

Predicting Stone-Free Status of Percutaneous Nephrolithotomy With Machine Learning System: Comparative Analysis With Guy's Stone Score and S.T.O.N.E Score System

Hong Zhao

Shanghai Xu-Hui Central Hospital

Wanling li

Shanghai Xu-Hui Central Hospital

Junsheng Li

Shanghai Xu-Hui Central Hospital

Li Li

Shanghai Xu-Hui Central Hospital

Hang Wang

Zhongshan Hospital

Jianming Guo (✉ dr gjm2021@126.com)

Zhongshan Hospital

Research Article

Keywords: Machine learning, Prediction, Percutaneous Nephrolithotomy, Stone-free status, Guy's Stone Score, S.T.O.N.E Score system

Posted Date: September 22nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-823907/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Purpose

Using machine learning methods (MLMs) to predict stone-free status after percutaneous nephrolithotomy (PCNL). We compared the performance of this system with Guy's stone score and S.T.O.N.E score system.

Materials and Methods

Data from 222 patients (90 females, 41%) who underwent PCNL at our center were used. Twenty-six parameters, including individual variables, renal and stone factors, surgical factors were used as input data for MLMS. We evaluate the efficacy of four different techniques: Lasso-logistic (LL), random forest (RF), support vector machine (SVM) and Naive Bayes. Model performance was evaluated using area under curve (AUC) and compared with Guy's stone score and S.T.O.N.E score system.

Results

Overall stone free rate was 50% (111/222). To predict the stone-free status, all receiver operating characteristic curves of the four MLMs were above the curve for the Guy's stone score. The AUCs of LL, RF, SVM and Naive Bayes were 0.879, 0.803, 0.818, 0.803 respectively. Those values were higher than the AUC of the Guy's score system, 0.800. The accuracies of the MLMs (0.803–0.818%) were also superior to S.T.O.N.E score system (0.788%). Among the MLMs, Lasso-logistic showed the most favorable AUC.

Conclusion

Machine learning methods can predict stone-free rate with AUCs no inferior to those of Guy's stone score and S.T.O.N.E score system.

Introduction

With the first description of the technique in 1976^[1], Percutaneous nephrolithotomy has been widespread for the treatment of renal calculi. As a golden standard for the treatment of 2 cm kidney stones^[2]. The success rate is between 56% – 96% in various series ^[3, 4, 5, 6]. Many factors contribute to the success of stone clearance including stone volume, location, grade of hydronephrosis, infection status of the patient and surgeon's experience as well. To predict the outcomes after PCNL, some scoring systems have been used including S.T.O.N.E Score system, Guy's stone score, CROES Nephrolithometry nomogram and S-ReSC score. The S.T.O.N.E. score is based on factors determined through CT imaging, it has five variables. Guy's stone score is easy to apply and has been validated in several studies. The CROES nomogram was developed from data in a large multicentre database and has a high statistical power.

The SReSC score was relied on stone location only, providing a simple approach to grading disease complexity. Each system has advantages and disadvantages, but their ability of predicting the stone-free rate is thought comparable^[7].

Machine learning techniques have been used extensively in the field of clinical medicine, especially when used for the construction of prediction models. The out performance of ML over conventional data analysis models has been shown in urology-oncology literature ^[8, 9, 10, 11].

In predicting post-lithotripsy outcomes with machine learning, there were only three studies published, until now ^[12, 13, 14]. Alireza et al. ^[12] first used machine learning method for the predicting post-PCNL outcomes compared to current scoring systems. They found the machine learning-based software was superior in predicting SFS after PCNL, with an AUC of 0.915 compared to 0.615 (GSS) and 0.621 (CROES nomograms) ($P < 0.01$). More than 20 variables of 146 patients were inputted for the training of machine learning in their study. Alireza used support vector machine (SVM) as the machine learning technique. We know, machine learning algorithm includes some other methods, like decision trees, random forests, artificial neural networks, Bayesian learning, Deep Learning and so on. In this study, we use four machine learning methods (Lasso logistic, random forests, SVM and Naive Bayes) to predict the SFS of PCNL with the information of 222 patients. We compare the out performance of ML to Guy's score and S.T.O.N.E score system at the same time.

Patients And Methods:

The study was approved by the independent ethics committee of Xu-hui Central Hospital. Between July 2017 and January 2020, 222 patients underwent PCNL performed by one single surgeon (Dr. G.J.M.) were included in this retrospective study. All patients had computed tomography (CT) scan and IVP before surgery. Normal preoperative coagulation and negative urine culture were verified.

All percutaneous accesses were performed under general anesthesia and in prone position after retrograde ureteral catheterization. Access to the selected calyx was performed by Dr. G.J.M with the aid of ultrasound guidance by using an 18-gauge needle. The tract was dilatated with serial dilators from 8F to 20F sheath. An 18F nephroscope (Wolf) was used to insect the sheath and we used holmium laser to fragment stones with the power ranges from 60 to 90 watts. Every case was demanded to place an internal ureteral stent on a suspect for the presence of mobile residual stones. A 14F nephrostomy tube was placed in the renal pelvis or the involved calyx for most patients.

Antibiotic prophylaxis was used with the second-generation cephalosporin. The medicine was ended after the nephrostomy tube was removed.

Plain radiography of the kidneys, ureters, and bladders was obtained on postoperative day 1 to day 3 according to the state of the patient. The nephrostomy tube was removed when there were no stone residual nor clinically insignificant residual fragments (diameter less than 4 mm).^[15]

All patients were asked to take out the stent for outpatient service in one or two months after the surgery. If there were residual stones, they would have repeated PCNL, ureteroscopy, and shockwave lithotripsy (SWL). After that, all patients were evaluated with an ultrasound test or non-contrast CT scan after 3 to 6 months postoperatively. All patients accepted follow-ups at least for one year. PCNL was considered successful when the patient was stone-free or did not need any further intervention (clinically insignificant residual stone fragments [CIRF]) [16]

2.2 Machine learning Methods:

Four types of supervised machine learning algorithms (Lasso logistic, random forests, SVM and Naive Bayes) were applied in this study. A set of input variables comprising individual variables (age, sex, hypertension, diabetes, hyperlipidemia, urinary infection, renal insufficiency, preoperative hemoglobin, use of anticoagulants or antiplatelet medications, renal and stone factors (previous surgery, stone burden, stone location, grade of hydronephrosis), surgical factors (postoperative fever, septicemia, need for transfusion, length of stay, stone-free status, ancillary procedures. The results of Stone free were entered as binary: 1 (stone residual) and 0 (clinically insignificant residual stone fragments).

The machine learning models were fitted using scikit-learn 0.18 modules of python throughout this study. Using lasso regularization and cross-validation (n fold = 10) to select the best regression. We selected lambda with 1se.lambda to screen characteristic variables. The selected variables include: stone size, stone location (top / middle / bottom), a total of four variables (Fig. 1).

The original data set is randomly divided into the training set, and the test set at 7:3 (156: 66)

Lasso-logistic, SVM, NaiveBayes took the results of lasso regression screening as independent variables to establish a model and calculate the prediction accuracy.

The RF model is a machine learning model built on decision trees. In the decision tree, each node of the tree splits the data into two groups using a cutoff value within one of the features. The RF method can minimize the effect of the overfitting problem by creating an ensemble of randomized decision trees, each of which overfits the data and averages the results to find a better classification.

2.3 Statistical analysis

Continuous variables were compared using the independent sample Student t test. Model performance was evaluated using area under the receiver operating characteristic (ROC) curve (AUC), which provides a measure of the discriminatory performance of the model. Sensitivity, which is the proportion of true positives that are classified; specificity, which measures the proportion of correctly identified true negatives; and accuracy, which is the proportion of correct predictions.

Results:

222 patients (132 males, 59.5%) were enrolled. Mean age was 54.8 ± 13.3 years and mean stone burden was 563.4 ± 517.6 mm². The mean Guy's score was 3.2 ± 0.9 , the mean S.T.O.N.E. score was 8.9 ± 1.8 . Table 1 shows the preoperative factors include individual variables and renal and stone factors. Table 2 shows the actual postoperative data for these patients. Overall SFS was 50% (111/222), Fig. 2 shows the stone-free rate in each subgroup of GSS grades and the S.T.O.N.E. score systems. Number of fever and infection during hospitalization was 18.9% (42) and 8.6% (19). Postoperative blood transfusion due to significant blood loss was happened in 9 patients (4.1%). With the follow-ups at least for one year, there were 12 patients (5.4%) accepted ancillary procedures to manage residual renal stones.

We have used 4 machine learning methods to analyze the outcomes to predict the stone-free status. Table 3 shows the AUC, sensitivity, specificity, and accuracy of each prediction method to the results of the stone-free status. When using AUC as a measure of predictive model performance, as shown in Table 3, the AUC of Lasso Logistic was 0.879. It was superior to those of RF, SVM and Naive Bayes (0.803, 0.818 and 0.803, respectively). The AUCs of the GSS, S.T.O.N.E. were 0.800 and 0.844, respectively, which were lower than Lasso Logistic. Figure 3 shows the ROC curves of the four MLMs as well as the GSS and S.T.O.N.E. score system.

As show in Table 3, the accuracies of the four MLMs were also superior to those of S.T.O.N.E. score system. The sensitivities of the MLMs were 75.8 to 83.3%, which were higher than the S.T.O.N.E. score system. The machine learning system of LL recognized stone burden and stone location as the most highly weighted preoperative factors affecting the post PCNL-SFR.

Discussion:

The incidence and prevalence of kidney stones have increased by three times over the past four decades^[17]. The prevalence of kidney stones is estimated about 5–10% in Europe, 4% in South America and 1–19% in Asia currently^[18, 19]. With no doubt, kidney stones represent a considerable burden for public health-care systems.

Thromas et al.^[20] were the first to introduce Guy's stone score (GSS) to predict the success of stone-free status (SFS) after PCNL. The model is based on the shape of stone and kidney, which is reproducible and provides easy categorization of renal stones in four grades. It has a good relationship with the SFS; however, it fails to take into account the volume and density of the stone. Okhunov et al.^[21] developed S.T.O.N.E. scoring system based on non-contrast CT (NCCT). The score varies from 5 to 13 including five variables and a lower score predicts a higher stone clearance rate. Also, greater S.T.O.N.E. scores are associated with serious complications, including longer operative times (OT), greater estimated blood loss (EBL), and increased length of stay (LOS). CROES (Clinical Research Office of the Endourological Society) nomogram of Smiths et al. [22] was based on a global database study of 5830 patients. Six characteristics (stone burden, number, location, multiple, staghorn and institute-level case volume) are included in this nomogram. It achieved a remarkable of 76% prediction accuracy, but it is cucumber and time consuming.

Many studies have compared the predictive performance of these score systems in post-PCNL SFR. Most studies have examined the performance of these scoring systems to predict SFR equally, but not equally to predict complications. Totally, the AUC ranges from 0.63 to 0.853^[7]. And different scoring system has its drawbacks or limitations. For example, in Guy's score system, Partial staghorn stone was not clearly defined. S.T.O.N.E. Score scoring system relies solely on preoperative CT. CROES nomogram requires information that might not be readily available (case volume and treatment history). So one more simpler and easier for application stone score system is needed nowadays. Alireza and his colleagues^[23] were the first using machine learning methods to evaluate stone-free rate and complications after PCNL. They used ANN to predict the stone-free rate. The accuracy was 81.0–98.2%. The AUC was 0.861. In 2019, his team^[12] reported they use software to predict SFR after PCNL with the AUC of 0.915. In our study, we used four machine learning methods to predict the SFR of PCNL compared with the Guy's system and S.T.O.N.E. Score system. The machine learning methods (MLMs) include Lasso logistic, random forests, SVM and Naive Bayes. The AUC of the MLMs were superior than that of the Guy's stone score system. And the sensitivity and accuracy of MLMs were superior than that of S.T.O.N.E. Score system.

Machine learning is built on statistical frame work. Different approaches are designed to make the most accurate prediction possible. It has been proved to have a good performance to predict the SFR post-PCNL. Although we didn't got an advantageous performance of AUC of 0.915^[12], in this study we found the MLMS could predict stone-free rate with AUC no inferior to those of Guy's stone score or S.T.O.N.E score system. Machine learning algorithm mainly includes random forests, decision trees, artificial neural networks, Bayesian learning, Deep learning et al. Each approach has its advantage and disadvantage. We have tried four methods to predict stone -free rate in this study. And all of them got a fairly superior performance as well as the clinical scoring systems currently available. Machine learning methods is a good tool to predict stone-free rate with AUCs after PCNL.

So far, in the field of urinary stones, there have been few studies using machine learning methods to predict operative outcomes or help to make operative decisions. As one author commented^[25], to improve the application of MLMs in uritholiasis, two things should be considered: First, more people include urologists, statistician, and computer experts need to be involved in this project; Second, more data from different regions or population should be collected for future events prediction. We need to establish, manage and share a cross-country or nation-wide database. Through which machine learning or AI would contribute to the field of calculi or other issues in the near future.

Abbreviations

PCNL: percutaneous nephrolithotomy

MLMs: machine learning methods

LL: lasso-logistic

RF: random forest

SVM: support vector machine

AUC: area under curve

Declarations

Ethics approval and consent to participate

All experimental protocols were approved by Shanghai Xu-Hui Central Hospital ethics committee. The ethics board approval number is SOP-IEC-069-02.0-AF02. All human subjects provided written informed consent with guarantees of confidentiality. All methods were carried out in accordance with relevant guidelines and regulations.

Consent for publication

All of the writers agreed for publication of the manuscript.

Availability of data and materials

The data was open and if you need we will share. Please communicate with Dr Hong Zhao, E-mail.: drzhaohong1986@gmail.com

Competing interests

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled

Funding

There is no funding for this study.

Authors' contributions

Hong Zhao: protocol, writing

Wan-Ling Li: data analysis

Jun-Sheng Li: data collection

Li Li: management

Hang Wang: editing

Jian-Ming Guo: project development

Corresponding author : Jian-Ming Guo

Hong Zhao, Wan-ling Li contributed equally to this work.

The word count was 2109, the number of figures was 3, the number of tables was 3,
the number of references was 24.

Acknowledgements

Not applicable

Authors' information (optional)

Hong Zhao¹, Wanling li², Junsheng Li¹, Li Li¹, Hang Wang³, Jianming Guo^{3,*}

1.Hong Zhao

drzhaohong1986@gmail.com

Department of Urology, Shanghai Xu-Hui Central Hospital, Shanghai 200032, China

No 966, Huaihai Road, Shanghai

2.Wanling li

lynn8203@outlook.com

Department of General surgery, Shanghai Xu-Hui Central Hospital, Shanghai 200032, China

No 966, Huaihai Road, Shanghai

3.Junsheng Li

drlijs0520@163.com

Department of Urology, Shanghai Xu-Hui Central Hospital, Shanghai 200032, China

No 966, Huaihai Road, Shanghai

Li Li

drlili1234@126.com

Department of Urology, Shanghai Xu-Hui Central Hospital, Shanghai 200032, China

No 966, Huaihai Road Shanghai

Hang Wang

drwanghang1234@163.com

Department of Urology, Zhongshan Hospital, Fudan University, Shanghai 200032, China

No 180, FengLin Road Shanghai

Jianming Guo

drjgm2021@126.com

Department of Urology, Zhongshan Hospital, Fudan University, Shanghai 200032, China

No 180, FengLin Road Shanghai

References

1. Fernström I, Johansson B. Percutaneous pyelolithotomy. A new extraction technique. *Scand J Urol Nephrol* 1976;10: 257–259
2. Miernik Arkadiusz, Schoenthaler Martin, Wilhelm Konrad, et al. Combined semirigid and flexible ureterorenoscopy via a large ureteral access sheath for kidney stones > 2 cm: a bicentric prospective assessment.[J]. *World J Urol*, 2014, 32: 697–702.
3. De la Rosette J., Assimos D., Desai M, et al. The clinical research office of the endourological society percutaneous nephrolithotomy global study: indications, complications, and outcomes in 5803 patients. *J Endourol*. 2011; 25:11–17.
4. Labadie K., Okhunov Z., Akhavein A., et al. Evaluation and comparison of urolithiasis scoring systems in percutaneous kidney stone surgery. *J Urol*. 2015;193:154–159.
5. Matlaga B.R., Kim S.C., Lingeman J.E. Improving outcomes of percutaneous nephrolithotomy access. *EAU Update Ser*. 2005; 3:37–43.
6. Akman T., Binbay M., Yuruk E., et al. Tubeless procedure is most important factor in reducing length of hospitalization after percutaneous nephrolithotomy: results of univariable and multivariable models. *Urology*. 2011; 77:299–304.
7. Wu Wayland J, Okeke Zeph, Current clinical scoring systems of percutaneous nephrolithotomy outcomes.[J]. *Nat Rev Urol*, 2017, 14: 459–469.
8. Suarez-Ibarrola R, Hein S, Reis G, et al. Current and future applications of machine and deep learning in urology: A review of the literature on urolithiasis, renal cell carcinoma, and bladder and prostate cancer. *World J Urol* 2019. DOI: 10.1007/s00345-019-03000-5.

9. Andras I, Mazzone E, van Leeuwen FWB, et al. Artificial intelligence and robotics: A combination that is changing the operating room. *World J Urol* 2019. DOI: 10.1007/s00345-019-03037-6.
10. Hung AJ, Chen J, Che Z, et al. Utilizing machine learning and automated performance metrics to evaluate robotassisted radical prostatectomy performance and predict outcomes. *J Endourol* 2018;32:438–444.
11. Ström Peter, Kartasalo Kimmo, Olsson Henrik, et al. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study.[J]. *Lancet Oncol*, 2020, 21: 222–232.
12. Aminsharififi A, Irani D, Tayebi S, 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 Endourol* 2020;34:692–699.
13. De Perrot T, Hofmeister J, Burgermeister S, et al (2019) Differentiating kidney stones from phleboliths in unenhanced low-dose computed tomography using radiomics and machine learning. *Eur Radiol* 29(9):4776–4782
14. Shabaniyan T, Parsaei H, Aminsharif A, et al (2019) An artificial intelligence-based clinical decision support system for large kidney stone treatment. *Australas Phys Eng Sci Med* 42(3):771–779
15. Harraz Ahmed M., Osman Yasser., El-Nahas Ahmed R., et al.(2017). Residual stones after percutaneous nephrolithotomy: comparison of intraoperative assessment and postoperative non-contrast computerized tomography. *World J Urol*, 35(8), 1241–1246. doi:10.1007/s00345-016-1990-4
16. Rassweiler JJ, Renner C, Eisenberger F. The management of complex renal stones. *BJU Int* 2000;86:919–928.
17. Thongprayoon Charat., Krambeck Amy E., Rule Andrew D.(2020). Determining the true burden of kidney stone disease. *Nat Rev Nephrol*, 16(12), 736–746. doi:10.1038/s41581-020-0320-7
18. Liu Yu, Chen Yuntian, Liao Banghua et al. Epidemiology of urolithiasis in Asia.[J]. *Asian J Urol*, 2018, 5: 205–214.
19. Sorokin Igor, Mamoulakis Charalampos, Miyazawa Katsuhito et al. Epidemiology of stone disease across the world.[J]. *World J Urol*, 2017, 35: 1301–1320.
20. Thomas K., Smith N.C., Hegarty N., et al. The Guy's stone score – grading the complexity of percutaneous nephrolithotomy procedures. *Urology*. 2011; 78:277–281.
21. Okhunov Z., Friedlander J.I., George A.K., et al.. S.T.O.N.E. nephrolithometry: novel surgical classification system for kidney calculi. *Urology*. 2013; 81:1154–1159.
22. Smith A., Averch T.D., Shahrour K., et al. A nephrolithometric nomogram to predict treatment success of percutaneous nephrolithotomy. *J Urol*. 2013; 190:149–156.
23. Aminsharifi A, Irani D, Pooyesh S, et al. Artificial Neural Network System to Predict the Postoperative Outcome of Percutaneous, *J Endourol*. 2017;31(5):461–467.

24. Peng L, J Pan, Yang D, et al. The prospect of machine learning in predicting post-lithotripsy outcomes[J]. World Journal of Urology, 2020(6).

Tables

Due to technical limitations, table 1,2,3 is only available as a download in the Supplemental Files section.

Figures

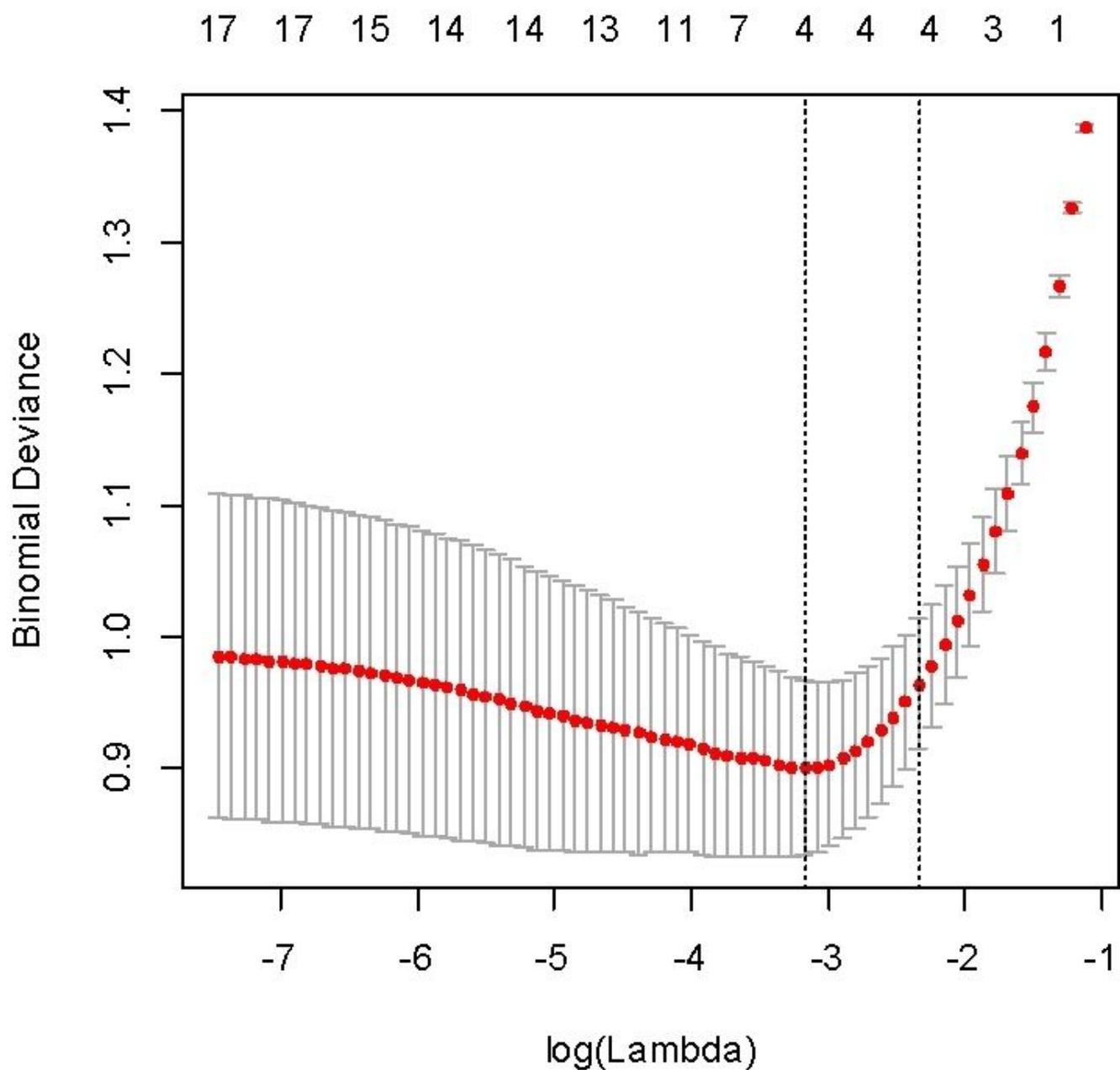


Figure 1

Using lasso regularization to select characteristic variables

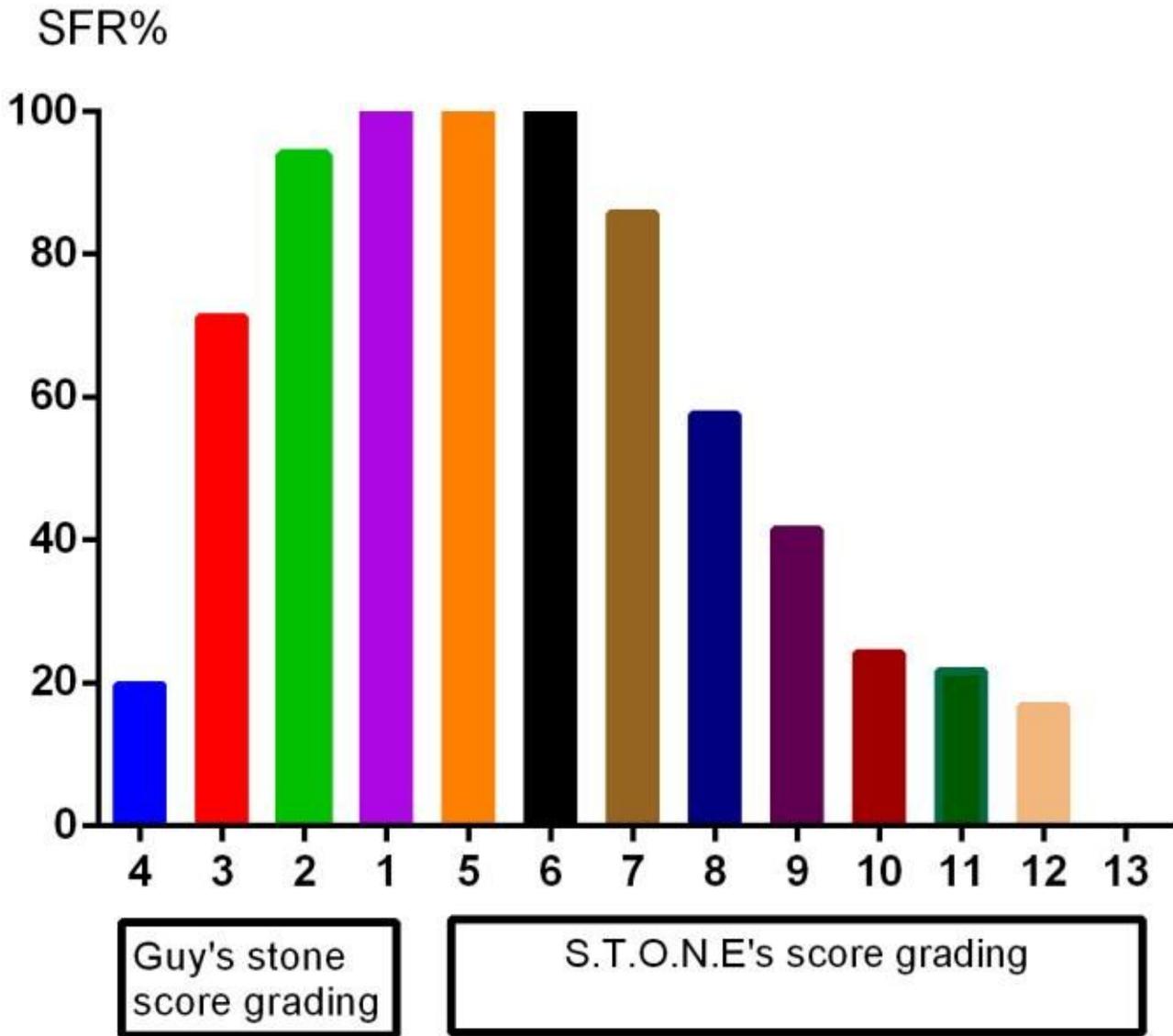


Figure 2

shows the stone-free rate in each subgroup of GSS grades and the S.T.ON.E score systems.

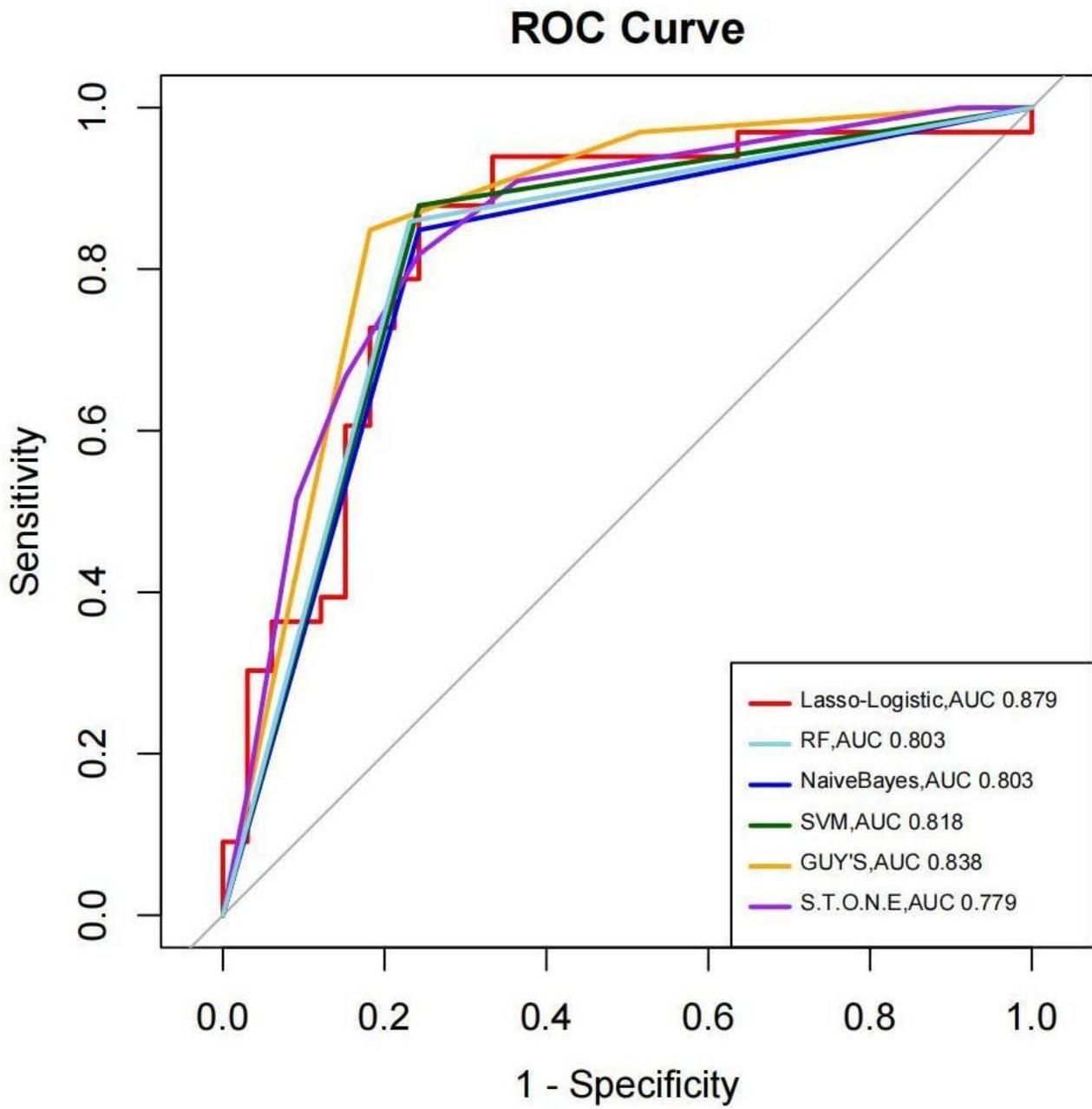


Figure 3

shows the ROC curves of the four MLMs as well as the GSS and S.T.O.N.E score system.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [table1.jpg](#)

- [table2.jpg](#)
- [table3.jpg](#)