



# Prediction of Renal Cancer Recurrence using Artificial Intelligence: A Systematic Review



Vahid Mehrnoush<sup>1</sup>, Fatemeh Darsareh<sup>2</sup>, Waleed Shabana<sup>3</sup> and Walid Shahrour<sup>3\*</sup>

<sup>1</sup>General Surgery Section, Division of Clinical Studies, Northern Ontario School of Medicine University, Thunder Bay, Ontario, Canada

<sup>2</sup>Mother and Child Welfare Research Center, Hormozgan University of Medical Sciences, Bandar Abbas, Iran

<sup>3</sup>Urology Section, Division of Clinical Studies, Northern Ontario School of Medicine University, Thunder Bay, Ontario, Canada

**Submission:** January 14, 2025; **Published:** January 27, 2025

**Corresponding author:** Walid Shahrour, Urology Department, Northern Ontario School of Medicine University, Thunder Bay, Ontario, Canada.  
Email: Walid.Shahrour@tbh.net

## Abstract

**Objective:** This systematic review seeks to offer a summary of artificial intelligence (AI) methods for predicting renal cancer recurrence (RCR).

**Methods** This review involved searching for published studies dating from inception to January 2024 in Cochrane Central Register, PubMed, EMBASE, ProQuest, Scopus, and Google Scholar. Search keywords consist of “renal cancer” OR “renal cell carcinoma” OR “kidney cancer” AND “artificial intelligence” OR “machine learning.” We examined all studies that utilized machine learning techniques to predict RCR. The PROBAST instrument was utilized to assess the bias risk and applicability of every study included.

**Results** The search method produced 143 citations; 105 full-text articles were assessed for eligibility following the screening. This review included four studies. According to the risk of bias evaluation, three studies were determined to have a low risk of bias, while one had a low to moderate risk. Ten unique models were selected as algorithms for machine learning. Age, gender, smoking status, BMI, tumor stage, histological type, presence of necrosis, lymphovascular invasion, capsular invasion, Fuhrman grade, type of surgery (radical or partial nephrectomy), surgical approach (laparoscopic or open), and tumor size were selected as potential predictors to develop the model. Naïve Bayes, Decision Tree, Adaboost, and Random Forest were some of the top models for predicting RCR. Each study utilized the area under the curve (AUC) to present the accuracy. The AUC of machine learning models ranged from 0.740 to 0.877 across four studies.

**Conclusion** The outcomes of AI models for predicting RCR are encouraging. Scientists ought to keep exploring its extensive possibilities.

**Keywords:** Artificial Intelligence; Machine learning; Renal cancer; Systematic review

## Introduction

Renal cancer ranks as the sixth most prevalent cancer in men and the tenth most prevalent cancer in women globally [1]. Renal cancer exhibits a significant recurrence rate of 20% to 30% post-surgery, with over 40% of patients succumbing to the illness, underscoring the importance of consistent monitoring and management [2]. At present, it is challenging to differentiate between individuals who will experience a recurrence and those who will not. Artificial intelligence (AI) and machine learning in healthcare represent two emerging fields of research that have generated considerable interest lately. AI often utilizes artificial

neural networks, which function as statistical models inspired by and somewhat based on biological neural networks. They are capable of modeling and processing nonlinear connections between inputs and outputs simultaneously [3]. Exploratory research has evaluated the possible uses of AI and machine learning across different areas of urology, mainly focusing on the diagnosis and prognosis of genitourinary cancers; insights from other urological domains, including urolithiasis, kidney transplantation, urinary infections, and functional urology, have also been shared [4]. However, it is important to recognize that AI medical applications extend well beyond the realm of urology

[5,6]. This review seeks to offer a summary of AI methods for forecasting renal cancer recurrence (RCR). It enhances the current literature by outlining the AI methods used, the most suitable features, prevalent training and testing approaches, standard evaluation metrics, and how systems are implemented in clinical settings. The primary aim of this review was to recognize and outline the predictive factors of RCR through machine learning models and to assess the diagnostic accuracy of these models in predicting RCR.

## Methods

### Review questions

1. Which machine learning models were used to predict RCR?
2. What predictive factors of RCR are used to train the machine learning model?
3. Which machine learning models had a better performance in predicting RCR?
4. How accurate are machine learning models for predicting RCR?

### Eligibility criteria

All research utilizing machine learning-based methods to forecast RCR was included. Articles that were not in English and those irrelevant to the topic were excluded. Letters to the editor and critiques were similarly omitted. Ethical approval was unnecessary since we gathered and examined data from earlier published studies, in which the primary researchers secured informed consent. Data Sources and Searches The 2020 guidelines for Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [7] were utilized to present this study. This approach involved looking for published research from the beginning until January 2024. Databases comprised the Cochrane Central Register, PubMed, EMBASE (Through Ovid), ProQuest, Scopus, and Google Scholar. Terms used in the search included: ("renal cancer" OR "renal cell carcinoma" OR "kidney cancer" AND "artificial intelligence" OR "machine learning." Selected words and phrases came from a controlled vocabulary (MeSH, ENTREE, and others) along with a free-text search for every database. Furthermore, the reference lists from the identified articles were examined alongside hand-searching to guarantee that all documents were collected, which are integrated using Boolean "OR" and "AND" operators. A knowledgeable researcher explored every database. Following the removal of duplicates, two researchers independently examined the titles and abstracts, and subsequently the full texts of studies that may qualify, based on the established eligibility criteria. Consensus or consulting with a senior researcher was utilized to settle disputes. A PRISMA flowchart was utilized to record the procedure for selecting studies.

### Data Collection

Data were independently gathered by two researchers, while a third party resolved any disputes. The subsequent items were gathered: (1) demographic details (the nation of data collection, the source of data, and the prediction timeframe); (2) the category of predictive model, which is machine learning; and (3) the outcomes of the predictions (such as accuracy, sensitivity, specificity, and area under the recurrence curve) (AUROC); (4) the characteristics utilized to train the models.

### Data Synthesis

Given that only observational studies were part of the research, a narrative synthesis of the information was performed instead of a quantitative analysis.

### Risk of Bias Assessment

Two researchers evaluated the quality of all included studies separately and deliberated over any differences until they reached an agreement. PROBAST [8], comprising 20 signaling questions categorized into four areas (participants, predictors, outcome, and analysis), was utilized to evaluate the risk of bias and relevance of each study included. PROBAST helps evaluate the studied outcome by examining how it was established, its objectivity, the inclusion of any predictor data, its consistency across different individuals, the timing of its assessment, its independence from prior knowledge of predictor information, and its alignment with the review question.

### Results

The search approach produced 143 citations; subsequent to eliminating duplicates and completing the screening, 105 full-text articles were assessed for eligibility (PRISMA Flow Diagram, Figure 1). This review encompassed four studies. According to the assessment of bias risk, three studies had a low risk of bias, [9-11] while one exhibited a low to moderate risk of bias [12]. (Table 1). Table 2 presents the features of the studies that were included. The included studies took place in hospitals located in South Korea, France, and Canada. All study designs included were retrospective cohort studies. Ten unique models were selected as machine learning algorithms from the four studies (ranging from 2 to 8 per study). Out of the four studies, 13 distinct variables including: age, BMI, gender, smoking, pathological tumor stage, histologic type, margin status, tumor size, lymphovascular invasion, capsular invasion, Fuhrman grade, operation type (partial or total nephrectomy), operative methods (open or laparoscopic surgery) were chosen as potential predictors to develop the model (between 7 and 10 for each study). Naïve Bayes, Decision Tree, AdaBoost, and Random Forest were some of the top models for predicting renal cancer recurrence. The accuracy in all four studies was reported using metrics like specificity, sensitivity, and the area under the curve (AUC). In four studies, the AUC for machine learning models ranged from 0.740 to 0.877.

**Table 1:** PROBAST risk of bias/applicability assessment.

Study	Risk of Bias				Applicability			Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of Bias	Applicability
1	+	+	+	+	+	+	+	+	+
2	+	+/-	+	+	+	+/-	+	+/-	+/-
3	+	+	+	+	+	+	+	+	+
4	+	+	+	+	+	+	+	+	+

(+) indicate low risk of bias, (+/-) indicate low/moderate risk of bias, (-) indicate high risk of bias and (?) indicate unclear risk of bias.

**Table 2:** Summary of studies.

Study Number	Author	Country	Data Source	Prediction Time	Type of ML	Findings		
						Best ML performance	Accuracy	The features used to train the model
1	Kim et al. [9]	Republic of Korea	Dataset of 6849 patients with renal cancer	5-10 Years after surgery	SVM LR DT KNN NB EGBM AdaBoost GB	NB	AUC: 0.84 for models of recurrence within 5 years  AUC: 0.79 for models of recurrence within 10 years	Age BMI Gender Smoking Pathological tumor stage Histologic type Margin status Lymphovascular invasion Capsular invasion Fuhrman grade
2	Guo et al. [12]	Canada	Dataset of 697 patients with renal cancer	Not mentioned	DT Neural Network	DT	AUC: 0.877	Age Gender Tumor laterality Operation type (Partial/Total) nephrectomy Pathological tumor stage Margin status Fuhrman grade
3	Kim et al. [10]	Republic of Korea	Dataset of 2956 patients with renal cancer	5 Years after surgery	SVM LR DT KNN NB EGBM AdaBoost GB	AdaBoost	AUC: 0.740 Sensitivity:0.673 Specificity:0.807 Accuracy: 0.799	Operation type (Partial/Total) nephrectomy Operative method (Open/Laparoscopy) Pathological tumor stage Pathological node stage Histologic type Lymphovascular invasion Tumor size
4	Khene et al. [11]	France	Dataset of 4067 patients with renal cancer	57 Months	SVM RF EGBM	RF	AUC: 0.794	Pathological tumor stage Pathological node stage Fuhrman grade

AUC: Area under curve; ML: Machine learning; GB: Gradient boost; EGBM: Extreme gradient boosting models; DL: Deep learning; LR: Logistic regression; BT: Boosted tree; DT: Decision tree; KNN: K nearest neighbor; SP: support vector; SVM: kernel support vector machine; NB: naïve Bayes, RF: Random Forest.

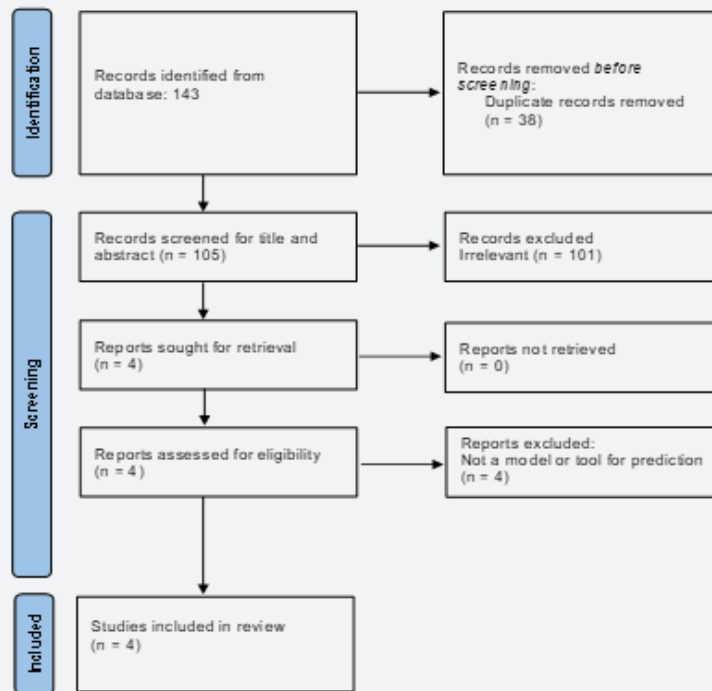


Figure 1: Flowchart of the study.

## Discussion

Estimating the total risk of cancer recurrence for patients can assist in customizing personal treatment strategies and improve patient guidance. Although only a limited number of studies employ AI models to forecast RCR, there exist multidisciplinary examinations of the various applications of AI in diagnosing and staging renal cancer. A recently published study demonstrated a wide range of AUCs, from 0.96 to 0.98, in validation cohort evaluations using different machine learning algorithms to differentiate low- from high-grade renal cancer, [13] although recent efforts have yielded AUCs between 0.91, [14] and 0.86, [15] down to 0.83 [16]. A few AI models have been evaluated for MRI classification of kidney cancer. Li et al. [17] found a favorable AUC of 0.842 for distinguishing low-grade from high-grade renal cancer using MRI, which showed a slight enhancement to 0.845 when incorporating patient history and imaging characteristics assigned by interpreting radiologists (clinicopathologic risk factors) into the model. To identify metastatic renal disease through AI imaging, Wen et al. found an AUC of 0.83 for forecasting the occurrence of synchronous distant metastases from renal cancer. Bai et al. incorporated clinicopathologic risk factors in their MRI radiomics machine learning model to identify synchronous distant metastases. They discovered AUCs of 0.854 and 0.814 derived from validation cohorts of MRI studies sourced from the same institution and outside institutions, respectively

[18,19].

Recent publications have discussed the use of AI in staging renal cancer to distinguish between early and advanced stages. Bhalla et al. utilized a 64-gene validation set to investigate gene expression in 523 samples of renal cancer. The aim was to pinpoint genes that are expressed differently in the early and late phases of renal cancer. The findings showed a peak accuracy of 72.64% and an AUC of 0.81 [20]. Regarding the application of AI to forecast overall survival in metastatic renal cancer, a retrospective analysis involving 322 Italian patients undergoing systemic therapy showed that the AI model reached (0.786 and 0.771) AUC and (0.675 and 0.558) specificity at a sensitivity of 0.90 for 3 and 5 years, respectively [21]. After our examination, we discovered only four studies regarding the use of AI for predicting RCR. Kim and colleagues employed machine learning algorithms to anticipate the probability of renal RCR five to ten years post-nephrectomy. Data from a Korean kidney cancer database were obtained, and eight distinct machine-learning models were created to estimate the probability of RCR. The maximum AUC in the five years following surgery was 0.836. In a decade, AUC had achieved 0.784. This approach showcases the capabilities of machine learning algorithms in helping clinical teams manage patients more effectively and deliver more personalized treatment strategies in patient care, especially after renal cancer surgery.

The final predictors of the optimized model comprised age, BMI, gender, smoking status, tumor stage, histologic type, presence of necrosis, lymphovascular invasion, capsular invasion, and Fuhrman grade.<sup>9</sup> Guo et al. conducted a study that compared a neural network with a boosted decision tree model to predict recurrence following curative treatment for kidney cancer. Data from 697 patients were available. The finalized predictors of the optimized model comprised age, gender, tumor laterality, radical or partial nephrectomy, pathological tumor stage, margin status, and Fuhrman grade, resulting in an AUC of 0.877.<sup>12</sup> A different research by Khene et al. revealed that the Random Forest model, achieving an AUC of 0.794, outperformed conventional statistical analysis in forecasting RCR. The most recent investigation by Kim et al. focused on creating a prediction model for late recurrence post-surgery in renal cancer patients, intended to serve as a clinical decision support system for the prompt identification of late recurrence. Late recurrence and non-recurrence were categorized using eight machine learning models. Among the eight models, the AdaBoost model displayed the best performance. The created algorithm demonstrated a sensitivity of 0.673, a specificity of 0.807, an accuracy of 0.799, and an AUC of 0.740 [10].

In summary, AI models are quickly advancing in various areas of renal cancer management and are currently achieving performance levels comparable to human professionals. Nevertheless, healthcare providers must initially establish a basic comprehension to standardize data sets, identify significant endpoints, and align interpretation. This requires collaboration across disciplines and the integration of AI courses into medical training. In the future, extensive and easily accessible databases containing high-quality data covering all facets of renal cancer care, from diagnosis to treatment, will be essential for facilitating external validation and ongoing training of AI models. Although they are effective in promoting innovation, competitions still require revision. Generally, the validation of the resulting algorithms has not been conducted separately from the algorithm creators. In a competitive landscape, the motivation for developers to deliberately or inadvertently introduce positive bias is arguably heightened. The absence of independent validation implies that the technical reproducibility of the suggested solutions must be confirmed. Additionally, competitions seldom include validation of the algorithms on extra international cohorts, prompting doubts about whether the solutions produced can genuinely address the fundamental clinical issue rather than being optimized for a particular competition format and dataset.<sup>22</sup> Despite significant attempts to perform a comprehensive and precise search in scientific databases, certain pertinent studies might have been missed due to resource constraints, the employment of particular search terms, or the restricted publication of certain articles. In addition, although the quality of the research was evaluated using established and validated methods, there remains a possibility of human error in rating or understanding the assessment standards.

These constraints could affect the final outcomes, regardless of attempts to minimize these biases.

## Conclusion

The outcomes of AI models designed to forecast RCR are encouraging. The identified prediction models displayed a low to moderate level of bias risk. Machine learning has the capacity to support various clinical areas within urology, and researchers ought to keep exploring its extensive possibilities. Ethical Approval Ethical approval was unnecessary since we gathered and examined data from earlier published studies, in which the primary researchers secured informed consent. Consent to Publish All authors gave informed consent for the publication of this study.

## Data Availability

Data will be provided upon request.

Declaration of Competing Interest The authors report no conflicts of interest in this study.

## Funding

None.

## References

1. Bukavina L, Bensalah K, Bray F, Carlo M, Challacombe B, et al. (2022) Epidemiology of Renal Cell Carcinoma: 2022 Update. *Eur Urol* 82(5): 529-542.
2. Chin AI, Lam JS, Figlin RA, Belldgrun AS (2006) Surveillance strategies for renal cell carcinoma patients following nephrectomy. *Rev Urol* 8(1): 1-7.
3. Van Booven DJ, Kuchakulla M, Pai R, Frech FS, Ramasahayam R, et al. (2021) A Systematic Review of Artificial Intelligence in Prostate Cancer. *Res Rep Urol* 13: 31-39.
4. Checcucci E, Autorino R, Cacciamani GE, Amparore D, De Cillis S, et al. (2020) Artificial intelligence and neural networks in urology: current clinical applications. *Minerva Urol Nefrol* 72(1): 49-57.
5. Boujarzadeh B, Ranjbar A, Banihashemi F, Mehrnough V, Darsareh F, et al. (2023) Machine learning approach to predict postpartum haemorrhage: a systematic review protocol. *BMJ Open* 13(1): e067661.
6. Ranjbar A, Ghamsari SR, Boujarzadeh B, Mehrnough V, Darsareh F (2023) Predicting risk of postpartum hemorrhage using machine learning approach: A systematic review. *Gynecology and Obstetrics Clinical Medicine* 3(3): 170-174.
7. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann T, et al. (2020) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 29: 372: n71.
8. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, et al. (2019) PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. *Ann Intern Med* 170(1): W1-W33.
9. Kim HM, Lee SJ, Park SJ, Choi IY, Hong SH (2021) Machine learning approach to predict the probability of recurrence of renal cell carcinoma after surgery: prediction model development study. *JMIR Med Inform* 9(3): 1-12.

10. Kim HM, Byun SS, Kim JK, Jeong CW, Kwak C, et al. (2022) Machine learning-based prediction model for late recurrence after surgery in patients with renal cell carcinoma. *BMC Med Inform Decis Mak* 22(1): 241.
11. Khene ZE, Bigot P, Doumerc N, Ouzaid I, Boissier R, et al. (2023) Application of Machine Learning Models to Predict Recurrence After Surgical Resection of Nonmetastatic Renal Cell Carcinoma. *Eur Urol Oncol* 6(3): 323-330.
12. Guo Y, Braga L, Kapoor A (2019) Machine learning to predict recurrence of localized renal cell carcinoma. *The Journal of Urology*.
13. Shu J, Wen D, Xi Y, Xia Y, Cai Z, et al. (2019) Clear cell renal cell carcinoma: Machine learning-based computed tomography radiomics analysis for the prediction of WHO/ISUP grade. *Eur J Radiol* 121: 108738.
14. Sun X, Liu L, Xu K, Li W, Huo Z, et al. (2019) Prediction of ISUP grading of clear cell renal cell carcinoma using support vector machine model based on CT images. *Medicine (Baltimore)* 98(14): e15022.
15. Bektas CT, Kocak B, Yardimci AH, Turkcanoglu MH, Yucetas U, et al. (2019) Clear Cell Renal Cell Carcinoma: Machine Learning-Based Quantitative Computed Tomography Texture Analysis for Prediction of Fuhrman Nuclear Grade. *Eur Radiol* 29(3): 1153-1163.
16. Nazari M, Shiri I, Hajianfar G, Oveisi N, Abdollahi H, et al. (2020) Noninvasive Fuhrman grading of clear cell renal cell carcinoma using computed tomography radiomic features and machine learning. *Radiol Med* 125(8): 754-762.
17. Li Q, Liu YJ, Dong D, Bai X, Huang QB, et al. (2020) Multiparametric MRI Radiomic Model for Preoperative Predicting WHO/ISUP Nuclear Grade of Clear Cell Renal Cell Carcinoma. *J Magn Reson Imaging* 52(5): 1557-1566.
18. Wen R, Huang J, Gao RZ, Wan D, Qin H, et al. (2021) CT Radiomics for the Prediction of Synchronous Distant Metastasis in Clear Cell Renal Cell Carcinoma. *J Comput Assist Tomogr* 45(5): 696-703.
19. Bai X, Huang Q, Zuo P, Zhang X, Yuan J, et al. (2021) MRI radiomics-based nomogram for individualised prediction of synchronous distant metastasis in patients with clear cell renal cell carcinoma. *Eur Radiol* 31(2): 1029-1042.
20. Bhalla S, Chaudhary K, Kumar R, Sehgal M, Kaur H, et al. (2017) Gene expression-based biomarkers for discriminating early and late stage of clear cell renal cancer. *Sci Rep* 7: 44997.
21. Barkan E, Porta C, Rabinovici-Cohen S, Tibollo V, Quaglini S, et al. (2023) Artificial intelligence-based prediction of overall survival in metastatic renal cell carcinoma. *Front Oncol* 13: 1021684.
22. Maier-Hein L, Eisenmann M, Reinke A, Onogur S, Stankovic M, et al. (2018) Why rankings of biomedical image analysis competitions should be interpreted with care. *Nat Commun* 9(1): 5217.



This work is licensed under Creative Commons Attribution 4.0 License  
DOI: [10.19080/CTOIJ.2025.28.556233](https://doi.org/10.19080/CTOIJ.2025.28.556233)

**Your next submission with Juniper Publishers  
will reach you the below assets**

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats  
( Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

**Track the below URL for one-step submission**  
<https://juniperpublishers.com/online-submission.php>