

Original Paper

# A Machine Learning Approach to Predict the Outcome of Urinary Calculi Treatment Using Shock Wave Lithotripsy: Model Development and Validation Study

Reihaneh Moghisi<sup>1\*</sup>, MSci; Christo El Morr<sup>2\*</sup>, PhD; Kenneth T Pace<sup>3,4</sup>, MD, PhD; Mohammad Hajiha<sup>5</sup>, MD; Jimmy Huang<sup>1</sup>, PhD

<sup>1</sup>School of Information Technology, York University, Toronto, ON, Canada

<sup>2</sup>School of Health Policy and Management, York University, Toronto, ON, Canada

<sup>3</sup>Division of Urology, St Michael's Hospital, Toronto, ON, Canada

<sup>4</sup>Department of Surgery, University of Toronto, Toronto, ON, Canada

<sup>5</sup>Department of Urology, Loma Linda University Health, Loma Linda, CA, United States

\*these authors contributed equally

**Corresponding Author:**

Reihaneh Moghisi, MSci

School of Information Technology

York University

4700 Keele St

Toronto, ON, M3J 1P3

Canada

Phone: 1 416 736 2100

Email: [rmoghisi88@gmail.com](mailto:rmoghisi88@gmail.com)

## Abstract

**Background:** Shock wave lithotripsy (SWL), ureteroscopy, and percutaneous nephrolithotomy are established treatments for renal stones. Historically, SWL has been a predominant and commonly used procedure for treating upper tract renal stones smaller than 20 mm in diameter due to its noninvasive nature. However, the reported failure rate of SWL after one treatment session ranges from 30% to 89%. The failure rate can be reduced by identifying candidates likely to benefit from SWL and manage patients who are likely to fail SWL with other treatment modalities. This would enhance and optimize treatment results for SWL candidates.

**Objective:** We proposed to develop a machine learning model that can predict SWL outcomes to assist practitioners in the decision-making process when considering patients for stone treatment.

**Methods:** A data set including 58,349 SWL procedures performed during 31,569 patient visits for SWL to a single hospital between 1990 and 2016 was used to construct and validate the predictive model. The AdaBoost algorithm was applied to a data set with 17 predictive attributes related to patient demographics and stone characteristics, with success or failure as an outcome. The AdaBoost algorithm was also applied to a training data set. The generated model's performance was compared to that of 5 other machine learning algorithms, namely C4.5 decision tree, naïve Bayes, Bayesian network, K-nearest neighbors, and multilayer perceptron.

**Results:** The developed model was validated with a testing data set and performed significantly better than the models generated by the other 5 predictive algorithms. The sensitivity and specificity of the model were 0.875 and 0.653, respectively, while its positive predictive value was 0.7159 and negative predictive value was 0.839. The C-statistics of the receiver operating characteristic (ROC) analysis was 0.843, which reflects an excellent test.

**Conclusions:** We have developed a rigorous machine learning model to assist physicians and decision-makers to choose patients with renal stones who are most likely to have successful SWL treatment based on their demographics and stone characteristics. The proposed machine learning model can assist physicians and decision-makers in planning for SWL treatment and allow for more effective use of limited health care resources and improve patient prognoses.

*(Interact J Med Res 2022;11(1):e33357) doi: [10.2196/33357](https://doi.org/10.2196/33357)*

**KEYWORDS**

lithotripsy; urolithiasis; machine learning; treatment outcome; ensemble learning; AdaBoost; renal stones; kidney disease

**Introduction**

Urinary stone disease, also known as urolithiasis, is a disease that occurs when a solid particle of minerals and salts is formed inside the urinary tract. A recent systematic review suggests an increasing prevalence of urolithiasis in North America over the past 3 decades [1]. In Canada, urinary stone disease is prevalent with a lifetime risk of 10% among both men and women, whereas there is a 75% chance of recurrence in 20 years after initial treatment [2].

Historically, shock wave lithotripsy (SWL) has been the most used procedure for treating upper tract urolithiasis and stones smaller than 20 mm in diameter due to its noninvasive nature, lower cost, fewer side effects, and faster recovery [3,4].

In Ontario, Canada, SWL is a regionalized and limited resource. St. Michael's Hospital in downtown Toronto is one of the only 3 centers in the province offering this service. Wait time to access SWL treatment in Canada ranges from 1 day to 1 year, with a mean wait time of 8.4 weeks in Ottawa and 8 weeks in Toronto [5]. Considering the intolerability of the pain associated with stone disease and long wait times, some patients opt for more invasive therapies such as ureteroscopy to gain access to faster treatment.

While SWL is the predominant treatment, the reported failure rate of SWL after the first session ranges from 30% to 89% [6-8]. The failure rate can be reduced significantly by identifying the candidates who are most likely to benefit from SWL, which would optimize treatment results for SWL candidates and allow for the most effective use of limited medical resources.

To identify the predictive factors of SWL outcome, several studies have focused on statistical analyses of patient

characteristics using bivariate and/or multivariate analysis [4,9-11]. The advantage and strength of machine learning is its ability to synthesize complex combinations of various attributes [12,13]. Our objective for this study was to construct a robust machine learning model that can predict SWL results to assist practitioners in their decision-making.

**Methods****Ethics Approval**

This study received ethics approval from the Office of Research Ethics at York University (certificate number STU 2019-139) and St. Michael Research Ethics Board (approval number 16-167).

**Data Set**

We assessed a data set of patients aged  $\geq 18$  years receiving SWL treatment at St. Michael's Hospital between 1998 and 2016. The data set comprised the records of 37,013 patients.

We excluded the data of patients with special conditions (eg, staghorn calculi, horseshoe kidney, caliceal diverticula, duplex collecting systems, solitary kidneys, musculoskeletal deformities) and stones larger than 25 mm in diameter. The remaining data set consisted of 57,485 SWL procedures that were performed on 31,569 patients during this period, which were used as a training data set to build the model. Several factors can impact SWL treatment outcome, including stone location and age; the choice of the attributes was guided by input from clinical experts and a literature review [4,10,11,14]. We retained 17 attributes that were most relevant to SWL success and were available in our database (Table 1).

**Table 1.** Training set attributes and corresponding values.

Attribute	Value
Kidney side	Left or right
Electrode used	Integer (1 to 3)
Stone treatment number	Integer
Number of shocks	Integer
Stone locations	Lower calyx, lower ureter, middle calyx, middle ureter, pelvis, upper calyx, upper ureter, ureterovesical junction, renal pelvis
Area of stone	Integer (mm <sup>2</sup> )
Gender	Female or male
BMI	Real number (kg/m <sup>2</sup> )
Age	18-95
Number of stones	Integer
Family history	True or false
Asymptomatic	True or false
Stent insertion	True or false
Shock frequency	120, 90, 60
Antibiotic	True or false
Shock maximum voltage	Integer
Lithotripter models	Dornier MFL 5000, Philips LithoTron, Storz Modulith SLX-F2
Outcome	Success or failure

### Defining Success and Failure of SWL on the Training Data Set

The failure or success of SWL in the training data set was based on whether there was a retreatment plan for the same patient and same stone within 90 days after initial treatment or not. The effectiveness of the lithotripter machine was measured by success rates on the training set.

### Ensemble Learning Technique

To predict the treatment outcome for SWL candidates, we used the AdaBoost algorithm based on the ensemble learning method, a machine learning technique that combines several base classifiers in various formats to produce a more robust and optimal classification model. Compared to other conventional machine learning algorithms, ensemble learning techniques are more stable, faster, simpler, and easier to program [15-19].

AdaBoost combines multiple weak classifiers that are sequentially applied to the data set. In each iteration, after the weak classifier is called, misclassified item sets are detected and given higher weight to increase the emphasis of the weak classifier on them in the next round. The final classification model is then generated as a linear combination of these weak classifiers with their assigned weights as their coefficient [19]. We used 10-fold cross-validation for AdaBoost.

### Performance Evaluation

To compare AdaBoost's performance to that of other classifiers, we used 5 classification algorithms to predict SWL failure (retreatment required <3 months), namely C4.5, naïve Bayes,

Bayesian network, K-nearest neighbors, and multilayer perceptron, and used *t* tests to perform pairwise comparisons of the performance of the AdaBoost algorithm against that of the other 5 classification models. The measurements used to determine the models' performance were sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) [20], accuracy, F1 score [14,21], and Matthews correlation coefficient [22]. Machine learning was performed using WEKA (version 3.9; University of Waikato) [23]. We used 10-fold cross-validation for performance evaluation.

### Generalizability of the Model

Classifiers were assessed for generalizability using the testing data set of 864 patients who had their preoperative and postoperative follow-ups conducted at the same center, and whose SWL procedure success and failure was determined by computed tomography (CT) scan of patients 3 months after the initial therapy. The testing data set was not included in the training set used to build the model. We employed the undersampling technique to resolve the imbalance in data by removing random examples from the majority class. *SpreadSubsample* was the Java class implemented for subsampling the original training set. We matched the ratio of success to failure in the training set to the ratio observed in our testing set, which was 40% to 60%.

## Results

### The AdaBoost Model

A total of 30 iterations were used for the AdaBoost model. Although increasing the number of iterations usually increases the accuracy of the model, we ceased adding more iterations to the model to avoid overfitting.

Research has shown that applying the boosting method to any weak classifier can drastically enhance the accuracy of the

classification model [24]. Indeed, the accuracy of applying the base learner (Decision Stump) alone on our data set was 67.8%. However, with the ensemble method, we could boost this accuracy by 9% to 76.38%, which demonstrates the superiority of the boosting method.

### Model Performance

Table 2 shows the comparison of the AdaBoost model against the other 5 classification techniques in terms of 4 different performance measurements. AdaBoost performed significantly better than all 5 other classifiers on all performance measures.

**Table 2.** Performance comparison of AdaBoost against 5 other classifiers.

Measurement	AdaBoostM1	C4.5	Naïve Bayes	Multilayer perceptron	Bayesian network	KNN <sup>a</sup>
Accuracy	77.59	75.26 <sup>b</sup>	75.82 <sup>b</sup>	69.11 <sup>b</sup>	76.49 <sup>b</sup>	57.52 <sup>b</sup>
MCC <sup>c</sup>	0.53	0.46 <sup>b</sup>	0.47 <sup>b</sup>	0.34 <sup>b</sup>	0.49 <sup>b</sup>	0.09 <sup>b</sup>
F1 score	0.84	0.82 <sup>b</sup>	0.83 <sup>b</sup>	0.76 <sup>b</sup>	0.83 <sup>b</sup>	0.66 <sup>b</sup>
Area under ROC <sup>d</sup>	0.80	0.74 <sup>b</sup>	0.75 <sup>b</sup>	0.74 <sup>b</sup>	0.78 <sup>b</sup>	0.54 <sup>b</sup>

<sup>a</sup>KNN: K-nearest neighbors.

<sup>b</sup>Statistically significant.

<sup>c</sup>MCC: Matthews correlation coefficient.

<sup>d</sup>ROC: receiver operating characteristic.

The sensitivity of the model was 0.875 (ie, 87.5% of all patients with successful SWL treatment were correctly identified by our model). On the other hand, the specificity was 0.6528 (ie, 65.3% of all patients with failed SWL treatment were correctly identified by our model).

Furthermore, the PPV (ie, the probability that subjects with a success prediction truly succeeded in the treatment) was 0.7159. Meanwhile, the NPV (ie, the probability that subjects with a failure prediction have truly failed the treatment) was 0.839.

Finally, we measured the correlation between the attributes and the class; the top 5 contributors detected were the number of stones, the area of the stone, the stone treatment number, the lithotripter machine, and the patient's age.

## Discussion

### Principal Findings

Our goal was to evaluate the ability of machine learning techniques to assist in effective decision-making for the treatment of urolithiasis with SWL by accurately predicting the SWL results. We have shown that AdaBoost provided superior prediction ability compared to 5 other classification techniques.

The AUC (area under the ROC [receiver operating characteristic] curve or C-statistic) of the ROC analysis for our prediction model was 0.843, which reflects an excellent test (a C-statistic value of 0.8-0.89 indicates an excellent test, 0.7-0.79 indicates a good test, and 0.51-0.69 indicates a poor test) [25].

The model had high sensitivity and medium specificity. Given that we are interested in identifying the patients for whom SWL has a low chance of success to plan for alternative procedures, the NPV of 0.839 demonstrated that the model can predict with

high probability if a subject will fail the treatment. Considering how scarce and expensive health care resources are, it is important to allocate those limited resources appropriately [26,27]; our model allows for appropriate allocation by informing physicians about patients who are not likely to benefit from SWL.

Recently, Choo et al [28] developed a decision tree algorithm C 5.0 for the same purpose of predicting treatment outcomes for SWL, including 15 predictive attributes on only 791 patients. Although their model had high accuracy (92.3%), some of its branches included fewer than 10 patients each. Considering that our AdaBoost-based model outperformed the decision tree algorithm in all performance measurements, we can expect it to yield better accuracy if other predictive attributes (ie, skin-to-stone distance, stone Hounsfield unit, creatinine level, stone composition, etc [3,4]) were included in the data set in a future study.

Our results show that the 3 different models of lithotripters did not significantly change the SWL treatment success rate ( $P=.81$ ). This finding suggests that frequently upgrading the technology of SWL machines does not necessarily result in a better outcome, whereas optimizing patient and stone selection is a more important factor in predicting the outcome of the SWL.

### Limitations

A limitation of this study was the lack of follow-up data for some of the patients enrolled. As a result, a treatment's failure was defined only based on having retreatment of a stone in the same center (St. Michael's Hospital) within 3 months of the initial SWL. However, to overcome this limitation and test the robustness of our model, we used 864 records that included only patients who had their complete preoperative and

postoperative follow-ups conducted at St. Michael's Hospital. This subset of the data set was not used for training the model. The follow-up data, the stone-free rate, and the success of treatment for these patients were assessed based on the follow-up CT scan administered at St. Michael's Hospital 3 months after the initial SWL.

Another limitation is that some attributes that have been shown to be predictive of SWL outcome in recent studies, such as stone density, skin-to-stone distance, and stone composition [10], were not available in our database since these data points were not known or collected 20 years ago.

## Conclusion

We built a machine learning model to assist physicians and decision-makers to choose the best treatment option for SWL

candidates based on their demographics and stone characteristics, which can result in improved prognoses. The model was generated based on the AdaBoost algorithm.

A pairwise comparison was performed between the AdaBoost classifier and 5 other classification techniques in terms of their accuracy, Matthews correlation coefficient, area under the ROC curve, and root mean squared error. The findings of these comparisons suggest the superiority of AdaBoost compared to those algorithms.

We aim to explore several meaningful research directions in the future. First, we will develop new models and architectures that are more robust and efficient by utilizing deep learning techniques. Second, our proposed ensemble learning approach can be applied to more comprehensive databases for more applications to ascertain the applicability of the model [29-32].

## Acknowledgments

We would like to thank Daniela Ghiculete for her continued support in establishing the research protocol, accessing the database, and assisting in getting the Research Ethics Board approval.

This research is supported by the Natural Sciences and Engineering Research Council of Canada, the York Research Chairs program, and the Faculty of Health at York University.

## Data Availability

The data sets generated and/or analyzed in this study are not publicly available due to patient data privacy reasons but are available from the corresponding author upon reasonable request.

## Authors' Contributions

RM preprocessed and compiled raw data, built the model, assessed the performance, and put in writing the findings. CE redacted, edited, and constructed the structure and content of the paper and study design, and assisted in the evaluation of the analysis. He gave continuous counsel and advised about the research objective. KTP provided support and training on the medical aspects of the research objective and contributed to the development of the research protocol and design. MH assisted and contributed to editing the draft, study design, and literature review. JH provided support and training for the study design, editing of the draft, and literature review.

## Conflicts of Interest

None declared.

## References

1. Chen Z, Proserpi M, Bird VY. Prevalence of kidney stones in the USA: The National Health and Nutrition Evaluation Survey. *J Clin Urol* 2018 Nov 26;12(4):296-302. [doi: [10.1177/2051415818813820](https://doi.org/10.1177/2051415818813820)]
2. Stamatelou KK, Francis ME, Jones CA, Nyberg LM, Curhan GC. Time trends in reported prevalence of kidney stones in the United States: 1976-1994. *Kidney Int* 2003 May;63(5):1817-1823 [FREE Full text] [doi: [10.1046/j.1523-1755.2003.00917.x](https://doi.org/10.1046/j.1523-1755.2003.00917.x)] [Medline: [12675858](https://pubmed.ncbi.nlm.nih.gov/12675858/)]
3. El-Assmy A, El-Nahas AR, Abo-Elghar ME, Eraky I, El-Kenawy MR, Sheir KZ. Predictors of Success after Extracorporeal Shock Wave Lithotripsy (ESWL) for Renal Calculi Between 20–30 mm: A Multivariate Analysis Model. *TSW Urology* 2006 Mar 23;1:93-100. [doi: [10.1100/tswurol.2006.74](https://doi.org/10.1100/tswurol.2006.74)]
4. Takahara K, Ibuki N, Inamoto T, Nomi H, Ubai T, Azuma H. Predictors of success for stone fragmentation and stone-free rate after extracorporeal shockwave lithotripsy in the treatment of upper urinary tract stones. *Urol J* 2012;9(3):549-552. [doi: [10.22037/uj.v9i3.1590](https://doi.org/10.22037/uj.v9i3.1590)]
5. Piggott K, Bell C. Looking for Lithotripsy: Accessibility and Portability of Canadian Healthcare. *Hcpol* 2013 Nov 27;9(2):65-75. [doi: [10.12927/hcpol.2013.23605](https://doi.org/10.12927/hcpol.2013.23605)]
6. Abe T, Akakura K, Kawaguchi M, Ueda T, Ichikawa T, Ito H, et al. Outcomes of shockwave lithotripsy for upper urinary-tract stones: a large-scale study at a single institution. *J Endourol* 2005 Sep;19(7):768-773. [doi: [10.1089/end.2005.19.768](https://doi.org/10.1089/end.2005.19.768)] [Medline: [16190825](https://pubmed.ncbi.nlm.nih.gov/16190825/)]

7. Park H, Park M, Park T. Two-year experience with ureteral stones: extracorporeal shockwave lithotripsy v ureteroscopic manipulation. *J Endourol* 1998 Dec;12(6):501-504. [doi: [10.1089/end.1998.12.501](https://doi.org/10.1089/end.1998.12.501)] [Medline: [9895251](https://pubmed.ncbi.nlm.nih.gov/9895251/)]
8. Srisubatt A, Potisat S, Lojanapiwat B, Setthawong V, Laopaiboon M. Extracorporeal shock wave lithotripsy (ESWL) versus percutaneous nephrolithotomy (PCNL) or retrograde intrarenal surgery (RIRS) for kidney stones. *Cochrane Database Syst Rev* 2014 Nov 24(11):CD007044. [doi: [10.1002/14651858.CD007044.pub3](https://doi.org/10.1002/14651858.CD007044.pub3)] [Medline: [25418417](https://pubmed.ncbi.nlm.nih.gov/25418417/)]
9. El-Assmy A, El-Nahas AR, Abou-El-Ghar ME, Awad BA, Sheir KZ. Kidney stone size and hounsfield units predict successful shockwave lithotripsy in children. *Urology* 2013 Apr;81(4):880-884. [doi: [10.1016/j.urology.2012.12.012](https://doi.org/10.1016/j.urology.2012.12.012)] [Medline: [23395121](https://pubmed.ncbi.nlm.nih.gov/23395121/)]
10. Gücük A, Uyetürk U. Usefulness of hounsfield unit and density in the assessment and treatment of urinary stones. *World J Nephrol* 2014 Nov 06;3(4):282-286 [FREE Full text] [doi: [10.5527/wjn.v3.i4.282](https://doi.org/10.5527/wjn.v3.i4.282)] [Medline: [25374823](https://pubmed.ncbi.nlm.nih.gov/25374823/)]
11. Nakasato T, Morita J, Ogawa Y. Evaluation of Hounsfield Units as a predictive factor for the outcome of extracorporeal shock wave lithotripsy and stone composition. *Urolithiasis* 2015 Feb 20;43(1):69-75. [doi: [10.1007/s00240-014-0712-x](https://doi.org/10.1007/s00240-014-0712-x)] [Medline: [25139151](https://pubmed.ncbi.nlm.nih.gov/25139151/)]
12. El Morr C, Ali-Hassan H. *Analytics in Healthcare: A Practical Introduction*. Cham, Switzerland: Springer International Publishing; 2019.
13. Imbus JR, Randle RW, Pitt SC, Sippel RS, Schneider DF. Machine learning to identify multigland disease in primary hyperparathyroidism. *J Surg Res* 2017 Nov;219:173-179 [FREE Full text] [doi: [10.1016/j.jss.2017.05.117](https://doi.org/10.1016/j.jss.2017.05.117)] [Medline: [29078878](https://pubmed.ncbi.nlm.nih.gov/29078878/)]
14. Bozzini G, Verze P, Arcaniolo D, Dal Piaz O, Buffi NM, Guazzoni G, et al. A prospective randomized comparison among SWL, PCNL and RIRS for lower calyceal stones less than 2 cm: a multicenter experience: A better understanding on the treatment options for lower pole stones. *World J Urol* 2017 Dec 5;35(12):1967-1975. [doi: [10.1007/s00345-017-2084-7](https://doi.org/10.1007/s00345-017-2084-7)] [Medline: [28875295](https://pubmed.ncbi.nlm.nih.gov/28875295/)]
15. Fern A, Givan R. Online Ensemble Learning: An Empirical Study. *Machine Learning* 2003;53:71-109. [doi: [10.1023/A:1025619426553](https://doi.org/10.1023/A:1025619426553)]
16. Freund Y, Schapire R. Experiments with a new boosting algorithm. In: *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning*. 1996 Presented at: ICML'96; July 3-6; Bari, Italy p. 148-156. [doi: [10.5555/3091696.3091715](https://doi.org/10.5555/3091696.3091715)]
17. Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors). *Ann Statist* 2000 Apr 1;28(2). [doi: [10.1214/aos/1016218223](https://doi.org/10.1214/aos/1016218223)]
18. Huang X, Huang Y, Wen M, An A, Liu Y, Poon J. Applying data mining to pseudo-relevance feedback for high performance text retrieval. 2006 Presented at: Sixth International Conference on Data Mining (ICDM'06); Dec 18-22; Hong Kong, China. [doi: [10.1109/icdm.2006.22](https://doi.org/10.1109/icdm.2006.22)]
19. Schapire R. Explaining AdaBoost. In: Schölkopf B, Luo Z, Vovk V, editors. *Empirical Inference*. Cham, Switzerland: Springer Berlin Heidelberg; 2013:37-52.
20. Qiu P. The Statistical Evaluation of Medical Tests for Classification and Prediction. *J Am Stat Assoc* 2005 Jun;100(470):705. [doi: [10.1198/jasa.2005.s19](https://doi.org/10.1198/jasa.2005.s19)]
21. Thornley C. *Advances in Information Retrieval*. *J Doc* 2012 Aug 31;68(5). [doi: [10.1108/jd.2012.27868eaa.002](https://doi.org/10.1108/jd.2012.27868eaa.002)]
22. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* 2020 Jan 02;21(1):6 [FREE Full text] [doi: [10.1186/s12864-019-6413-7](https://doi.org/10.1186/s12864-019-6413-7)] [Medline: [31898477](https://pubmed.ncbi.nlm.nih.gov/31898477/)]
23. Frank E, Hall M, Holmes G, Kirkby R, Pfahringer B, Witten I. *Weka-A Machine Learning Workbench for Data Mining*. In: *Data Mining and Knowledge Discovery Handbook*. Cham, Switzerland: Springer; 2009:1269-1277.
24. Drucker H, Cortes C. Boosting Decision Trees. *Advances in neural information processing systems*. In: *Proceedings of the 8th International Conference on Neural Information Processing Systems*. 1996 Presented at: NIPS'95; Nov 27-Dec 2; Denver, Colorado p. 479-485.
25. Scott AJ, Hosmer DW, Lemeshow S. *Applied Logistic Regression*. *Biometrics* 1991 Dec;47(4):1632. [doi: [10.2307/2532419](https://doi.org/10.2307/2532419)]
26. El Morr C, editor. *Research Perspectives on the Role of Informatics in Health Policy and Management*. Hershey, PA: IGI Global; 2014.
27. El Morr C. *Introduction to Health Informatics: A Canadian Perspective*. Toronto, ON: Canadian Scholars' Press; 2018.
28. Choo MS, Uhm S, Kim JK, Han JH, Kim D, Kim J, et al. A Prediction Model Using Machine Learning Algorithm for Assessing Stone-Free Status after Single Session Shock Wave Lithotripsy to Treat Ureteral Stones. *J Urol* 2018 Dec;200(6):1371-1377. [doi: [10.1016/j.juro.2018.06.077](https://doi.org/10.1016/j.juro.2018.06.077)] [Medline: [30036513](https://pubmed.ncbi.nlm.nih.gov/30036513/)]
29. Huang X, Zhong M, Si L. York University at TREC 2005: Genomics Track. 2005. URL: <https://trec.nist.gov/pubs/trec14/papers/yorkku-huang2.geo.pdf> [accessed 2022-03-11]
30. Liu Y, An A, Huang X, editors. Boosting Prediction Accuracy on Imbalanced Datasets with SVM Ensembles. In: *Advances in Knowledge Discovery and Data Mining*. Berlin, Heidelberg: Springer; 2006.
31. Liu Y, Huang X, An A, Yu X. ARSA: a sentiment-aware model for predicting sales performance using blogs. In: *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*. 2007 Presented at: SIGIR '07; July 23-27; Amsterdam, The Netherlands p. 607-614. [doi: [10.1145/1277741.1277845](https://doi.org/10.1145/1277741.1277845)]

32. Yu X, Liu Y, Huang X, An A. Mining Online Reviews for Predicting Sales Performance: A Case Study in the Movie Domain. *IEEE Trans Knowl Data Eng* 2012 Apr;24(4):720-734. [doi: [10.1109/tkde.2010.269](https://doi.org/10.1109/tkde.2010.269)]

## Abbreviations

**AUC:** area under the receiver operating characteristic curve

**CT:** computed tomography

**NPV:** negative predictive value

**PPV:** positive predictive value

**ROC:** receiver operating characteristic

**SWL:** shock wave lithotripsy

*Edited by T Leung; submitted 03.09.21; peer-reviewed by CY Yang, J Kim; comments to author 13.12.21; revised version received 24.12.21; accepted 11.02.22; published 16.03.22*

*Please cite as:*

*Moghisi R, El Morr C, Pace KT, Hajiha M, Huang J*

*A Machine Learning Approach to Predict the Outcome of Urinary Calculi Treatment Using Shock Wave Lithotripsy: Model Development and Validation Study*

*Interact J Med Res* 2022;11(1):e33357

URL: <https://www.i-jmr.org/2022/1/e33357>

doi: [10.2196/33357](https://doi.org/10.2196/33357)

PMID:

©Reihaneh Moghisi, Christo El Morr, Kenneth T Pace, Mohammad Hajiha, Jimmy Huang. Originally published in the Interactive Journal of Medical Research (<https://www.i-jmr.org/>), 16.03.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.i-jmr.org/>, as well as this copyright and license information must be included.