ORIGINAL RESEARCH

# The Use of Artificial Intelligence Machine Learning Models to Predict Stone-Free Status After Percutaneous Nephrolithotomy: A Meta-Analysis

[1]Rajiv H. Kalbit, MD, FPUA; [1]Enrique Ian S. Lorenzo, MD, FPUA and [2]Karl Marvin M. Tan, MD, FPUA

Department of Surgery, [1]Jose R. Reyes Memorial Medical Center and [2]Veterans Memorial Medical Center

**Objective**: This meta-analysis aimed to evaluate the diagnostic capability of machine learning (ML) models in predicting stone-free status following percutaneous nephrolithotomy (PCNL).
**Methods**: A comprehensive literature search was conducted across MEDLINE, Embase, Scopus, Cochrane, Google Scholar and supplementary databases was undertaken until June 2023. Inclusion criteria were English publications assessing the sensitivity and specificity of ML in predicting post-PCNL stone-free status. Studies on non-human subjects or with incomplete data sets were excluded. Quality assessment utilized the Cochrane Risk of Bias Tool. Pooled sensitivity, specificity, and other diagnostic metrics were calculated using Meta-Disc 1.4 software.
**Results**: Of the 65 initial articles, 5 met the inclusion criteria, representing a total of 1,773 participants. The accuracy of ML models ranged from 44% to 94.8%. The pooled sensitivity and specificity were 0.60 (95% CI [0.57, 0.63]) and 0.87 (95% CI [0.84, 0.89]), respectively. The pooled positive likelihood ratio was 4.69 (95% CI [3.82, 5.77]) and the negative likelihood ratio was 0.45 (95% CI [0.41, 0.48]). The diagnostic odds ratio was 10.93 (95% CI [8.35, 14.33]). The area under the curve (AUC) stood at 0.9372, signifying an excellent diagnostic performance.
**Conclusion**: Machine learning models demonstrate significant potential in accurately predicting stone-free status post-PCNL. However, the small number of included studies, retrospective designs, and heterogeneity in ML approaches limit generalizability. Standardized definitions, larger multicenter datasets, and prospective validation are required before routine clinical adoption.

**Key words**: Machine learning, percutaneous nephrolithotomy, stone-free status, diagnostic capability, meta-analysis.

## Introduction

Urolithiasis remains one of the most prevalent urologic disorders worldwide, with an estimated lifetime incidence ranging from 1.7% to 14.8%.[1,2] It is more common among males and continues to rise globally, imposing a significant socioeconomic and quality-of-life burden.[3,4] Technological innovations have transformed the management of renal calculi, shifting from open surgery toward minimally invasive procedures such as extracorporeal shock wave lithotripsy (ESWL), retrograde intrarenal surgery (RIRS) and percutaneous nephrolithotomy (PCNL).[1,3,5] PCNL remains the standard of care for renal stones ≥20 mm, offering high clearance rates while preserving renal function.[4]

Predicting postoperative stone-free status (SFS) after PCNL is crucial for clinical decision-

making, patient counseling, and optimizing surgical outcomes. Several conventional scoring systems, such as Guy's Stone Score, the Clinical Research Office of the Endourological Society (CROES) nomogram, and the S.T.O.N.E. nephrolithometry score, have been developed to estimate stone-free outcomes.[5-7] However, these tools are limited by their reliance on a fixed set of variables, subjective grading, and an assumption of linear relationships among predictors. They may fail to capture the complex, nonlinear interactions between patient demographics, stone characteristics and intraoperative parameters that influence surgical success.[5-7]

Artificial Intelligence (AI) refers to any computer technology that examines intricate patterns and solves complex problems by imitating human cognitive functions, such as thinking, learning and problem solving. Machine learning (ML) is a subtype of AI that analyzes and understands complex patterns using data-driven dynamic algorithms and semi-automatically improves its analysis.[4-6] Training data sets are used to create algorithms for rapid identification of complex patterns and relationships of future data.[4,6] Deep learning (DL), a variant of ML artificial neural network (ANN), is patterned on the function and structure of the human brain, wherein artificial neurons are arranged and are interconnected in complex architectural layers.[2,5] It uses computer vision in conjunction with DL algorithms to examine medical images. It provides precise and reliable anatomical models for operational support, and predicts outcomes and success rates of treatment when used alongside computed tomography (CT) images. It aids medical practitioners in decision making, thereby decreasing iatrogenic errors.[2] In recent years, AI has been in the forefront of medical diagnostics and analytics research; and image-based diagnostic systems have been developed for many medical specialties.[1]

In the field of urology, there has been increasing use of ML in predicting the outcome of renal calculi following ESWL and PCNL. However, there are only a few studies looking into the sensitivity and specificity in the application of urologic condition. Thus, this study aimed to determine the diagnostic characteristic of ML in predicting stone-free status after PCNL.

**Methods**

A comprehensive literature search was conducted across MEDLINE, Embase, Scopus, Cochrane, Google Scholar and supplementary databases was undertaken until June 2023. To enhance comprehensiveness, additional databases and gray literature sources were explored. The MeSH terms used for the search included: [[artificial intelligence] OR [machine learning] OR AI OR [deep learning] OR [neural network]] AND [[urinary calculi] OR [kidney calculi] OR [renal calculi] OR urolithiasis OR renal OR ureteric OR stones] AND [percutaneous nephrolithotomy] OR nephrolithotomy. Supplementary studies were identified through manual scanning of reference lists from relevant articles.

Inclusion Criteria:
- Published in English.
- Assessed the sensitivity and specificity of ML in predicting stone-free status after PCNL.

Exclusion criteria encompassed:
- Studies on non-human subjects.
- Studies with incomplete data sets.

All the identified articles were screened for eligibility by three independent researchers (RHK, EIL, KMT). The PRISMA flowchart was adopted to map out the study selection process, ensuring transparency and replicability. In cases of disagreement between reviewers, consensus was reached through a majority vote.

A standardized form was used for data extraction. This form included key variables such as study design, sample size, ML model used, sensitivity, specificity, positive and negative likelihood ratios, diagnostic odds ratio, and any other relevant metrics. When data were not explicitly mentioned, they were derived and computed based on available information.

The quality and risk of bias for each study were assessed using the Cochrane Risk of Bias Tool. Elements of assessment included selection bias, performance bias, detection bias, attrition bias and reporting bias. Studies were classified as having low, unclear, or high risk of bias.

Heterogeneity among the studies was evaluated using the $I^2$ statistics. Studies with an $I^2$ value

over 60% were considered to have substantial heterogeneity, and in such cases, the DerSimonian Laird's random effects model was applied. A sensitivity analysis was also undertaken, and the rationale for excluding specific studies, such as that by Geraghty et al., was explicitly documented.

Several ML models were detailed and analyzed:

1. **Artificial Neural Network (ANN)**: This emulates interconnected neural synapses and networks, akin to the human brain. Its mechanism and parameters, including the number of neurons in the input layer and output calculation method, were outlined.
2. **Support Vector Machine (SVM)**: A widely applied supervised learning model for regression and classification, its cross-validation process and training phase were highlighted.
3. **Other Models**: Additional models used in the studies, like logistic regression, sequential forward selection (SFS), Fisher discriminant analysis (FDA), quadratic discriminant analysis (QDA), K-nearest NEIGHBORS (KNN), multilayer perceptron neural network (MLPNN), and random forest (RF) were described with specific references to their core mechanics, algorithms and applications in the context of the research.

Meta-Disc 1.4 software was employed for statistical analysis. Pooled sensitivity, specificity, positive and negative likelihood ratio, and diagnostic odds ratio were calculated, and a random effects model was employed in the presence of high heterogeneity.

## Results

Initial literature search found 63 articles screened for this study. Ten additional articles were identified through cross-referencing and review of bibliography of the included articles. Eight duplicate articles were identified and removed, leaving 65 articles. A review of the abstracts of each articles was done, and 59 articles failed to meet the inclusion criteria. Of the 6 articles that underwent full text review, 1 article was excluded due to incomplete data for analysis. Overall, there were a total of five articles that were included for analysis.[7-11] The PRISMA flow diagram of literature search is shown in Figure 1. To ensure the quality of each included article, a risk of bias analysis of all the included articles was done. All included articles had low risk of bias (Figure 2).

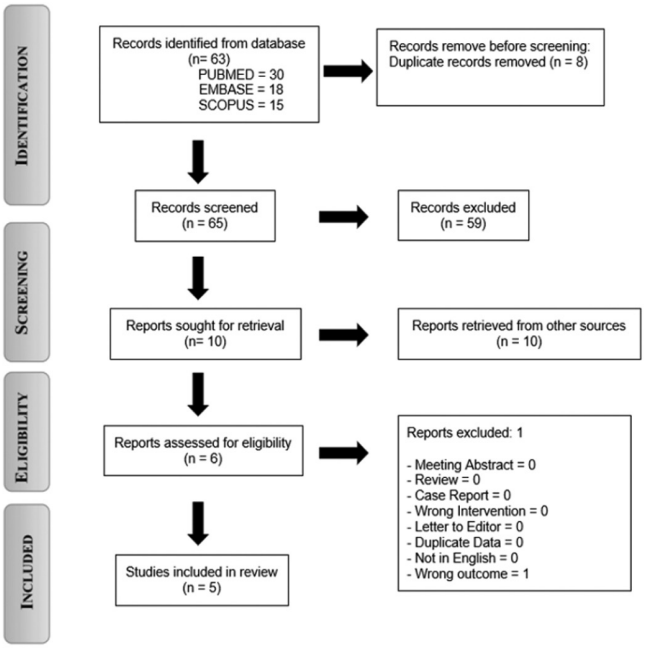


**Figure 1.** PRISMA flow diagram of literature search.

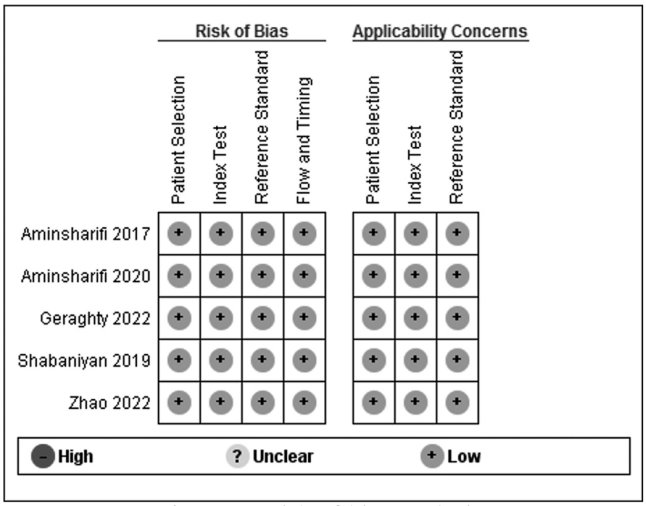


**Figure 2.** Risk of bias analysis.

The details of each included study are summarized in Table 1. There were a pooled total of 1,773 participants in this metaanalysis, with mean age between 40 and 50 years old. Various

ML algorithms were used, including artificial neural networks (ANN), support vector machines (SVM), random forest (RF), extreme gradient boosting (XGBoost), and logistic regression models. Validation approaches varied, with most studies performing internal cross-validation. Only the study by Aminsharifi et al. (2020) conducted external validation using an independent dataset to compare its ML model against Guy's Stone Score and the CROES nomogram. Definitions of stone-free status (SFS) varied considerably across studies. Some defined SFS using non-contrast CT scans, whereas others relied on plain KUB radiography or ultrasound, with follow-up intervals ranging from immediate postoperative imaging to six weeks after surgery. This inconsistency in imaging modality and timing likely contributed to the observed heterogeneity in pooled diagnostic estimates. Table 2 shows the summary of diagnostic testing of the included studies. The accuracy of the ML model ranges from 44% to 94.8%. The sensitivity varies from 0% to 100%, while specificity varies from 21% to 100%. The positive predictive value (PPV) and negative predictive value (NPV) ranges from 0% to 97.3% and 31.2% to 100%, respectively. The false positive rate (FPR) varies from 0% to 78.6%. Lastly, the area under the curve (AUC) ranges from 0.50 to 0.915.

**Table 1.** Summary of included studies.

| Study | ML models | N | Mean age (years) | Mean stone size (mm) | No. stone free status |
|---|---|---|---|---|---|
| Artificial neural network system to predict the postoperative outcome of percutaneous nephrolithotomy A. Aminsharifi et al. (2017) | ANN | 254 | 46.64 ± 12.16 | 21.587 ± 9.09 | 194 |
| Predicting the postoperative outcome of percutaneous nephrolithotomy with machine learning system: Software validation and comparative analysis with Guy's stone score and the CROES nomogram A. Aminsharifi et al. (2020) | SVM | 146 | 49.3 ± 12.6 | 451.2 ± 427.8 | 106 |
| Use of internally validated and deep learning models to predict outcomes of percutaneous nephrolithotomy using data from the BAUS PCNL audit R. Geraghty et al. (2022) | 1. LR<br>2 .RF<br>3. XGBoost<br>4. BGLM<br>5. Partitioning<br>6. Neural networks | 778 | 56.5 ± 19.4 | Not stated | 535 |
| An artificial intelligence-based clinical decision support system for large kidney stone treatment T. Shabaniyan et al. (2019) | 1. SFS<br>2. FDA<br>a. QDA<br>b. KNN<br>c. MLP<br>d. SVM | 254 | 46.6 ± 12.2 | 21.587 ± 9.09 | 194 |
| Predicting the stone-free status of percutaneous nephrolithotomy with the machine learning system: comparative analysis with Guy's stone score and the S.T.O.N.E score system H. Zhao et al. (2022) | 1. Lasso logistic<br>2. RF<br>3. SVM<br>4. Naïve Bayes | 222 | 54.81 ± 13.31 | Not stated | 111 |
| Legend: machine learning (ML), artificial neural network (ANN), support vector model (SVM), logistic regression (LR), random forest (RF), extreme gradient boosting (XGBoost), Bayesian generalized linear model (BGLM), partitioning, sequential forward selection (SFS), Fisher discriminant approach (FDA), Quadratic discriminant analysis (QDA), K-nearest NEIGHBORS (KNN), Multilayer perception neural network (MLP) | | | | | |

**Table 2.** Summary of diagnostic testing of included studies.

| Study/ ML model | Accuracy | Sensitivity | Specificity | PPV | NPV | FPR | AUC |
|---|---|---|---|---|---|---|---|
| Aminsharifi 2017 | | | | | | | |
| ANN | 82.8% | 83% | 81% | 83% | 81% | 19% | 0.861 |
| Aminsharifi 2020 | | | | | | | |
| SVM | 91.8% | 92% | 88.9% | 95% | 83.3% | 11% | 0.915 |
| Geraghty 2022 | | | | | | | |
| RF | 70% | 0% | 100% | 0% | 31.2% | 0% | 0.69 |
| Partitioning | 70% | 0% | 100% | 0% | 31.2% | 0% | 0.55 |
| XGBoost | 65% | 20% | 87% | 77.2% | 33.1% | 13.2% | 0.70 |
| LR | 62% | 30% | 78% | 75% | 33.6% | 21.8% | 0.61 |
| Neural network | 70% | 0% | 100% | 0% | 31.2% | 0% | 0.50 |
| BGLM | 69% | 30% | 88% | 84.6% | 36.4% | 11.9% | 0.67 |
| Deep neural network (single otcome) | 59% | 43% | 78% | 81,1% | 38.3% | 21.8% | 0.62 |
| Deep neural network (multiple outcome) | 44% | 92% | 21% | 71.9% | 54.4% | 78.6% | 0.60 |
| Shabaniyan 2019 | | | | | | | |
| QDA_SFS | 70.3% | 71.1% | 69.5% | 88.3% | 42.7% | 30% | |
| QDA_SFSFDA | 78% | 72.3% | 84.7% | 93.9% | 48.6% | 15% | |
| KNN_SFS | 73.5% | 80.7% | 65.3% | 88.3% | 51.1% | 35% | |
| KNN_SFSFDA | 79.4% | 85.6% | 72.3% | 90.9% | 60.8% | 28.3% | Not stated |
| MLP_SFS | 63.9% | 71.1% | 55.6% | 83.8% | 37.3% | 45% | |
| MLP_SFSFDA | 75.5% | 81.9% | 68.1% | 89.3% | 53.8% | 31.7% | |
| SVM_SFS | 92.3% | 92.7% | 91.6% | 97.3% | 79.5% | 8.3% | |
| SVM_SFSFDA | 94.8% | 100% | 88.9% | 96.7% | 100% | 11.7% | |
| Zhao 2022 | | | | | | | |
| Lasso logistic | 81.81% | 75.76% | 87.77% | 86.1% | 78.36% | 12.61% | 0.879 |
| RF | 80.3% | 75.76% | 84.85% | 83.34% | 77.78% | 15.32% | 0.803 |
| SVM | 81.82% | 75.76% | 87.88% | 86.21% | 78.38% | 11.71% | 0.818 |
| Naïve Bayes | 80.3% | 83.33% | 77.78% | 78.95% | 82.35% | 22.52% | 0.803 |
| Rami AlAzab 2023 | | | | | | | 0.761 |
| Predicting the Stone-Free | 0.74 | | | | | | 0.769 |
| Status of Percutaneous | 0.72 | Not stated | Not stated | Not | Not | Not | 0.751 |
| Nephrolithotomy with the | 0.74 | | | stated | stated | stated | 0.666 |
| Machine Learning System | | | | | | | 0.71 |
| Zeeshan Hameed 2021 | | | | | | | |
| With MRMR treatment extracting top 3 features | | 96% | | 60 % | | | LDA 0.81 |
| With MRMR treatment extracting top 5 features | 81% | 84% | Not Stated | 60 % | Not Stated | Not Stated | MRMR 0.64 |
| With MRMR treatment extracting top 10 features | | 68% | | 60 % | | | |

Legend: machine learning (ML), positive predictive value (PPV), negative predictive value (NPV), false positive rate (FPR), area under the curve (AUC), artificial neural network (ANN), support vector model (SVM), logistic regression (LR), random forest (RF), extreme gradient boosting (XGBoost), Bayesian generalized linear model (BGLM), partitioning, sequential forward selection (SFS), Fisher discriminant approach (FDA), Quadratic discriminant analysis (QDA), K-nearest NEIGHBORS (KNN), Multilayer perception neural network (MLP)

Overall, ML is an excellent diagnostic tool capable of predicting stone-free status after PCNL, with a pooled sensitivity of 0.60, 95% CI [0.57, 0.63], pooled specificity of 0.87, 95% CI [0.84, 0.89], pooled positive likelihood ratio of 4.69, 95% CI [3.82, 5.77], pooled negative likelihood ratio of 0.45, 95% CI [0.41, 0.48], and pooled diagnostic odds ratio of 10.93, 95% CI [8.35, 14.33] (Figures 3-7). The SROC curve for the diagnostic performance of ML in predicting stone-free status after PCNL is shown in Figure 8. The size of the circle represents the sample size of the study and its location represent the different sensitivity and specificity of ML in the included studies. The area under the curve (AUC) is 0.9372.

Moderate heterogeneity was observed across studies ($I^2 > 60\%$), likely due to differences in datasets, model architectures, and SFS definitions.

Sensitivity analysis excluding the outlier study by Geraghty et al., which included multiple models with very low sensitivities, increased the pooled sensitivity to 0.86 (95% CI: 0.83–0.89) without significantly altering specificity. (Figure 9)

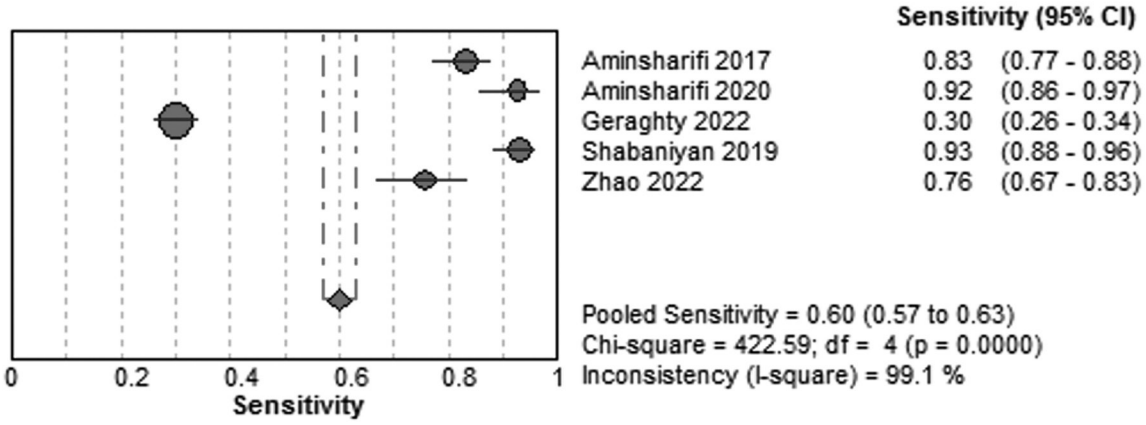


**Figure 3.** Pooled sensitivity of ML in predicting stone-free status after PCNL.

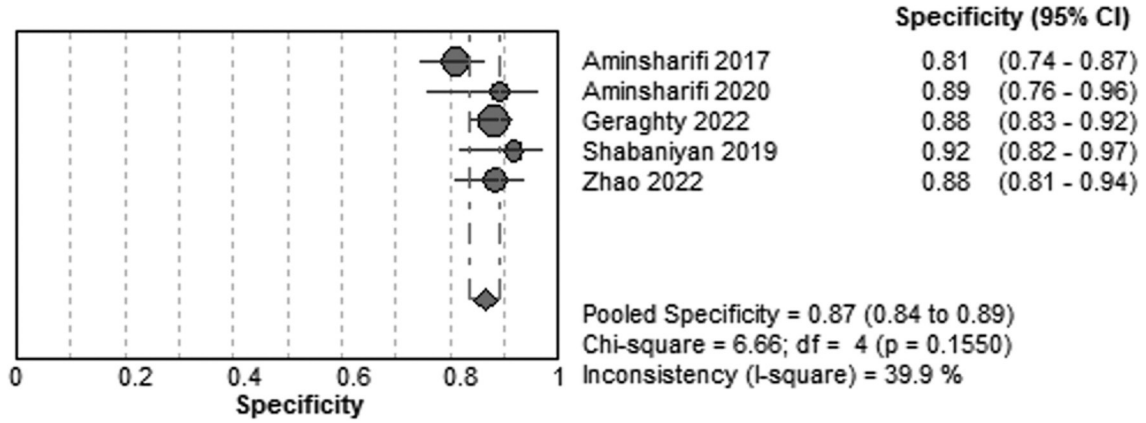


**Figure 4.** Pooled specificity of ML in predicting stone-free status after PCNL.

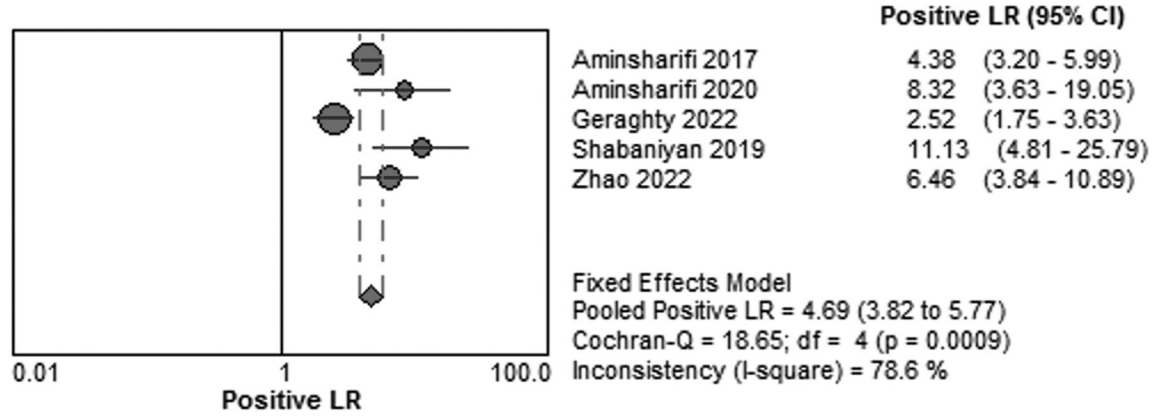


**Figure 5.** Pooled positive likelihood ratio of ML in predicting stone-free status after PCNL.

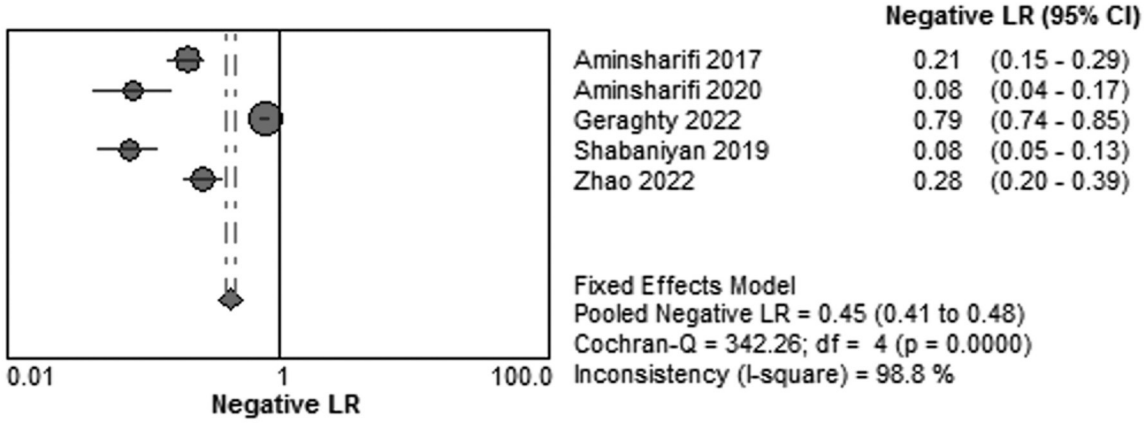


**Figure 6.** Pooled negative likelihood ratio of ML in predicting stone-free status after PCNL.

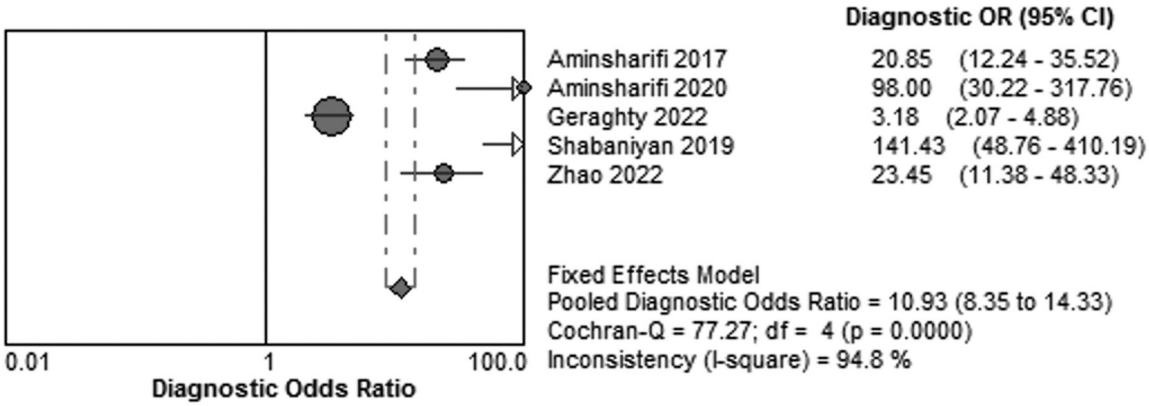


**Figure 7.** Pooled diagnostic ratio of ML in predicting stone-free status after PCNL.

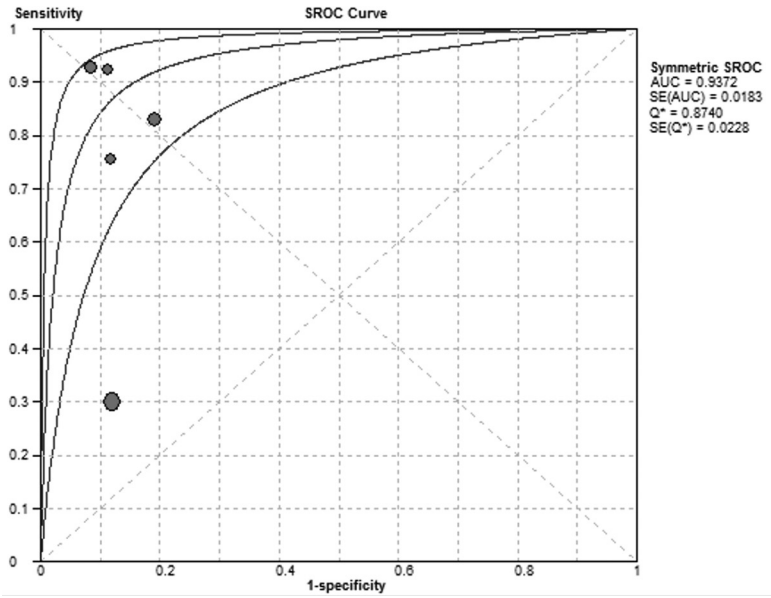


**Figure 8.** SROC of ML in predicting stone-free status after PCNL.

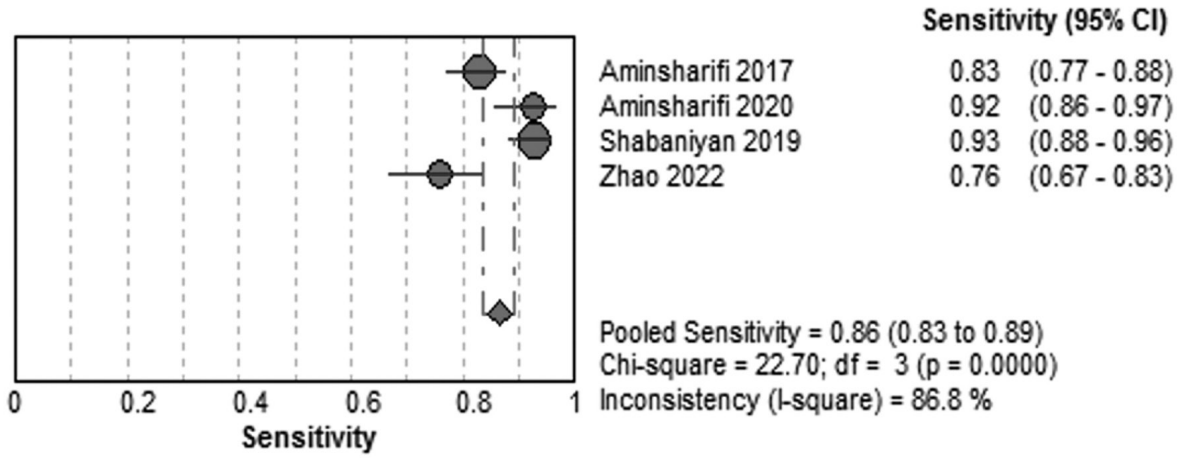


**Figure 9.** Sensitivity analysis of the sensitivity of ML models.

## Discussion

This meta-analysis demonstrates that machine learning (ML) models show considerable potential in predicting stone-free status (SFS) following percutaneous nephrolithotomy (PCNL). The pooled analysis revealed a sensitivity of 0.60 and a specificity of 0.87, with an area under the SROC curve (AUC) of 0.94, indicating excellent overall discriminative ability. The diagnostic capabilities of ML models in this regard were further validated by a positive likelihood ratio of 4.69 and a negative likelihood ratio of 0.45. These results suggest that ML algorithms are highly effective at identifying patients who will achieve stone-free status after PCNL. Furthermore, the diagnostic odds ratio stood at 10.93, hinting that patients deemed stone-free by ML models are nearly 11 times more likely to achieve that status than those predicted otherwise.

These results have important clinical implications. Accurate preoperative prediction of SFS can aid in patient counseling, individualized treatment planning, and efficient use of operative resources. Machine learning (ML)-based models can provide a more flexible and nonlinear analysis than traditional methods by taking into account a wide range of factors, such as patient demographics, stone features, operating parameters and imaging data. On the other hand, traditional scoring tools like the Guy's Stone Score, CROES nomogram, and S.T.O.N.E. nephrolithometry score are limited since they depend on static, linearly weighted data and subjective interpretation. Several studies, such as Aminsharifi et al. (2020), have shown that ML models can outperform these conventional scores in predictive accuracy when externally validated. The wide variability observed in model accuracy (44–94.8%) and sensitivity (0–100%) across included studies reflects differences in algorithm architecture, dataset size, variable selection and validation strategy. The heterogeneity is further compounded by inconsistencies in defining "stone-free status," with some studies relying on KUB radiography or ultrasound, while others used non-contrast CT scans at varying follow-up intervals from immediate postoperative to six weeks. Moreover, most studies utilized internal cross-validation, with only one study (Aminsharifi et al., 2020) performing external validation. This limits the generalizability of the models and may lead to overly optimistic estimates of diagnostic performance.

Although ML models demonstrate significant potential, numerous practical obstacles must be overcome before they can be widely adopted in the clinical setting. First, most published models are trained on single-center, retrospective datasets with limited external applicability. Second, the "black box" problem, which makes it hard to understand how models make decisions, is still a problem since doctors may not want to use models that are hard to

understand. Third, for ML technologies to function with clinical workflows, they need to work with imaging systems, have consistent data inputs, and follow data privacy laws like HIPAA and GDPR. Variation in imaging modality, scanner parameters and feature extraction protocols can further affect model performance and reproducibility.

Despite these limitations, ML represents a transformative step toward personalized, data-driven urolithiasis management. The ability to process complex multidimensional data offers the potential for improved prediction of surgical success and tailored postoperative care. Future research should focus on prospective, multicenter studies using standardized definitions of SFS, transparent model architectures, and external validation across diverse populations. Collaborative registries that integrate clinical, radiologic, and intraoperative data may enhance generalizability and accelerate clinical translation.

## Conclusion

Machine learning (ML) models demonstrate promising diagnostic capability in predicting stone-free status (SFS) following percutaneous nephrolithotomy (PCNL), with high specificity and excellent overall discriminative performance. Compared with conventional scoring systems, ML offers greater flexibility in analyzing complex clinical and imaging variables. Nevertheless, the current evidence remains limited by the small number of studies, heterogeneity in definitions of SFS, and reliance on retrospective, single-center data with predominantly internal validation. To guarantee clinical reliability and reproducibility, future research should prioritize transparent model reporting, standardized imaging-based definitions, and external multicenter validation using large, prospective datasets.

Although ML is not yet capable of replacing clinical expertise or established scoring tools, it is a promising complementary tool for individualized management, patient counseling, and preoperative planning. To transform these algorithms into practical, real-world decision-support systems in endourology, it will be important to maintain collaboration among urologists, data scientists and engineers.

## References


1. Onal E and Tekgul H. Assessing kidney stone composition using smartphone microscopy and deep neural networks. BJUI Compass 2022; 3: 310-5.
2. Caglayan A, Horsanali M, Kocadurdu K, Ismailoglu E et al. Deep learning model-assisted detection of kidney stones on computed tomography. Int J Brazilian Urol 2022; 48: 830-9.
3. Veser J, Jahrreiss V and Seitz C. Innovations in urolithiasis management. Curr Opin Urol 2021; 31(2): 130-4.
4. Rice P, Pugh M, Geraghty R, Hameed B et al. Machine learning models for predicting stone-free status after shockwave lithotripsy: a systematic review and meta-analysis. Urol 2021; 00: 1-7.
5. Checcucci E, De Cillis S, Granato S, Chang P et al. Applications of neural network in urology: a systematic review. Curr Opin Urol 2020; 30(6): 788-807.
6. Liu H, Wang X, Tang K, Peng E et al. Machine learning-assisted decision-support models to better predict patients with calculous pyonephrosis. Transl Androl Urol 2021; 10(2): 710-23.
7. Aminsharifi A, Irani D, Pooyesh S, Parvin H et al. Artificial neural network system to predict the postoperative outcome of percutaneous nephrolithotomy. J Endo 2017; 31 (5): 461-7.
8. Aminsharifi A, Irani D, Tayebi S, Kafash T et al. Predicting the postoperative outcome of percutaneous nephrolithotomy with machine learning system: software validation and comparative analysis with Guy's stone score and the CROES nomogram. J Endo 2020; 34 (6): 1-24.
9. Geraghty R, Finch W, Fowler S, Sriprasad S et al. Use of internally validated machine and deep learning models to predict outcomes of percutaneous nephrolithotomy using data from the BAUS PCNL audit. MedRxiv 2022; doi: 10.1101/2022.06.16.22276481.
10. Shabaniyan T, Parsaei H, Aminsharifi A, Movahedi M et al. An artificial intelligence-based clinical decision support system for large kidney stone treatment. Australasian Phys Eng Sci Med 2019; 42(3): 771-9.
11. Zhao H, Li W, Li J, Li L et al. Predicting the stone-free status of percutaneous nephrolithotomy with the machine learning system: comparative analysis with Guy's stone score and the S.T.O.N.E Score System. Front Pharmacol 2022; 9: 880291.
12. Alghafees M, Rab S, Aljurayyad A, Alotaibi T et al. A retrospective cohort study on the use of machine learning to predict stone-free status following percutaneous nephrolithotomy: an experience from Saudi Arabia. Ann Med Surg 2022; 84: 1-5.
13. Ganesan V and Pearle M. Artificial intelligence in stone disease. Curr Opin Urol 2021; 31 (4): 391-6.
14. Yang B, Veneziano D and Somani B. Artificial intelligence in the diagnosis, treatment and prevention of urinary stones. Curr Opin Urol 2020; 30 (6): 782-7.

15. AlAzab R, Ghammaz O, Ardah N, Al-Bzour A, Zeidat L, Mawali Z, Ahmed YB, Alguzo T, Al-Alwani A & Samara M. Predicting the Stone-Free Status of Percutaneous Nephrolithotomy with the Machine Learning System. 2023 https://doi.org/10.21203/rs.3.rs-2550836/v1

16. Zeeshan Hameed BM, Somani BTPR, Raza SZ, Paul R, Naik N, Singh H, Shah M & Reddy S. Application of Artificial Intelligence-based classifiers to predict the outcome measures and stone-free status following percutaneous nephrolithotomy for staghorn calculi: Cross-validation of data and estimation of accuracy. Eur Urol 2021; 79: S1375. https://doi.org/10.1016/s0302-2838(21)01348-8