Predicting acute kidney injury following open partial nephrectomy treatment using SAT-pruned explainable machine learning model

Teddy Lazebnik1†, Zaher Bahouth2†, Svetlana Bunimovich-Mendrazitsky3 and Sarel Halachmi2

Abstract

Background: One of the most prevalent complications of Partial Nephrectomy (PN) is Acute Kidney Injury (AKI), which could have a negative impact on subsequent renal function and occurs in up to 24.3% of patients undergoing PN. The aim of this study was to predict the occurrence of AKI following PN using preoperative parameters by applying machine learning algorithms.

Methods: We included all adult patients (n = 723) who underwent open PN in our department since 1995 and on whom we have data on the pre-operative renal function. We developed a random forest (RF) model with Boolean satisfaction-based pruned decision trees for binary classification (AKI or non-AKI). Hyper-parameter grid search was performed to optimize the model's performance. Fivefold cross-validation was applied to evaluate the model. We implemented a RF model with greedy feature selection to binary classify AKI and non-AKI cases based on pre-operative data.

Results: The best model obtained a 0.69 precision and 0.69 recall in classifying the AKI and non-AKI groups on average (k = 5). In addition, the model's probability to correctly classify a new prediction is 0.75. The proposed model is available as an online calculator.

Conclusions: Our model predicts the occurrence of AKI following open PN with (75%) accuracy. We plan to externally validate this model and modify it to minimally-invasive PN.

Keywords: AKI prediction, SAT pruned random forest, PN treatment complication prediction

Background

Renal cell carcinoma (RCC) represents about 3% of all cancer-related cases in 2018, with the highest incidence occurring in Western countries [1]. During the last decades, stage migration towards localized disease has occurred [2]. Partial nephrectomy (PN) is the treatment of choice for localized cT1 renal masses [3]. The main advantage of PN is the preservation of renal function compared to radical nephrectomy [4]. One of the adverse effects of PN is post-operative acute kidney injury (AKI), which increases the risk of long-term chronic kidney disease (CKD) with its consequences, including decreased overall survival [5], although some studies questioned its impact on long-term renal function [6]. The prevalence of AKI following PN is reported to be up around 25% and is dependent on the surgical approach, patient baseline...
characteristics, and the definition of AKI used in each study [7].


decision-Making          (2022) 22:133

Machine learning (ML) based models which predict different clinical properties have been shown to be a useful tool [8] and particularly in clinical practice [9]. ML models can be classified into three main subtypes: classification, search, and prediction. In this paper, we focus on the latter in order to predict AKI following PN.

Prediction ML models provided with retrospective data are able to find complex statistical connections between different parameters (this step is usually referred to as the learning process) [8]. As a result, upon providing a new set of parameters, these models are able to predict, with fair accuracy, the outcome one wishes to retrieve. Weng et al. [10] used four ML algorithms to predict cardiovascular risk, showing improvement in all four compared to standard algorithms. A study by Wu et al. [11] developed an ML model to predict fatty liver disease. The authors used five different ML algorithms on the same data, where RF showed the best results. Several authors used different ML models to predict medical AKI in hospitalized patients [12, 13]. For instance, Kate el al. presented a framework in which AKI is continually predicted automatically from EHR data over the entire hospital stay using [14]. In addition, Gameiro et al. reviewed multiple artificial intelligence based models for AKI risk prediction [15]. The authors concluded that real-time implementation of ML-based AKI risk models is a promising approach as these do not require additional AKI biomarker testing.

Specifically, previous studies investigating the performance of ML models in predicting AKI have yielded promising results [16, 17]. However, the accuracy of these models is not optimal, and we thought that by using an ML model, we can increase the accuracy of these models. In this study, we aim to apply an explainable ML model to predict AKI in patients undergoing open PN. We hypothesized that ML models could identify and learn from pre-operative parameters and predict the AKI outcome. A self-explainable prediction system that is based on ML was then built and deployed online. A schematic view of the workflow of the proposed framework is shown in Fig. 1.

Methods
Data acquisition
Since 1995, we have been continuously extending our open PN database to include surgical and oncological parameters. For this particular study, we included all adult (> 18 years) patients who underwent open PN for enhancing solid renal mass and then split the data into AKI and non-AKI. Patients with a solitary kidney or multiple tumors were excluded from this study. Therefore, the PN database includes 723 patients. Renal function was assessed the day before surgery, on the day of the surgery, and on a daily basis after the surgery until discharge which more often than not was on post-operative day 3.

Operative technique
An extraperitoneal, extrapleural supra-11th rib incision was done on the operated side. IV Mannitol was given before clamping the renal artery. In situ renal hypothermia was done by cooling the surface of the kidney with ice slush for 10–15 min immediately after clamping the renal artery. The tumor was enucleated with a minimal rim of normal parenchyma. Renorrhaphy was done using either 2/0 VICRYL interrupted sutures or tissue adhesive BioGlue (CryoLife, Atlanta, GA). A more detailed surgical technique has been previously published by our group [18].

Renal function assessment
Baseline serum Creatinine (sCr) was measured the day before surgery. We used both the RIFLE (risk, injury, failure, loss of kidney function, and end-stage renal failure) [19] and AKIN (Acute Kidney Injury Network) [20] criteria to define AKI, comparing each of the
post-operative renal function assessments to the baseline level. AKI was defined as the occurrence of one of the following conditions: (1) an increase in serum Creatinine of ≥ 0.5 times above baseline in the first week following surgery, (2) an increase in sCr by ≥ 0.3 mg/dl(≥ 26.5 mmol/l) above baseline in the 48 h window post-operatively, or (3) reduction of more than 25 percent of the estimated Glomerular Filtration Rate in the 7 days period after surgery. In total, 231 patients developed AKI based on the aforementioned criteria and constituted the AKI group, and 492 did not develop AKI and therefore were classified as non-AKI. 723 patients is considered a large enough set to use for the methods shown in the following sections [21].

**Data split**
In order to develop ML algorithm, the study population was compiled into a data set, split into a training cohort from which the proposed algorithm was derived and a validation cohort on which the model was applied and tested. The training cohort was derived from a random sampling of 80% of the data set, and the validation cohort comprised the remaining 20%. The division process was repeated 1000 times looking for the optimal split that ensures no statistically significant differences between the two cohorts in demographics or AKI outcome. The split was carried on such that the divisions are minimizing the differences of the age, gender, smoking years, and AKI parameters in both the training and validation cohorts. The distribution of the parameters age, smoking, gender, and AKI in both these cohorts are shown in Eq. (1).

| Parameter | Training Cohort | Validation Cohort |
|-----------|----------------|------------------|
| Age       | 61.23 ± 12.05  | 60.92 ± 13.14    |
| Smoking   | 17.61% ± 38.15%| 18.34% ± 38.74%  |
| Gender    | Male : 37.98%, Female : 62.02% | Male : 35.45%, Female : 64.55% |
| AKI       | 46.84% with AKI | 38.54% with AKI  |

(1)

**Algorithm**
We used the random forest (RF) ML algorithm [22]. We selected the RF algorithm because it can provide a simple explanation of the model’s prediction to healthcare professionals while obtaining a good accuracy on a relatively small data set [23]. We applied the proposed binary AKI prediction decision tree (DT) algorithm on the training cohort and then validated it on the validation cohort that was completed using the sklearn library with Python 3.5. The model’s hyper parameters were determined using the grid search method [24] (see Sect. 2.9) and fivefold cross-validation on the training cohort to determine the values which led to the best performance.

**Feature selection**
We performed a feature selection in the following order: first, we manually filtered the features available before the surgery (marked as F). Afterward, we evaluated the model’s accuracy, picking one feature from F. The feature that resulted in the model’s highest accuracy was chosen F1. Then, an additional feature from the remaining feature set (F\F1) was added to the chosen feature set from the previous step such that the model’s accuracy was the highest between all combinations. The process was repeated until the gain in the model’s accuracy upon adding a new feature became less than 1%.

**Model pruning**
After training the model, we transformed each DT in the RF into a respective Boolean satisfaction problem (SAT). Each branch was converted into a Boolean condition: \( x_1 \land x_2 \land \cdots \land x_n \) where \( x_i \)’s \( i=1 \) are the conditions in each node in the branch and \( r \) was the result label node. Branches with the same result label \( r \) were stitched together using the’or’ logical operator (\( \lor \)). Afterward, each Boolean condition was reduced to the minimal Boolean condition that satisfied the same inputs. The result of Boolean condition was converted back into a DT.

**Statistical analysis**
We performed a fivefold cross-validation to evaluate the model’s accuracy. The data was divided into five cohorts where four cohorts were used for the training cohort and one for the testing cohort. The process was repeated five times, allowing each patient to be included in both the training and test cohorts. The receiver operating characteristic (ROC) curve was used to measure the model’s classification ability. At each point, the recall and precision were presented in correspondence with a specific decision threshold. The area under the ROC curve (AUC) was used to quantify the model’s classification ability. Finally, the importance of each feature depended on the reduction of classification accuracy caused by removing the feature (e.g., information gain) [25].
Hyper-parameter fine-tuning

We performed hyper-parameter fine-tuning using the grid search method, based on the model’s accuracy \[24\]. The grid search was performed on

\[ H := [\text{depth}, \text{MSPL}, \text{LC}, n], \]

\[ F := [\text{size}, \text{renal}, \text{age}, \text{baseHB}, \text{IT}, \text{weight}, \text{height}, \text{creatinine}], \]

where depth is an individual DT tree depth; MSPL is the minimal number of samples for a leaf; LC is the leaf count; and n is the number of trees in the RF model.

Results

Decision features

Implementation of the method described in Sect. 2.6 on 31 features (see Additional file 1: Appendix) resulted in a set of eight features

\[ F := [\text{size}, \text{renal}, \text{age}, \text{baseHB}, \text{IT}, \text{weight}, \text{height}, \text{creatinine}], \]

where size is the size of the tumor in centimeters; renal is the RENAL score; age is the patient’s age in years at the time of the surgery; baseHB is the baseline hemoglobin in g/dL; IT is the ischemia time in minutes; weight is the patient’s weight in kilograms at the time of the surgery; height is the patient’s height in centimeters at the time of the surgery; and creatinine is the baseline pre-operative Creatinine in mg/dL. The model found that IT contributed significantly to the accuracy of the model. However, IT is surgical parameter, and is not available beforehand. In order to overcome this, we defined a feature called IT* which is an estimate of the real IT. The IT* is obtained using the k-nearest neighbors (KNN) algorithm (where \( k = 5 \) and the distance metric is weighted by distance and the average is obtained using the grid search method) on the other seven features which are available before the surgery. To evaluate the quality of the IT* feature compared to the original IT feature, we performed a fivefold test on the IT feature with the KNN algorithm. A linear regression on the values (IT, IT*) was obtained, resulting in a coefficient of determination \( R^2 \) of 0.879. Namely, the IT* feature well estimate the IT feature and therefore, we were able to replace the feature space to:

\[ F^* := [\text{size}, \text{renal}, \text{age}, \text{baseHB}, \text{IT}^*, \text{weight}, \text{height}, \text{creatinine}]. \]

A Pearson correlation coefficient between all pairs of \( F \) was then obtained, showing a 0.57 correlation between the renal and size features and 0.45 correlation between patient height and weight. The first correlation is trivial as the renal score includes the size. In addition, the second correlation is already reported in other studies \[26\]. All other combinations of features from \( F \) have absolute correlation below 0.3, supporting the fact that the feature-space is mostly linearly independent.

AKI classification model

We trained the RF model (see Sect. 2.5) on the clinical data set as described in Sect. 2.4 on the feature set \( F \). Afterward, we performed hyper-parameter fine-tuning as described in Sect. 2.9. Then, we carried out pruning on the best model (see Sect. 2.7). As a result, we obtained a model with 107 DTs; each one of these DTs had up to five levels of depth. The number of leaves is different for each tree in the RF due to the pruning process. The model was validated on the validation set using fivefold cross-validation analysis. The precision obtained 0.69 ± 0.085. Similarly, the recall obtained 0.69 ± 0.062. The features’ importance is presented in Table 1.

Furthermore, we derived the ROC curve of the model, as shown in Fig. 2. The AUC was found to be 0.75. In addition, the average confusion matrix was: 

![Confusion Matrix]

Table 1  Model’s features importance

| Feature | Size | Renal | Age | BaseHB | Weight | Height | Creatinine | IT | IT* |
|---------|------|-------|-----|--------|--------|--------|------------|----|-----|
| Original importance | 0.13 | 0.11 | 0.17 | 0.05  | 0.15  | 0.05  | 0.21  | 0.13 | 0  |
| Estimated importance | 0.14 | 0.11 | 0.19 | 0.06  | 0.12  | 0.03  | 0.23  | 0   | 0.12 |
| Core importance | 0.17 | 0.11 | 0.21 | 0.06  | 0.16  | 0.06  | 0.24  | 0   | 0   |

The original importance is the one obtained by a model that uses IT. The estimated importance is the one obtained by estimating IT (e.g., IT* using the other seven features). The core importance is the importance of the seven features as the weighted average in contribution to the final prediction of IT*.


Discussion

AKI following PN is a unique entity, which significantly differs from medical and post-surgical AKI; in addition to the common risk factors for medical AKI, patients undergoing PN have increased risk for AKI due to the associated blood loss and relative hypovolemia and, more importantly, the clamping of the renal artery and the loss of functional tissue. Several studies reported an increased risk of chronic kidney disease (and mortality) in patients who develop AKI [5, 27–29]. The incidence of AKI following PN varies depending on several parameters, including surgical approach, the definition used for AKI, and the cohort reported in each study. In a recent study by Tachibana et al., the authors reported less than 11% AKI following robotic PN and almost 50% incidence following open PN [30]. Our results demonstrate that the development of AKI following PN can be accurately predicted based on clinical information routinely collected before surgery. The proposed models performed well according to all evaluation criteria and achieved a higher AUC and accuracy score compared to the classical scoring methods. The proposed model has similar scores to the modern ML-based models [16, 32] and AUC in accordance with previous large studies on general intensive-care unit patients developing medical AKI [31].

Our results agree with a wide range of ML-based AKI risk prediction models from recent years [12, 13].

\[
\begin{bmatrix}
\text{True} & \text{False} \\
\text{Positive} & 0.34 & 0.41 \\
\text{Negative} & 0.07 & 0.18
\end{bmatrix}
\]

(4)

Interface

The model has been deployed as a web service.1 Figure 3 shows the model’s interface as a web service.

1 https://teddy4445.github.io/cancer-AKI-predictor-gui/predictor.html.
particular, our model can be used at any point before the surgery. As such, it is agnostic in time, similar to the approach of several AKI prediction tools [15]. We used a RF model (as an ensemble of DT models) as the ML algorithm for our model in order to take advantage of the explainable property of this model. In addition, by using the SAT pruning algorithm, we were able to obtain the shortest, and therefore, easiest to understand explanation for each prediction. This explanation provides the treating Urologist with the ability to agree or disagree with the model on unique cases and has a second validation process on the model’s prediction based on the personnel’s wider understanding of the patient’s condition (i.e., man-in-the-loop) [33].

The proposed AKI prediction model could be publicly available as an online prognostic calculator, providing a platform for future AKI-prediction studies, and complementing existing risk assessment scores [34]. One could argue that a better endpoint would be the risk of CKD following PN, which is the most important endpoint. In this study, we aimed to predict AKI as it was demonstrated to increase the risk for CKD. A model that predicts CKD is harder to build, and this is one of our future projects.

The main limitation of our study is being an open PN cohort, and it is yet to be determined if it will apply to patients who undergo minimally invasive surgery. The second limitation is the relatively small cohort for this model, although others used smaller cohorts and reported good results [21]. Another limitation is the lack of external validation. Moreover, we used an estimated parameter, ischemia time, to predict the AKI based on pre-operative parameters. However, despite these limitations, our model can predict AKI with relatively high accuracy (75%). In conclusion, our ML model can reliably predict AKI following open PN. Future possible research is to extend the size of the database and perform stability analysis for controversial cases in order to improve the robustness of the proposed model.

**Conclusion**

Machine learning algorithms can predict the risk of AKI following open PN.

**Abbreviations**

PN: Partial nephrectomy; AKI: Acute kidney injury; CKD: Chronic kidney disease; ML: Machine learning; RF: Random forest; DT: Decision tree; SAT: Boolean satisfaction problem; ROC: Receiver operation characteristic; AUC: Area under the ROC curve; IT: Ischemia time; KNN: K-nearest neighbors.

**Supplementary Information**

The online version contains supplementary material available at https://doi.org/10.1186/s12911-022-01877-8.

**Acknowledgements**

Not applicable.

**Author contributions**

T.L.: Project development, Data analysis, Manuscript writing. Manuscript editing. Z.B.: Project development, Data collection, Manuscript writing. S.B-M.: Manuscript editing. S.H.: Data management, Data collection, Manuscript editing. All the authors have read and approved the final manuscript.

**Funding**

No funds, grants, or other support were received.

**Availability of data and materials**

All the data that has been used is available in the code repository of the project in GitHub. Upon acceptance, we will publish all the code used in a GitHub repository. The final outcome can be reviewed at https://teddy4445.github.io/cancer-AKI-predictor.gui/predictor.html.

**Declarations**

**Ethics approval and consent to participate**

This research study was conducted retrospectively from data obtained for clinical purposes. We consulted extensively with the Bnai Zion Medical Center who provided our study approval numbered 82-0049-10.

**Consent for publication**

Informed consent was obtained from all individual participants included in the study.

**Competing interests**

The authors have no conflicts of interest to declare that are relevant to the content of this article.

**Author details**

1 Department of Cancer Biology, Cancer Institute, University College London, London, UK. 2 Department of Urology, Bnai Zion Medical Center, Haifa, Israel. 3 Department of Mathematics, Ariel University, Ariel, Israel.

**Received: 13 July 2021 Accepted: 10 May 2022**

**Published online: 16 May 2022**

**References**

1. Ferlay J, Colombet M, Soerjomataram I, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries and 25 major cancers in 2018. Eur J Cancer. 2018;103:356.
2. Chow WH, Devesa SS. Contemporary epidemiology of renal cell cancer. Cancer J. 2008;14:288–301.
3. EAU, url: https://uroweb.org/guidelines/ (13.12.2020)
4. Scosyrev E, Messing EM, Sylvester R, et al. Renal function after nephron-sparing surgery versus radical nephrectomy: results from EORTC randomized trial 30904. Eur Urol. 2014;66:372.
5. Bravi CA, Vertosick E, Benfante N, et al. Impact of acute kidney injury and its duration on long-term renal function after partial nephrectomy. Eur Urol. 2019;76(3):398–403.
6. Zabell J, Isharwal S, Dong W, et al. Acute kidney injury after partial nephrectomy of solitary kidneys: impact on long-term stability of renal function. J Urol. 2018;200(6):1295–301.
7. Patel HO, Pierozzo PM, Johnson MH, et al. Renal functional outcomes after surgery, ablation, and active surveillance of localized renal tumors: a systematic review and meta-analysis. Am Soc Nephrol. 2017;15:1555–9041.
8. Shah P, Kendall F, Khoszin S, et al. Artificial intelligence and machine learning in clinical development: a transnational perspective. NPJ Digit Med. 2019;2:69.
9. Boyko N, Sviridova T, Shakhovskaya N. Use of machine learning in the forecast of clinical consequences of cancer diseases. In: 7th Mediterranean conference on embedded computing (MECO). 2018; 1–6.
10. Weng SF, Reps J, Kai J, et al. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS ONE. 2017;12:e0174944.
11. Wu CC, Yeh WC, Hsu WD, et al. Prediction of fatty liver disease using machine learning algorithms. Comput Methods Programs Biomed. 2019;170:23–9. https://doi.org/10.1016/j.cmpb.2018.12.032.
12. Lee TH, Chen JJ, Cheng CT, Chang CH. Does artificial intelligence make clinical decision better? A review of artificial intelligence and machine learning in acute kidney injury prediction. Healthcare. 2021;9(12):1662. https://doi.org/10.3390/healthcare9121662.
13. Mistry NS, Koyner JL. Artificial intelligence in acute kidney injury: from static to dynamic models. Adv Chronic Kidney Dis. 2021;28(1):74–82. https://doi.org/10.1053/j.ackd.2021.03.002.
14. Kate RJ, Pearce N, Mazumdar D, et al. A continual prediction model for inpatient acute kidney injury. Comput Biol Med. 2020;116:103580. https://doi.org/10.1016/j.compbiomed.2019.103580.
15. Gameiro J, Branco T, Lopes JA. Artificial intelligence in acute kidney injury risk prediction. J Clin Med. 2020;9(3):678. https://doi.org/10.3390/jcm9030678.
16. Thottakkara P, Ozrazag-Baslanti T, Hupf BB, et al. Application of machine learning techniques to high-dimensional clinical data to forecast postoperative complications. PLoS ONE. 2016;11:e0155705.
17. Flechet M, Guza F, Schetz M, et al. AKI predictor, an online prognostic calculator for acute kidney injury in adult critically ill patients: development, validation and comparison to serum neutrophil gelatinase-associated lipocalin. Intensive Care Med. 2017;43(6):764–73.
18. Bahouth Z, Halachmi S, Getzler I, et al. Functional and oncological outcomes of open nephron-sparing surgery for complex renal masses. Urol Oncol Semin Orig Invest. 2015;33:427.
19. Bellomo R, Ronco C, Kellum JA, et al. Acute renal failure — definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. Crit Care. 2004;8:1–9.
20. Mehta RL, Kellum JA, Shah SV, et al. Acute Kidney Injury Network: report of an initiative to improve outcomes in acute kidney injury. Crit Care. 2007;11:R31.
21. Mangasarian OL, Setiono R, Wolberg WH. Pattern recognition via linear programming: theory and application to medical diagnosis. In: Large-scale numerical optimization. 1990; 22–30.
22. Breiman L. Random Forests. Mach Learn. 2001;45:5–32.
23. Guyon I, Cawley G. An improved Random Forests approach with application to the performance prediction challenge datasets. 2009; 1.
24. Liu R, Liu E, Yang J, et al. Optimizing the Hyper-parameters for SVM by Combining Evolution Strategies with a Grid Search. Intelligent Control and Automation. Lecture Notes in Control and Information Sciences, Springer, Berlin, Heidelberg. 2006; 344.
25. Wu G, Xu J. Optimized approach of feature selection based on information gain. In: International conference on computer science and mechanical automation (CSMA), Hangzhou. 2015; 157–161, https://doi.org/10.1109/C-SMA.2015.38.
26. Slaqm M, Shafique IB, Rahman K, et al. A Simple Study on Weight and Height of Students. Eur Sci J. 2017;13:63–71.
27. Coca SG, Yusuf B, Shlipak MG, et al. Long-term risk of mortality and other adverse outcomes after acute kidney injury: a systematic review and meta-analysis. Am J Kidney Dis. 2009;53(6):961–73.
28. Bucaloiu ID, Kirchner HL, Norfolk ER, et al. Increased risk of death and de novo chronic kidney disease following reversible acute kidney injury. Kidney Int. 2012;81(5):477–85.
29. Greenberg JH, Coca S, Parikh CR. Long-term risk of chronic kidney disease and mortality in children after acute kidney injury: a systematic review. BMC Nephrol. 2014;15:184.
30. Hidekazu T, Tsunenori K, Kazuhiro Y, et al. Lower incidence of postoperative acute kidney injury in robot-assisted partial nephrectomy than in open partial nephrectomy: a propensity score-matched study. J Endourol. 2020;34(7):754–62.
31. Cruz DN, de Cal M, Garzotto F, et al. Plasma neutrophil gelatinase-associated lipocalin is an early biomarker for acute kidney injury in an adult ICU population. Intensive Care Med. 2010;36:446–51.
32. Rank N, Pfähinger B, Kempter J, et al. Deep-learning-based real-time prediction of acute kidney injury outperforms human predictive performance. NPJ Digit Med. 2020;3:139.
33. Batchman LE, Foster CG. Missile system incorporating a targeting aid for man-in-the-loop missile controller. U.S. Patent 605307A. Issued February 25, 1997.
34. Mehta RL, Cercda J, Burdmann EA, et al. International Society of Nephrology’s 0 by 25 initiative for acute kidney injury (zero preventable deaths by 2025): a human rights case for nephrology. Lancet. 2015;385(9987):2616–43.

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