 66% at 5 years, respectively. For both thresholds, KM analysis revealed a significant recurrence-free probability between the groups ( $p < 0.01$ , Fig. 2a, b).

Of the 13 predictors used by the ROKS nomogram, family history of stone disease, any calcium oxalate monohydrate stone, and having a stone  $> 6$  mm in diameter all associated with risk of recurrence in the cohort (Table 2, HRs: 1.03, 1.8, 0.3, 0.61, and 0.26, respectively).

## Discussion

Our study suggests that the ROKS nomogram modestly predicts recurrent stone events when including a broad definition of recurrence. This further emphasizes the limitations of utilizing the nomogram for counseling kidney stone patients. However, we found that the ROKS nomogram successfully stratified patients by risk-level (i.e., low vs. high) for stone recurrence based on different cutoff points. For example, using a sensitive cutoff point for stratification predicted the risk of kidney stone recurrence for the low and high-risk groups to be 16 and 38% at 2 years, and 42 and 66% at 5 years, respectively. This finding could facilitate the development of surveillance protocols for kidney stone recurrence by risk group.

Currently, the American Urologic Association (AUA) recommends postoperative imaging after stone surgery to assess for residual stone and silent hydronephrosis with periodic imaging thereafter to assess for new stone formation or stone growth [1, 10]. However, there is limited evidence supporting specific protocols for stone surveillance.

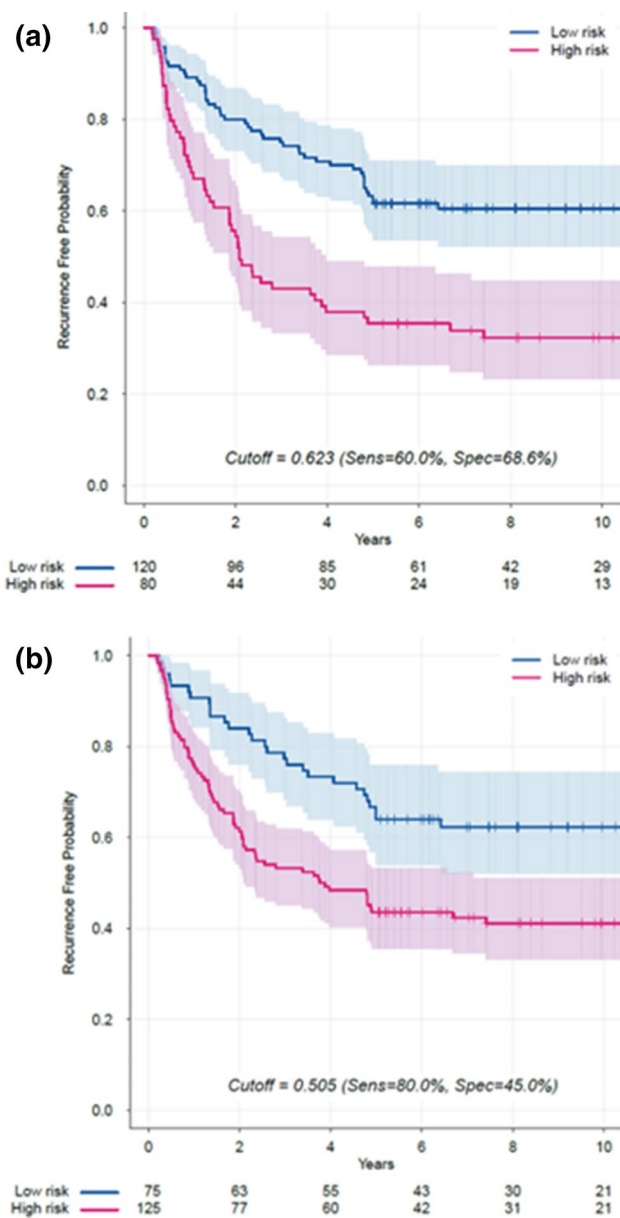


**Fig. 1** Prediction of overall recurrence at **a** Two-years (the blue point represents the optimized cutoff point with sensitivity (0.60) and specificity (0.68); the red point represents the sensitive cutoff point with sensitivity (0.80) with corresponding specificity of 0.45). **b** Five-years (the blue point represents the optimized cutoff point with sensi-

tivity (0.54) and specificity (0.72); the red point represents the sensitive cutoff point with sensitivity (0.80) with corresponding specificity of 0.26), and **c** nomogram prediction of recurrence at 2 and 5-years for each cutoff point. (AUC = area under the curve)

Unnecessary imaging can be costly and lead to excess radiation exposure [11, 12]. The lack of data supporting long term imaging use has led to a wide variation follow-up imaging utilization after stone surgery. For example, in an evaluation of MarketScan data, 29 and 15% of patient who underwent percutaneous nephrolithotomy do not undergo postoperative

imaging by 3 and 12 months, respectively [13]. Comparatively, 55 and 39% of patients who underwent ureteroscopy, and 23 and 16% of patients who underwent shock wave lithotripsy had no postoperative imaging at 3 and 12 months, respectively [14]. Both studies, moreover, reported significant variations in postoperative imaging modalities as well,



**Fig. 2** KM curves evaluating time to recurrence between low and high risk groups for using a ROC-cutoff point threshold of **a** 0.623 by optimizing sensitivity and specificity, and **b** 0.505 for high sensitivity, low specificity. The shaded areas represent the 95% confidence interval of the respective curves

emphasizing a lack of consensus for kidney stone surveillance. Thus, a rational approach for follow-up of kidney stone patients is to tailor the frequency of surveillance imaging to specific patient risk factors.

As recurrence risk varies across patients, it is necessary to personalize stone care to optimize stone surveillance

strategies. However, clinicians may not be able to distinguish patients with low or high risk of recurrence, particularly compared to the ROKS nomogram [2]. Additionally, prior attempts at external validation of the ROKS nomogram has been limited. Previously, Iremashvili et al. demonstrated only moderate discrimination of the nomogram (0.655 at 2 years and 0.605 at 5 years) for prediction of recurrence risk [15]. We found a similar, modest prediction of risk with a broader definition of recurrence (i.e., including both symptomatic and radiographic recurrence). However, despite limitations of the nomogram as an external prediction tool, our findings suggest it could be used for stone risk stratification. By stratifying patients as low or high-risk for recurrence, the nomogram could help guide follow-up management and improve care for kidney stone patients.

There are several limitations to this study. The retrospective design cannot account for confounders and bias due to omitted variables. All patients were identified after being referred to a tertiary medical center and may have a higher baseline risk for stone recurrence. Furthermore, the nomogram performance may differ for other definitions of kidney stone recurrence [6]. Though we reviewed imaging reports, we were unable to review kidney stone imaging directly. Therefore, we could not differentiate whether recurrence was from a residual fragment or a new stone. We additionally did not include the impact of medication on stone recurrence as it is not an included parameter on the ROKS nomogram. Additionally, as this study was done at a single institution, we cannot assess for recurrence events that occurred at other institutions or at home. It is possible further analysis with a larger dataset could refine our models. Despite these limitations, this study demonstrates the feasibility of kidney stone risk stratification using the ROKS nomogram.

### Conclusion

We found that the ROKS nomogram has potential to serve as a tool for recurrence risk stratification into lower and higher risk groups, despite only modest prediction of stone events when including a broad definition of stone recurrence, including symptomatic and radiographic recurrence. This could facilitate adherence to risk-based surveillance protocols.

**Table 2** Performance of individual predictors from the ROKS nomogram in the cohort

|   | No recurrence<br>( <i>N</i> = 100) | Recurrence<br>( <i>N</i> = 100) | HR (95% CI)      | <i>p</i> |
|---|------------------------------------|---------------------------------|------------------|----------|
| Age at last stone episode (years, mean, SD)                       | 52 ± 15                            | 47 ± 15                         | 1.01 (0.99–1.02) | 0.5      |
| BMI at last stone episode   | 31 ± 9                             | 30 ± 7                          | 1.03 (0.99–1.1)  | 0.06     |
| Sex (male, <i>N</i> )   | 56                                 | 45                              | 0.83 (0.49–1.44) | 0.52     |
| Family history of stone disease, <i>N</i>                         | 37                                 | 49                              | 1.8 (1.07–3.10)  | 0.03     |
| Any stone to be found uric acid or brushite or struvite, <i>N</i> | 20                                 | 22                              | 0.38 (0.15–0.97) | 0.04     |
| Any stone to be found calcium oxalate monohydrate, <i>N</i>       | 80                                 | 78                              | 0.30 (0.13–0.67) | < 0.01   |
| Ureterovesical junction stone, <i>N</i>                           | 8                                  | 28                              | .087 (0.48–1.57) | 0.64     |
| Renal pelvis/lower pole stone, <i>N</i>                           | 31                                 | 53                              | 0.61 (0.34–1.1)  | 0.1      |
| Prior incidental stone seen on imaging, <i>N</i>                  | 28                                 | 41                              | 0.85 (0.49–1.48) | 0.57     |
| Prior suspected event without actual stone seen, <i>N</i>         | 81                                 | 85                              | 1.78 (0.84–3.7)  | 0.13     |
| Number of stones in both kidneys, <i>N</i>                        |                                    |                                 |                  |          |
| 1   | 36                                 | 27                              | Ref              |          |
| > 2   | 64                                 | 73                              | 1.2 (0.66–2.04)  | 0.6      |
| Diameter of largest stone, <i>N</i>                               |                                    |                                 |                  |          |
| < 3 mm  | 6                                  | 4                               | Ref              |          |
| 3–6 mm  | 18                                 | 20                              | 0.67(0.21–2.1)   | 0.51     |
| > 6 mm  | 75                                 | 77                              | 3.84 (1.3–10)    | 0.02     |

*CI* confidence interval, *HR* hazard ratio

## Appendix 1

(See Table 3).

**Table 3** ICD-9 and ICD-10 codes used to identify incident stones

|                              |                                       |
|------------------------------|---------------------------------------|
| Kidney stone, ureteral stone | 592, 592.1, 592.9                     |
| Hematuria                    | 5.997                                 |
| Flank Pain                   | 789.0, 789.01, 789.02, 789.03, 789.04 |
| Hydronephrosis               | 591                                   |

## Appendix 2

(See Table 4).

**Table 4** ICD-9 and ICD-10 codes used to identify comorbidities

|   |  |
|---|--|
| Bowel resection                               | V45.3, V45.86, Z98.0, Z98.84                               |
| Cerebrovascular accident                      | 434.00–434.91, I63.0–63.9                                  |
| Coronary artery disease/Myocardial infarction | 414.00–414.07, 410.0–410.92, I121.0–121.01, I125.110–125.9 |
| Cystinuria                                    | 270.0, E72.01  |
| Diarrhea                                      | 564.1–564.5, K58.0–59.1                                    |
| Epilepsy                                      | 345.00–345.91, G40.A01–40.909                              |
| Gastroesophageal reflux disease               | 530.81, K21.9 Gout 274–274.9, M10.411–10.9                 |
| Hyperlipidemia                                | 272.0–272.4, E78.00–78.5                                   |
| Hyperparathyroidism                           | 252.01, E21.0  |
| Hypertension                                  | 401–401.9, I10–16.9  |
| Immobility                                    | 728.3, M62.3–62.89   |
| Inflammatory bowel disease                    | 555.0–556.9, K50.011–51.919                                |
| Intestinal malabsorption                      | 579.0–579.9, K90.1–90.9                                    |
| Migraine                                      | 339.00–339.89, G44.001–44.89                               |
| Osteoporosis                                  | 733.00–733.03, M81.0–M81.8                                 |
| Type 2 diabetes                               | 250.0–250.92, E11.0–11.9, E13.0–13.9                       |

**Author contributions** Nicholas Kavoussi and Ryan Hsi conceived the project, did data analysis and writing the main manuscript text. Tatsuki Koyama did the statistical analysis and generated the figures. Allison McCoy, Alexandre Da Silva, and Chase Floyd did data collection, analysis and contributed to the main text.

**Data Availability** The derived data supporting the findings of this study are available from the corresponding author (NLK) on request.

## Declarations

**Conflict of interest** The authors declare no competing interests.

## References

- Pearle MS, Goldfarb DS, Assimos DG et al (2014) Medical management of kidney stones: AUA guideline. *J Urol* 192:316–324. <https://doi.org/10.1016/j.juro.2014.05.006>
- Forbes CM, McCoy AB, Hsi RS (2021) Clinician versus nomogram predicted estimates of kidney stone recurrence risk. *J Endourol* 35:847–852. <https://doi.org/10.1089/end.2020.0978>
- Rule AD, Lieske JC, Li X et al (2014) The ROKS nomogram for predicting a second symptomatic stone episode. *J Am Soc Nephrol JASN* 25:2878–2886. <https://doi.org/10.1681/ASN.2013.91011>
- Vaughan LE, Enders FT, Lieske JC et al (2019) Predictors of symptomatic kidney stone recurrence after the first and subsequent episodes. *Mayo Clin Proc* 94:202–210. <https://doi.org/10.1016/j.mayocp.2018.09.016>
- Iremashvili V, Li S, Penniston KL et al (2019) External validation of the recurrence of kidney stone nomogram in a surgical cohort. *J Endourol* 33:475–479. <https://doi.org/10.1089/end.2018.0893>
- D’Costa MR, Pais VM, Rule AD (2019) Leave no stone unturned: defining recurrence in kidney stone formers. *Curr Opin Nephrol Hypertens* 28:148–153. <https://doi.org/10.1097/MNH.0000000000000478>
- Roden DM, Pulley JM, Basford MA et al (2008) Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin Pharmacol Ther* 84:362–369. <https://doi.org/10.1038/clpt.2008.89>
- Assel M, Sjoberg D, Elders A et al (2019) Guidelines for reporting of statistics for clinical research in urology. *J Urol* 201:595–604. <https://doi.org/10.1097/JU.0000000000000001>
- R Core Team (2021) R: A Language and environment for statistical computing
- Fulgham PF, Assimos DG, Pearle MS, Preminger GM (2013) Clinical effectiveness protocols for imaging in the management of ureteral calculous disease: AUA technology assessment. *J Urol* 189:1203–1213. <https://doi.org/10.1016/j.juro.2012.10.031>
- Rodger F, Roditi G, Aboumarzouk OM (2018) Diagnostic accuracy of low and ultra-low dose ct for identification of urinary tract stones: a systematic review. *Urol Int* 100:375–385. <https://doi.org/10.1159/000488062>
- Brisbane W, Bailey MR, Sorensen MD (2016) An overview of kidney stone imaging techniques. *Nat Rev Urol* 13:654–662. <https://doi.org/10.1038/nrur.2016.154>
- Dai JC, Ahn JS, Holt SK et al (2018) National imaging trends after percutaneous nephrolithotomy. *J Urol* 200:147–153. <https://doi.org/10.1016/j.juro.2018.01.078>
- Ahn JS, Holt SK, May PC, Harper JD (2018) National imaging trends after ureteroscopic or shock wave lithotripsy for nephrolithiasis. *J Urol* 199:500–507. <https://doi.org/10.1016/j.juro.2017.09.079>
- Iremashvili V, Li S, Penniston KL et al (2019) Role of residual fragments on the risk of repeat surgery after flexible ureteroscopy and laser lithotripsy: single center study. *J Urol* 201:358–363. <https://doi.org/10.1016/j.juro.2018.09.053>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.