

Original Paper

Automated Generation of Personalized Shock Wave Lithotripsy Protocols: Treatment Planning Using Deep Learning

Zhipeng Chen¹, PhD; Daniel D Zeng², PhD; Ryan G N Seltzer³, PhD; Blake D Hamilton⁴, MD

¹Shenzhen Artificial Intelligence and Data Science Institute (Longhua), Longhua, Shenzhen, China

²The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China

³Translational Analytics and Statistics, Tucson, AZ, United States

⁴School of Medicine, University of Utah, Salt Lake City, UT, United States

Corresponding Author:

Zhipeng Chen, PhD

Shenzhen Artificial Intelligence and Data Science Institute (Longhua)

Building 26, Technology Innovation Center

Hongshan 6979

Longhua, Shenzhen, 518110

China

Phone: 86 21071934

Email: zhipengchen@saidi.org.cn

Abstract

Background: Though shock wave lithotripsy (SWL) has developed to be one of the most common treatment approaches for nephrolithiasis in recent decades, its treatment planning is often a trial-and-error process based on physicians' subjective judgement. Physicians' inexperience with this modality can lead to low-quality treatment and unnecessary risks to patients.

Objective: To improve the quality and consistency of shock wave lithotripsy treatment, we aimed to develop a deep learning model for generating the next treatment step by previous steps and preoperative patient characteristics and to produce personalized SWL treatment plans in a step-by-step protocol based on the deep learning model.

Methods: We developed a deep learning model to generate the optimal power level, shock rate, and number of shocks in the next step, given previous treatment steps encoded by long short-term memory neural networks and preoperative patient characteristics. We constructed a next-step data set (N=8583) from top practices of renal SWL treatments recorded in the International Stone Registry. Then, we trained the deep learning model and baseline models (linear regression, logistic regression, random forest, and support vector machine) with 90% of the samples and validated them with the remaining samples.

Results: The deep learning models for generating the next treatment steps outperformed the baseline models (accuracy = 98.8%, F1 = 98.0% for power levels; accuracy = 98.1%, F1 = 96.0% for shock rates; root mean squared error = 207, mean absolute error = 121 for numbers of shocks). The hypothesis testing showed no significant difference between steps generated by our model and the top practices ($P=.480$ for power levels; $P=.782$ for shock rates; $P=.727$ for numbers of shocks).

Conclusions: The high performance of our deep learning approach shows its treatment planning capability on par with top physicians. To the best of our knowledge, our framework is the first effort to implement automated planning of SWL treatment via deep learning. It is a promising technique in assisting treatment planning and physician training at low cost.

(*JMIR Med Inform* 2021;9(5):e24721) doi: [10.2196/24721](https://doi.org/10.2196/24721)

KEYWORDS

nephrolithiasis; extracorporeal shock wave therapy; lithotripsy; treatment planning; deep learning; artificial intelligence

Introduction

Shock wave lithotripsy (SWL, or extracorporeal shock wave lithotripsy) has been considered as a safe and effective noninvasive treatment option for nephrolithiasis since its introduction in early 1980s [1]. Reported SWL stone-free rates

approach 74%-88% [2,3]; however, it is not without risk. Common contraindications to SWL include pregnancy, coagulopathy or use of platelet aggregation inhibitors, aortic aneurysms, severe untreated hypertension, and untreated urinary tract infections [4]. Failure of SWL treatment results in

unnecessary exposure to various complications, such as loin pain, dysuria, analgesia, hematuria, and infection [3,5].

Given such risks, previous studies have identified proper patient selection, modifications in treatment technique, and employment of adjunctive measures as elements to improve SWL outcomes [6]. The treatment outcomes are strongly affected by a variety of preoperative patient characteristics (PPC), including BMI [7-9], stone location, overall stone burden [4], skin-to-stone distance [10,11], stone composition [12,13], stone density [14-17], and variation coefficients of stone density [18]. Various studies have also demonstrated that precise targeting [19,20] and tight coupling [21,22] increase fragmentation probability.

Appropriate control over shock wave delivery has a strong impact on treatment success and minimal complications. A treatment plan for shock wave delivery is a series of shock wave delivery steps with a specified power level, shock rate, and number of shocks; a successful sample SWL treatment plan is

shown in Table 1. A plan precisely specifies step-by-step power levels, shock rates, and number of shocks. Each treatment step has a single power level, a constant shock rate, and shocks usually between 500-2500 [23-28]. Physicians are obliged to design plans that both deliver sufficient energy for breaking stones and minimize damage to body tissues. While the range of shock rates is typically 30-180 shocks/minute, a shock rate of 60-90 shocks/minute has been shown to improve efficacy [29-31] and decrease potential injury risks. The main reason is that the slower shock rate of 60-90 shocks/minute allows time for cavitation bubbles caused by the shock to disperse before the next shock arrives. Physicians can check stone fragmentation via x-ray. If the fragments of treated stones are ≤ 4 mm, they typically pass on their own without further treatment. An SWL treatment has to be stopped to reduce risks of tissue damage when the shock number reaches the maximum limit, even though the treated stone has not broken up.

Table 1. A sample shock wave lithotripsy (SWL) treatment plan.

Shock wave delivery steps	Power level	Shock rate (per minute)	Number of shocks
Step 1	1	120	100
Step 2	2	120	100
Step 3	3	120	100
Step 4	4	120	100
Step 5	5	120	100
Step 6	6	120	100
Step 7	7	120	100
Step 8	8	120	2300

Effective fragmentation leads to fewer shocks overall and therefore less damage to tissue [32,33]. In order to maximize treatment effect and control tissue damage, ramping protocols have been developed. The low-energy pretreatment allows for better pain management, thus preventing movement and subsequent decoupling of the shock head [34]. Clinical trials support that stepwise voltage ramping is associated with less tissue damage compared with a fixed maximal voltage protocol [23,25,26,35].

Although the strength, rates, and total number of shock waves are identified as the important factors of SWL treatment outcomes, there is no case-by-case guideline for physicians to optimize shock wave delivery protocols that take into account patient demographics and stone characteristics. The optimal energy delivery strategy remains controversial. In vitro and in vivo studies suggest that the strategy of ramping up shock wave energy is beneficial to improve fragmentation and stone clearance and limit renal damage, but clinical results are discordant [6,23]. In the current planning process, physicians adopt a trial-and-error approach to tune treatment plans. This approach involves nonintuitive iterations based on physicians'

subjective decisions. Inexperienced physicians using this method may be more apt to produce inefficient or ineffective treatment plans. Such dependence on physicians' unique experience also leads to significant variability in the quality and consistency of treatment delivery. Moreover, different types of machines have different designs and different sources for generating shock waves. Therefore, an effective treatment plan for one machine may not transfer to a different machine.

As a result, SWL success rates are significantly different among physicians. Table 2 shows the percentiles of success rates of 171 physicians who recorded outcomes in the International Stone Registry, a database of accumulated treatment records for all patients treated within a national network of SWL services provided by Translational Analytics and Statistics, a lithotripsy service provider. Here, treatment success is defined as treated stone fragments ≤ 4 mm that typically pass on their own without further treatment. The top 20% of physicians have success rates higher than 94.3%, while the success rates of the bottom 20% of physicians are lower than 79.1%. Such variation indicates that the inexperience with and subjectivity of SWL treatment could lead to unnecessary damage to patients.

Table 2. Percentiles of treatment success rates.

Percentiles	Treatment success rates, %
Minimum	54.5
10th percentile	74.8
20th percentile	79.1
30th percentile	82.6
40th percentile	84.7
50th percentile	86.6
60th percentile	88.9
70th percentile	91.4
80th percentile	94.3
90th percentile	100
Maximum	100

Machine learning techniques have been applied in the planning process of high-quality personalized treatments, such as radiation therapies [36-39], chemotherapies [40,41] and diabetes treatments [42]. Most machine learning models only take independent vectors as inputs, so they are not suited to the sequential nature of SWL treatment plans. However, recurrent neural networks (RNNs) are naturally suited to temporal sequence inputs. Several variants like long short-term memory (LSTM) [43] and gated recurrent unit [44] have been developed for sequential features and applied to disease diagnosis [45,46]. Following these recent works, we aimed to validate the deep learning approach to generate next SWL treatment steps by learning the practices of top physicians and, based on the deep learning approach, develop a system to automatically produce personalized, unbiased, and consistent SWL treatment plans. The generated treatment plans can help physicians minimize the trial-and-error process and develop evidence-based personalized treatment based on PPC, including patient demographics and stone characteristics. An additional benefit is that this treatment planning framework can be generalized to different machine types, so physicians can easily adapt to new generations of SWL machines.

Methods

Data

To train and evaluate our models, we used a dataset of renal treatments with Storz SLX-T from the International Stone Registry provided by Translational Analytics and Statistics. Each treatment consisted of PPC and several treatment steps (ie, ternaries of a power level, a shock rate, and number of shocks). The power level ranged from 1 to 9. The options for shock rates were 60, 90, 120, and 180 shocks per minute. The maximum number of shocks was typically set at 3000 for renal stones. The PPC in our dataset included patient gender, age, stone location (one-hot encoding), stone size, mean arterial pressure before treatment, anticoagulant use, sedation use, whether multiple stones existed, and whether strapping was applied.

Our deep learning models were trained with the best treatment plans for obtaining the best planning capability. We selected 54 physicians in the top quartile of treatment success rates. These physicians had more than 91.4% treatment success rates. Then, we selected their successful treatment cases with no reported complications, in which they were stone free or had fragments ≤ 4 mm and typically passed on their own without further treatment. We identified 1216 cases in total and assumed these cases are the best practices in SWL treatment planning.

We then built the step dataset from the identified successful cases to train and evaluate the step generation model. We identified steps by power level change or shock rate change and limited the number of shocks to 1000 for each step, a natural step length in previous literature [25]. If more than 1000 shocks were delivered under the same power level and the same shock rate, we broke them into multiple steps with 1000 shocks maximum.

Then, we exhaustively decomposed each case into samples by step for the step generation task, where the ternary of each step was generated by its previous steps and PPC. An n -step treatment case was decomposed into $n - 1$ samples: we used the first i step(s) and PPC as the model inputs and the power level, shock rate, and number of shocks in the $(i + 1)$ th step as the model outputs, where $0 < i < n$. For example, the SWL treatment case in Table 1 that consisted of 10 steps after the last 2300 shocks at power level 8 was split into 3 steps: (1) power level = 8, shock rate = 120, number of shocks = 1000; (2) power level = 8, shock rate = 120, number of shocks = 1000; and (3) power level = 8, shock rate = 120, number of shocks = 300. Then, we decomposed this case into 9 samples: (1) The input is the first step (power level = 1, shock rate = 120, number of shocks = 100) and PPC, and the output is the second step (power level = 2, shock rate = 120, number of shocks = 100); (2) the input is the first 2 steps and PPC, and the output is the third step (power level = 3, shock rate = 120, number of shocks = 100); ...; and (9) the input is the first 9 steps and PPC, and the output is the last step (power level = 8, shock rate = 120, number of shocks = 300).

At last, we constructed 8583 samples for step generation. We randomly chose 90% of the samples for model training and used

the remaining samples for validation. In the data split, we enforced that samples from the same treatment case were only contained within the same split.

Deep Learning for Step Generation

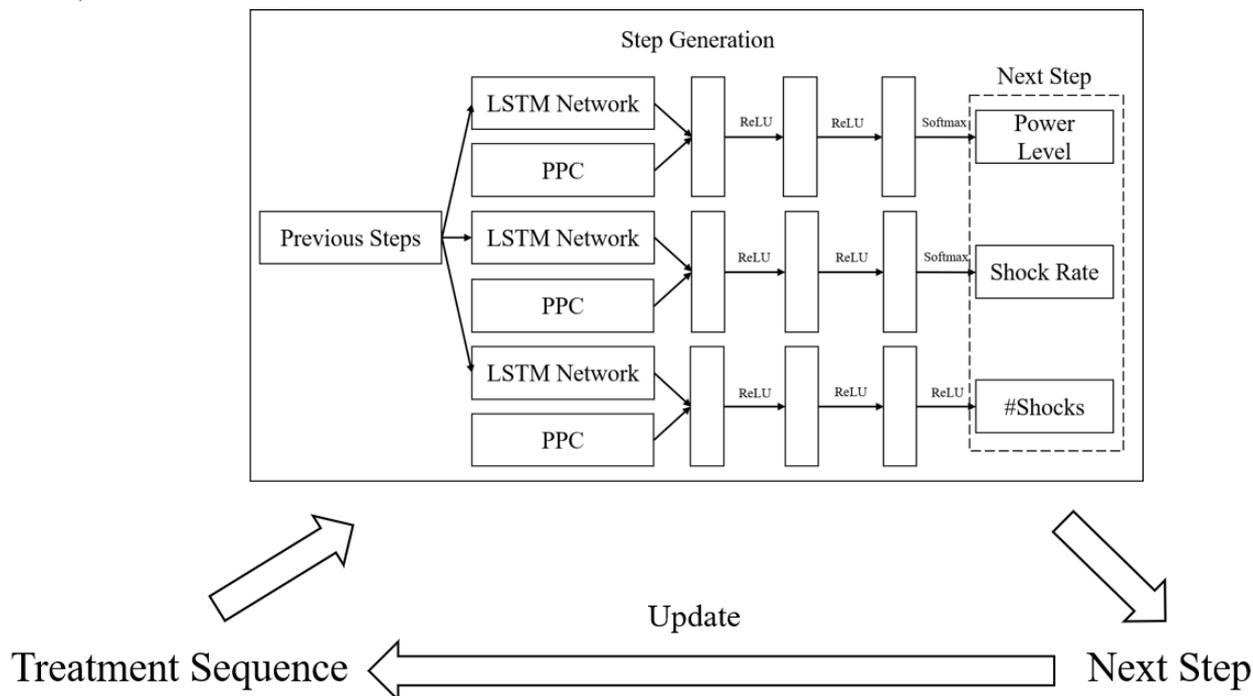
We first built deep neural networks to separately generate power levels, shock rates, and numbers of shocks for the next steps, given previous steps and PPC (Figure 1). Most off-the-shelf machine learning models only take inputs represented as independent vectors rather than a sequence of previous steps. However, RNNs are naturally suited to temporal sequence inputs, so we adopted an RNN variant, the LSTM model [43], which can keep track of arbitrary long-term dependencies in

the input sequences, to encode the treatment sequences to vectors. More specifically, assume the i -th step is encoded as a vector x_i , then the LSTM model is defined iteratively as follows:

$$\begin{aligned}
 f_i &= \sigma(W_f x_i + U_f h_{i-1} + b_f) \\
 g_i &= \sigma(W_g x_i + U_g h_{i-1} + b_g) \\
 o_i &= \sigma(W_o x_i + U_o h_{i-1} + b_o) \\
 c_i &= f_i \circ c_{i-1} + g_i \circ \tanh(W_c x_i + U_c h_{i-1} + b_c) \\
 h_i &= o_i \circ \tanh(c_i)
 \end{aligned}$$

where the initial values c_0 and h_0 are zero vectors, \circ denotes the element-wise product, σ is the sigmoid function, and h_i is the representation of the first i treatment steps.

Figure 1. The framework for automated shock wave lithotripsy (SWL) treatment planning. LSTM: long short-term memory; PPC: preoperative patient characteristics; ReLU: rectifier linear unit.



Then, the encoded previous steps were concatenated to PPC vectors and fed to deep neural networks. In our implementation, we used 2 fully connected layers with a rectifier linear unit (ReLU) function as activation functions, because ReLU functions are nonsaturated and make the model less likely to overfit [47]. At last, we used different classifiers or regressors to generate power levels, shock rates and shock numbers. The formula are as follows.

$$\begin{aligned}
 q &= \text{ReLU}(W_2 \cdot \text{ReLU}(W_0 h_n + W_1 p + b_1) + b_2) \\
 \text{PowerLevel} &= \text{softmax}(W_3 q + b_3) \\
 \text{ShockRate} &= \text{softmax}(W_4 q + b_4) \\
 \text{ShockNumber} &= \text{ReLU}(W_5 q + b_5)
 \end{aligned}$$

where h_n is the n previous steps encoded by LSTM, and p denotes the PPC vector. The classifiers at the end of the networks were softmax functions for generating power level and shock rate because they are categorical, and we used categorical cross-entropy as the loss functions; for shock number generation in which the output is an integer, we used ReLU as the regressor and mean squared error (MSE) as the loss function.

For all the deep neural networks, we chose the Adam SGD optimizer [48,49] in model training.

Statistical Analysis

We hypothesized that the deep learning approach is comparable to the treatment practices of top physicians and that it outperforms machine learning models which do not take treatment sequences as inputs. Thus, we compared the performance of the deep learning model and other up-to-date machine learning models.

Three classical machine learning approaches were selected as baselines for generating power level, shock rate, and number of shocks, respectively. We used logistic regression, random forest classifier (RFC), and support vector classifier (SVC) as the baseline models for power level generation and shock rate generation. We chose linear regression, random forest regression (RFR), and support vector regression (SVR) as the baseline models to generate the number of shocks. As these baseline models could not be fed with sequential data directly, the features for the baseline models were (1) the average power level, average shock rate, and average number of shocks in

previous steps; (2) the power level, shock rate, and number of shocks in the last step; and (3) PPC.

We trained the deep learning models and baseline models with 90% of the samples. Then, we validated them with the remaining samples and calculated evaluation metrics. In the multiclass tasks of power level generation and shock rate generation, we used accuracy, macro-averaged precision, macro-averaged recall, and macro-averaged F1 score as the evaluation metrics [50,51]. Accuracy was defined as a ratio of correctly generated observations to the total observations. Suppose the number of categories is n and the confusion matrix of a classifier is a $n \times n$ matrix C , where C_{ij} is the number of samples that is labeled as i but generated as j , then the accuracy is defined as

$$Accuracy = \frac{\sum_k C_{kk}}{\sum_{i,j} C_{ij}}$$

The precision and recall of category k are defined as

$$Precision_k = \frac{C_{kk}}{\sum_i C_{ik}}$$

$$Recall_k = \frac{C_{kk}}{\sum_j C_{kj}}$$

Macro-averaged precision and recall are the average of precisions and recalls for all categories:

$$Precision = \frac{1}{n} \sum_k Precision_k$$

$$Recall = \frac{1}{n} \sum_k Recall_k$$

The F1 score of category k is defined as the harmonic mean of precision and recall of category k

$$F1_k = 2 \cdot \frac{Precision_k \cdot Recall_k}{Precision_k + Recall_k}$$

and macro-averaged F1 score is defined as the average of F1 scores for all categories:

Table 3. Model performance in power level generation.

Model	Accuracy	Precision	Recall	F1	t statistic	P value
Deep learning	0.988	0.980	0.980	0.980	0.707	.480
Logistic regression	0.974	0.964	0.964	0.964	1.257	.209
RFC ^a	0.708	0.823	0.859	0.803	4.976	<.001
SVC ^b	0.981	0.969	0.976	0.972	2.205	.028

^aRFC: random forest classifier.

^bSVC: support vector classifier.

$$F1 = \frac{1}{n} \sum_k F1_k$$

Because the number of shocks is an integer, we used the root mean squared error (RMSE) and mean absolute error (MAE) as the metrics to evaluate the models generating the number of shocks and to measure the average magnitude of errors. At last, we conducted paired t test to detect the difference between treatment steps generated by machine learning models and treatment practices of top physicians.

Results

The deep learning models generated high-quality treatment steps and outperformed the baselines, as summarized in Tables 3-5. In power level generation (Table 3), the accuracy of the deep learning model was 0.988, and the precision, recall, and F1 scores were all 0.980. The best baseline was the SVC, for which the accuracy was 0.981, precision was 0.969, recall was 0.976, and F1 score was 0.972, lower than the performance of the deep learning model. For shock rate generation (Table 4), our model achieved an F1 score of 0.960 along with an accuracy of 0.981, precision of 0.963, and recall of 0.957. Among the baseline models, the logistic regression performed the best in accuracy and precision, at 0.978 and 0.932, respectively, while the RFC had the best recall and F1 score, at 0.986 and 0.956, respectively. Though the recall of the RFC and the logistic regression was better than that of the deep learning model, the accuracy, precision, and F1 score of our proposed model outperformed all the baseline models. The RMSE of the generation of the number of shocks (Table 5) by the deep learning model was 207, about 19% less than the best baseline model RFR. The MAE of the deep learning model was 121, about 23% less than the best baseline model.

Table 4. Model performance in shock rate generation.

Model	Accuracy	Precision	Recall	F1	<i>t</i> statistic	<i>P</i> value
Deep learning	0.981	0.963	0.957	0.960	0.277	.782
Logistic regression	0.978	0.932	0.960	0.945	2.331	.020
RFC ^a	0.952	0.930	0.986	0.956	2.064	.039
SVC ^b	0.976	0.926	0.956	0.939	2.510	.012

^aRFC: random forest classifier.

^bSVC: support vector classifier.

Table 5. Model performance in shock number generation.

Model	RMSE ^a	MAE ^b	<i>t</i> statistic	<i>P</i> value
Deep learning	207	121	0.350	.727
Linear regression	265	206	0.917	.359
RFR ^c	255	158	0.628	.530
SVR ^d	350	173	9.427	<.001

^aRMSE: root mean squared error.

^bMAE: mean absolute error.

^cRFR: random forest regression.

^dSVR: support vector regression.

The analysis also tested the difference between the generated step and the ground truth. In the paired *t* test result, there was no evidence indicating a difference between the generated steps of the deep learning model and treatment steps planned by top physicians, while the outputs of some baseline models significantly deviated from the ground truth. The power levels generated by the RFC and SVC, the shock rates generated by all the baseline models, and the numbers of shocks generated by the SVR were significantly different from the treatment steps in the successful SWL cases of top physicians.

Furthermore, we analyzed the performance of the deep learning models on samples of various treatment sequence lengths to gain a better understanding of how the treatment sequence information could aid decision making. We partitioned the validation dataset into 9 sets by the number of previous treatment steps and summarized the validation results in Tables 6-8. As shown in Table 6, the deep learning model was able to perfectly generate power levels when previous treatment steps were fewer than 6. As the number of previous treatment steps increases, the treatment becomes more complicated and leads to lower performance of power level generation by the deep learning model. The deep learning model reached the lowest

accuracy (accuracy = 0.875) and lowest recall (recall = 0.500) in samples containing 9 previous treatment steps and the lowest precision (precision = 0.873) and lowest F1 score (F1 = 0.888) in samples containing 8 previous treatment steps. Similarly, the deep learning model generated highly accurate shock rates in samples with previous treatment steps fewer than 6; the model reached the lowest accuracy (accuracy = 0.889), lowest precision (precision = 0.857), and lowest recall (recall = 0.631) in samples containing 7 previous treatment steps and the lowest F1 score (F1 = 0.861) in samples containing 6 previous treatment steps (Table 7). For the performance of generating the number of shocks (Table 8), the RMSE and MAE generally increased as the number of previous treatment steps increased, and the maximum number of errors appeared in samples with 5 previous treatment steps (RMSE = 365; MAE = 310). The results show the excellent performance of deep learning models in step generation in the first 4-6 steps, where most successful cases end. It reflects the reliability of deep learning models in aiding treatment decision making. Longer treatment lengths typically indicate treatment difficulties; even our deep learning models cannot generate treatment steps with high accuracy in these rare cases.

Table 6. Power level generation performance in samples containing different numbers of previous treatment steps.

Number of previous treatment steps	Accuracy	Precision	Recall	F1
1	1.000	1.000	1.000	1.000
2	1.000	1.000	1.000	1.000
3	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1.000
5	1.000	1.000	1.000	1.000
6	0.983	0.980	0.980	0.980
7	0.926	0.915	0.939	0.925
8	0.889	0.873	0.914	0.888
9	0.875	0.875	0.500	0.933

Table 7. Shock rate generation performance in samples containing different numbers of previous treatment steps.

Number of previous treatment steps	Accuracy	Precision	Recall	F1
1	1.000	1.000	1.000	1.000
2	0.992	0.997	0.972	0.984
3	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1.000
5	0.992	0.997	0.974	0.985
6	0.975	0.976	0.802	0.861
7	0.889	0.857	0.631	0.888
8	0.917	0.864	0.642	0.902
9	1.000	1.000	1.000	1.000

Table 8. Performance of the generation of the number of shocks in samples containing different numbers of previous treatment steps.

Number of previous treatment steps	RMSE ^a	MAE ^b
1	31	25
2	32	17
3	34	24
4	139	60
5	365	310
6	317	233
7	273	190
8	275	242
9	99	76

^aRMSE: root mean squared error.

^bMAE: mean absolute error.

The validation showed that the capability of the deep learning model for step generation is on par with that of top physicians. Based on the high-quality step generation, we generated treatment plans by iteratively generating steps with the trained models (Figure 1). We started from an empty treatment sequence. We fed PPC and the current treatment sequence into the step generation model. The generated next step was then added to the current treatment sequence. We repeated such a process until the total number of shocks reached the upper limit.

If a physician confirms stone fragmentation via x-ray before reaching the maximum limit of the number of shocks, they can stop immediately; if the number of shocks reaches the maximum limit, the physician has to stop for risk control. Thus, the generated treatment sequence is enough to guide practice. The specifications of individual lithotripters limit the maximum number of shocks per session to 2000–4500 [4], and for the majority of treatments of upper ureteral and renal stones, the range is 2000–3500 [52]. We used 3000 as the upper limit in

our implementation, which is a typical shock limit in renal stone treatment practices and can be adjusted according to shock wave generating machines.

Discussion

Principal Findings

Previous literature has shown a series of work on standardizing SWL treatment [2,53]; however, energy delivery is still controversial and unclear [4,6], relying on physicians' subjective judgement. Manual treatment design is significantly affected by nonstandardizable radiographic appearance of stones, bias to a low power level for fear of complications, and preconceived expectations. Our study utilized deep learning to generate treatment steps and developed a framework for automated SWL treatment planning.

The analysis results revealed that deep learning models for treatment step generation effectively learn from SWL treatment plans and achieve the step generation capability of top physicians. The performance comparison indicated that utilization of a previous treatment sequence in deep learning improves the quality of generated steps. By iteratively generating treatment steps, our automated planning framework can avoid human biases and generate personalized, high-quality, and consistent SWL treatment plans based on PPC, including patient demographics and stone characteristics. With the help of these automatically generated treatment plans, physicians can minimize the trial-and-error process and implement evidence-based personalized treatment. This framework can be

generalized to different machine types, so physicians can easily adapt to new generations of SWL machines.

Limitations

Our proposed model only learns and imitates the best practices, but cannot perform better than them. Even the best physician cannot plan successful SWL treatment plans for all cases, so successful difficult cases, including those requiring long treatment sequences, are rare for model training. Therefore, our model may be good at planning easier cases, but less adept in rare difficult cases, similar to physicians' actual practice. As the treatment cases, especially successful difficult cases, accumulate, our model is likely to gain an expert-level planning capability to handle difficult cases.

Due to data limitations, we were only able to consider a small set of patient demographics and stone characteristics. However, our framework can be easily extended to utilize a larger set of parameters than has previously been used. Moreover, the data are retrospective. Therefore, clinical studies are warranted to confirm the effectiveness and efficiency of this framework.

Conclusions

To the best of our knowledge, our framework is the first effort to implement automated planning of SWL treatment via deep learning. Its assistance for inexperienced urologists in designing SWL treatment plans is useful in both SWL treatment planning and physician training. While the applications of machine learning in diagnosis are becoming more mature, few studies exist in automated treatment plan generation. Our approach is a step forward in exerting the potential of machine learning in medical sciences.

Authors' Contributions

All authors designed the study. RGNS provided the data. ZC and DDZ developed the deep learning model for treatment plan generation. ZC implemented and evaluated the deep learning model and drafted the manuscript. All authors revised the manuscript.

Conflicts of Interest

RGNS was an employee of Translational Analytics and Statistics. BDH is a consultant for NextMed Management Services. The remaining authors declare no conflicts of interest.

References

1. Chaussy C, Brendel W, Schmiedt E. Extracorporeally induced destruction of kidney stones by shock waves. *Lancet* 1980 Dec 13;2(8207):1265-1268. [doi: [10.1016/s0140-6736\(80\)92335-1](https://doi.org/10.1016/s0140-6736(80)92335-1)] [Medline: [6108446](https://pubmed.ncbi.nlm.nih.gov/6108446/)]
2. Assimos D, Krambeck A, Miller NL, Monga M, Murad MH, Nelson CP, et al. Surgical Management of Stones: American Urological Association/Endourological Society Guideline, PART I. *J Urol* 2016 Oct;196(4):1153-1160. [doi: [10.1016/j.juro.2016.05.090](https://doi.org/10.1016/j.juro.2016.05.090)] [Medline: [27238616](https://pubmed.ncbi.nlm.nih.gov/27238616/)]
3. Al-Marhoon MS, Shareef O, Al-Habsi IS, Al Balushi AS, Mathew J, Venkiteswaran KP. Extracorporeal Shock-wave Lithotripsy Success Rate and Complications: Initial Experience at Sultan Qaboos University Hospital. *Oman Med J* 2013 Jul;28(4):255-259 [FREE Full text] [doi: [10.5001/omj.2013.72](https://doi.org/10.5001/omj.2013.72)] [Medline: [23904918](https://pubmed.ncbi.nlm.nih.gov/23904918/)]
4. Reynolds LF, Krocak T, Pace KT. Indications and contraindications for shock wave lithotripsy and how to improve outcomes. *Asian J Urol* 2018 Oct;5(4):256-263 [FREE Full text] [doi: [10.1016/j.ajur.2018.08.006](https://doi.org/10.1016/j.ajur.2018.08.006)] [Medline: [30364729](https://pubmed.ncbi.nlm.nih.gov/30364729/)]
5. Skolarikos A, Alivizatos G, de la Rosette J. Extracorporeal shock wave lithotripsy 25 years later: complications and their prevention. *Eur Urol* 2006 Nov;50(5):981-90; discussion 990. [doi: [10.1016/j.eururo.2006.01.045](https://doi.org/10.1016/j.eururo.2006.01.045)] [Medline: [16481097](https://pubmed.ncbi.nlm.nih.gov/16481097/)]
6. McClain PD, Lange JN, Assimos DG. Optimizing shock wave lithotripsy: a comprehensive review. *Rev Urol* 2013;15(2):49-60 [FREE Full text] [Medline: [24082843](https://pubmed.ncbi.nlm.nih.gov/24082843/)]
7. Pareek G, Armenakas NA, Panagopoulos G, Bruno JJ, Fracchia JA. Extracorporeal shock wave lithotripsy success based on body mass index and Hounsfield units. *Urology* 2005 Jan;65(1):33-36. [doi: [10.1016/j.urology.2004.08.004](https://doi.org/10.1016/j.urology.2004.08.004)] [Medline: [15667858](https://pubmed.ncbi.nlm.nih.gov/15667858/)]

8. Perks AE, Schuler TD, Lee J, Ghiculete D, Chung D, D'A Honey RJ, et al. Stone attenuation and skin-to-stone distance on computed tomography predicts for stone fragmentation by shock wave lithotripsy. *Urology* 2008 Oct;72(4):765-769. [doi: [10.1016/j.urology.2008.05.046](https://doi.org/10.1016/j.urology.2008.05.046)] [Medline: [18674803](https://pubmed.ncbi.nlm.nih.gov/18674803/)]
9. Hatiboglu G, Popeneciu V, Kurosch M, Huber J, Pahernik S, Pfitzenmaier J, et al. Prognostic variables for shockwave lithotripsy (SWL) treatment success: no impact of body mass index (BMI) using a third generation lithotripter. *BJU Int* 2011 Oct;108(7):1192-1197 [FREE Full text] [doi: [10.1111/j.1464-410X.2010.10007.x](https://doi.org/10.1111/j.1464-410X.2010.10007.x)] [Medline: [21342413](https://pubmed.ncbi.nlm.nih.gov/21342413/)]
10. Wiesenthal JD, Ghiculete D, Ray AA, Honey RJD, Pace KT. A clinical nomogram to predict the successful shock wave lithotripsy of renal and ureteral calculi. *J Urol* 2011 Aug;186(2):556-562. [doi: [10.1016/j.juro.2011.03.109](https://doi.org/10.1016/j.juro.2011.03.109)] [Medline: [21684557](https://pubmed.ncbi.nlm.nih.gov/21684557/)]
11. Patel T, Kozakowski K, Hruba G, Gupta M. Skin to stone distance is an independent predictor of stone-free status following shockwave lithotripsy. *J Endourol* 2009 Sep;23(9):1383-1385. [doi: [10.1089/end.2009.0394](https://doi.org/10.1089/end.2009.0394)] [Medline: [19694526](https://pubmed.ncbi.nlm.nih.gov/19694526/)]
12. Dretler SP. Special article: calculus breakability--fragility and durability. *J Endourol* 1994 Feb;8(1):1-3. [doi: [10.1089/end.1994.8.1](https://doi.org/10.1089/end.1994.8.1)] [Medline: [8186775](https://pubmed.ncbi.nlm.nih.gov/8186775/)]
13. Ringdén I, Tiselius H. Composition and clinically determined hardness of urinary tract stones. *Scand J Urol Nephrol* 2007;41(4):316-323. [doi: [10.1080/00365590601154551](https://doi.org/10.1080/00365590601154551)] [Medline: [17763224](https://pubmed.ncbi.nlm.nih.gov/17763224/)]
14. Ouzaid I, Al-qahtani S, Dominique S, Hupertan V, Fernandez P, Hermieu J, et al. A 970 Hounsfield units (HU) threshold of kidney stone density on non-contrast computed tomography (NCCT) improves patients' selection for extracorporeal shockwave lithotripsy (ESWL): evidence from a prospective study. *BJU Int* 2012 Dec;110(11 Pt B):E438-E442. [doi: [10.1111/j.1464-410X.2012.10964.x](https://doi.org/10.1111/j.1464-410X.2012.10964.x)] [Medline: [22372937](https://pubmed.ncbi.nlm.nih.gov/22372937/)]
15. El-Nahas AR, El-Assmy AM, Mansour O, Sheir KZ. A prospective multivariate analysis of factors predicting stone disintegration by extracorporeal shock wave lithotripsy: the value of high-resolution noncontrast computed tomography. *Eur Urol* 2007 Jun;51(6):1688-93; discussion 1693. [doi: [10.1016/j.eururo.2006.11.048](https://doi.org/10.1016/j.eururo.2006.11.048)] [Medline: [17161522](https://pubmed.ncbi.nlm.nih.gov/17161522/)]
16. Joseph P, Mandal A, Singh S, Mandal P, Sankhwar S, Sharma S. Computerized Tomography Attenuation Value of Renal Calculus: Can It Predict Successful Fragmentation of the Calculus by Extracorporeal Shock Wave Lithotripsy? A Preliminary Study. *Journal of Urology* 2002 May;167(5):1968-1971. [doi: [10.1016/s0022-5347\(05\)65064-1](https://doi.org/10.1016/s0022-5347(05)65064-1)] [Medline: [11956419](https://pubmed.ncbi.nlm.nih.gov/11956419/)]
17. Abdelhamid M, Mosharafa AA, Ibrahim H, Selim HM, Hamed M, Elghoneimy MN, et al. A Prospective Evaluation of High-Resolution CT Parameters in Predicting Extracorporeal Shockwave Lithotripsy Success for Upper Urinary Tract Calculi. *J Endourol* 2016 Nov;30(11):1227-1232. [doi: [10.1089/end.2016.0364](https://doi.org/10.1089/end.2016.0364)] [Medline: [27597174](https://pubmed.ncbi.nlm.nih.gov/27597174/)]
18. Yamashita S, Kohjimoto Y, Iguchi T, Nishizawa S, Iba A, Kikkawa K, et al. Variation Coefficient of Stone Density: A Novel Predictor of the Outcome of Extracorporeal Shockwave Lithotripsy. *J Endourol* 2017 Apr;31(4):384-390. [doi: [10.1089/end.2016.0719](https://doi.org/10.1089/end.2016.0719)] [Medline: [28052698](https://pubmed.ncbi.nlm.nih.gov/28052698/)]
19. Bohris C, Stief CG, Strittmatter F. Improvement of SWL Efficacy: Reduction of the Respiration-Induced Kidney Motion by Using an Abdominal Compression Plate. *J Endourol* 2016 Apr;30(4):411-416. [doi: [10.1089/end.2015.0681](https://doi.org/10.1089/end.2015.0681)] [Medline: [26558296](https://pubmed.ncbi.nlm.nih.gov/26558296/)]
20. Honey RJ, Healy M, Yeung M, Psihramis KE, Jewett MA. The Use of an Abdominal Compression Belt to Reduce Stone Movement During Extracorporeal Shock Wave Lithotripsy. *Journal of Urology* 1992 Sep;148(3 Part 2):1034-1035. [doi: [10.1016/s0022-5347\(17\)36808-8](https://doi.org/10.1016/s0022-5347(17)36808-8)] [Medline: [1507324](https://pubmed.ncbi.nlm.nih.gov/1507324/)]
21. Pishchalnikov YA, Neucks JS, VonDerHaar RJ, Pishchalnikova IV, Williams JC, McAteer JA. Air pockets trapped during routine coupling in dry head lithotripsy can significantly decrease the delivery of shock wave energy. *J Urol* 2006 Dec;176(6 Pt 1):2706-2710 [FREE Full text] [doi: [10.1016/j.juro.2006.07.149](https://doi.org/10.1016/j.juro.2006.07.149)] [Medline: [17085200](https://pubmed.ncbi.nlm.nih.gov/17085200/)]
22. Jain A, Shah TK. Effect of air bubbles in the coupling medium on efficacy of extracorporeal shock wave lithotripsy. *Eur Urol* 2007 Jun;51(6):1680-6; discussion 1686. [doi: [10.1016/j.eururo.2006.10.049](https://doi.org/10.1016/j.eururo.2006.10.049)] [Medline: [17112655](https://pubmed.ncbi.nlm.nih.gov/17112655/)]
23. Rabah DM, Mabrouki MS, Farhat KH, Seida MA, Arafa MA, Talic RF. Comparison of escalating, constant, and reduction energy output in ESWL for renal stones: multi-arm prospective randomized study. *Urolithiasis* 2017 Jun;45(3):311-316. [doi: [10.1007/s00240-016-0912-7](https://doi.org/10.1007/s00240-016-0912-7)] [Medline: [27687681](https://pubmed.ncbi.nlm.nih.gov/27687681/)]
24. Connors BA, Evan AP, Handa RK, Blomgren PM, Johnson CD, Liu Z, et al. Using 300 Pretreatment Shock Waves in a Voltage Ramping Protocol Can Significantly Reduce Tissue Injury During Extracorporeal Shock Wave Lithotripsy. *J Endourol* 2016 Sep;30(9):1004-1008 [FREE Full text] [doi: [10.1089/end.2016.0087](https://doi.org/10.1089/end.2016.0087)] [Medline: [27307070](https://pubmed.ncbi.nlm.nih.gov/27307070/)]
25. Lambert EH, Walsh R, Moreno MW, Gupta M. Effect of escalating versus fixed voltage treatment on stone comminution and renal injury during extracorporeal shock wave lithotripsy: a prospective randomized trial. *J Urol* 2010 Feb;183(2):580-584. [doi: [10.1016/j.juro.2009.10.025](https://doi.org/10.1016/j.juro.2009.10.025)] [Medline: [20018316](https://pubmed.ncbi.nlm.nih.gov/20018316/)]
26. Handa RK, McAteer JA, Connors BA, Liu Z, Lingeman JE, Evan AP. Optimising an escalating shockwave amplitude treatment strategy to protect the kidney from injury during shockwave lithotripsy. *BJU Int* 2012 Dec;110(11 Pt C):E1041-E1047 [FREE Full text] [doi: [10.1111/j.1464-410X.2012.11207.x](https://doi.org/10.1111/j.1464-410X.2012.11207.x)] [Medline: [22612388](https://pubmed.ncbi.nlm.nih.gov/22612388/)]
27. McAteer JA, Evan AP, Williams JC, Lingeman JE. Treatment protocols to reduce renal injury during shock wave lithotripsy. *Curr Opin Urol* 2009 Mar;19(2):192-195 [FREE Full text] [doi: [10.1097/mou.0b013e32831e16e3](https://doi.org/10.1097/mou.0b013e32831e16e3)] [Medline: [19195131](https://pubmed.ncbi.nlm.nih.gov/19195131/)]
28. Honey RJD, Ray AA, Ghiculete D, University of Toronto Lithotripsy Associates, Pace KT. Shock wave lithotripsy: a randomized, double-blind trial to compare immediate versus delayed voltage escalation. *Urology* 2010 Jan;75(1):38-43. [doi: [10.1016/j.urology.2008.12.070](https://doi.org/10.1016/j.urology.2008.12.070)] [Medline: [19896176](https://pubmed.ncbi.nlm.nih.gov/19896176/)]

29. Pishchalnikov YA, McAteer JA, Williams JC, Pishchalnikova IV, Vonderhaar RJ. Why stones break better at slow shockwave rates than at fast rates: in vitro study with a research electrohydraulic lithotripter. *J Endourol* 2006 Aug;20(8):537-541 [FREE Full text] [doi: [10.1089/end.2006.20.537](https://doi.org/10.1089/end.2006.20.537)] [Medline: [16903810](https://pubmed.ncbi.nlm.nih.gov/16903810/)]
30. Pishchalnikov YA, McAteer JA, Williams JC. Effect of firing rate on the performance of shock wave lithotriptors. *BJU Int* 2008 Dec;102(11):1681-1686 [FREE Full text] [doi: [10.1111/j.1464-410X.2008.07896.x](https://doi.org/10.1111/j.1464-410X.2008.07896.x)] [Medline: [18710450](https://pubmed.ncbi.nlm.nih.gov/18710450/)]
31. Kang DH, Cho KS, Ham WS, Lee H, Kwon JK, Choi YD, et al. Comparison of High, Intermediate, and Low Frequency Shock Wave Lithotripsy for Urinary Tract Stone Disease: Systematic Review and Network Meta-Analysis. *PLoS One* 2016;11(7):e0158661 [FREE Full text] [doi: [10.1371/journal.pone.0158661](https://doi.org/10.1371/journal.pone.0158661)] [Medline: [27387279](https://pubmed.ncbi.nlm.nih.gov/27387279/)]
32. Delius M, Jordan M, Eizenhoefer H, Marlinghaus E, Heine G, Liebich HG, et al. Biological effects of shock waves: Kidney haemorrhage by shock waves in dogs—Administration rate dependence. *Ultrasound in Medicine & Biology* 1988 Jan;14(8):689-694. [doi: [10.1016/0301-5629\(88\)90025-7](https://doi.org/10.1016/0301-5629(88)90025-7)] [Medline: [3212839](https://pubmed.ncbi.nlm.nih.gov/3212839/)]
33. Willis LR, Evan AP, Connors BA, Shao Y, Blomgren PM, Pratt JH, et al. Shockwave lithotripsy: dose-related effects on renal structure, hemodynamics, and tubular function. *J Endourol* 2005;19(1):90-101. [doi: [10.1089/end.2005.19.90](https://doi.org/10.1089/end.2005.19.90)] [Medline: [15735392](https://pubmed.ncbi.nlm.nih.gov/15735392/)]
34. Mobley TB, Myers DA, Grine WB, Jenkins JM, Jordan WR. Low Energy Lithotripsy with the Lithostar: Treatment Results with 19,962 Renal and Ureteral Calculi. *Journal of Urology* 1993 Jun;149(6):1419-1424. [doi: [10.1016/s0022-5347\(17\)36404-2](https://doi.org/10.1016/s0022-5347(17)36404-2)] [Medline: [8501779](https://pubmed.ncbi.nlm.nih.gov/8501779/)]
35. Skuginna V, Nguyen DP, Seiler R, Kiss B, Thalmann GN, Roth B. Does Stepwise Voltage Ramping Protect the Kidney from Injury During Extracorporeal Shockwave Lithotripsy? Results of a Prospective Randomized Trial. *Eur Urol* 2016 Feb;69(2):267-273. [doi: [10.1016/j.eururo.2015.06.017](https://doi.org/10.1016/j.eururo.2015.06.017)] [Medline: [26119561](https://pubmed.ncbi.nlm.nih.gov/26119561/)]
36. Fan J, Wang J, Chen Z, Hu C, Zhang Z, Hu W. Automatic treatment planning based on three-dimensional dose distribution predicted from deep learning technique. *Med Phys* 2019 Jan;46(1):370-381. [doi: [10.1002/mp.13271](https://doi.org/10.1002/mp.13271)] [Medline: [30383300](https://pubmed.ncbi.nlm.nih.gov/30383300/)]
37. Smith WP, Kim M, Holdsworth C, Liao J, Phillips MH. Personalized treatment planning with a model of radiation therapy outcomes for use in multiobjective optimization of IMRT plans for prostate cancer. *Radiat Oncol* 2016 Mar 11;11:38 [FREE Full text] [doi: [10.1186/s13014-016-0609-7](https://doi.org/10.1186/s13014-016-0609-7)] [Medline: [26968687](https://pubmed.ncbi.nlm.nih.gov/26968687/)]
38. Nicolae A, Morton G, Chung H, Loblaw A, Jain S, Mitchell D, et al. Evaluation of a Machine-Learning Algorithm for Treatment Planning in Prostate Low-Dose-Rate Brachytherapy. *Int J Radiat Oncol Biol Phys* 2017 Mar 15;97(4):822-829. [doi: [10.1016/j.ijrobp.2016.11.036](https://doi.org/10.1016/j.ijrobp.2016.11.036)] [Medline: [28244419](https://pubmed.ncbi.nlm.nih.gov/28244419/)]
39. Mak RH, Endres MG, Paik JH, Sergeev RA, Aerts H, Williams CL, et al. Use of Crowd Innovation to Develop an Artificial Intelligence-Based Solution for Radiation Therapy Targeting. *JAMA Oncol* 2019 May 01;5(5):654-661 [FREE Full text] [doi: [10.1001/jamaoncol.2019.0159](https://doi.org/10.1001/jamaoncol.2019.0159)] [Medline: [30998808](https://pubmed.ncbi.nlm.nih.gov/30998808/)]
40. Lee S, Celik S, Logsdon BA, Lundberg SM, Martins TJ, Oehler VG, et al. A machine learning approach to integrate big data for precision medicine in acute myeloid leukemia. *Nat Commun* 2018 Jan 03;9(1):42 [FREE Full text] [doi: [10.1038/s41467-017-02465-5](https://doi.org/10.1038/s41467-017-02465-5)] [Medline: [29298978](https://pubmed.ncbi.nlm.nih.gov/29298978/)]
41. Lin H, Wei N, Chou T, Lin C, Lan Y, Chang S, et al. Building personalized treatment plans for early-stage colorectal cancer patients. *Oncotarget* 2017 Feb 21;8(8):13805-13817 [FREE Full text] [doi: [10.18632/oncotarget.14638](https://doi.org/10.18632/oncotarget.14638)] [Medline: [28099153](https://pubmed.ncbi.nlm.nih.gov/28099153/)]
42. Doubleday K, Zhou H, Fu H, Zhou J. An Algorithm for Generating Individualized Treatment Decision Trees and Random Forests. *J Comput Graph Stat* 2018;27(4):849-860 [FREE Full text] [doi: [10.1080/10618600.2018.1451337](https://doi.org/10.1080/10618600.2018.1451337)] [Medline: [32523325](https://pubmed.ncbi.nlm.nih.gov/32523325/)]
43. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput* 1997 Nov 15;9(8):1735-1780. [doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735)] [Medline: [9377276](https://pubmed.ncbi.nlm.nih.gov/9377276/)]
44. Cho K, van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. Cornell University. 2014. URL: <https://arxiv.org/abs/1406.1078> [accessed 2021-04-24]
45. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016 Dec 13;316(22):2402-2410. [doi: [10.1001/jama.2016.17216](https://doi.org/10.1001/jama.2016.17216)] [Medline: [27898976](https://pubmed.ncbi.nlm.nih.gov/27898976/)]
46. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. *J Am Med Inform Assoc* 2017 Mar 01;24(2):361-370 [FREE Full text] [doi: [10.1093/jamia/ocw112](https://doi.org/10.1093/jamia/ocw112)] [Medline: [27521897](https://pubmed.ncbi.nlm.nih.gov/27521897/)]
47. Xu L, Choy CS, Li YW. Deep sparse rectifier neural networks for speech denoising. 2016 Presented at: IEEE International Workshop on Acoustic Signal Enhancement (IWAENC); September 13-16, 2016; Xi'an, China. [doi: [10.1109/iwaenc.2016.7602891](https://doi.org/10.1109/iwaenc.2016.7602891)]
48. Kingma DP, Ba J. Adam: A method for stochastic optimization. Cornell University. 2014. URL: <https://arxiv.org/abs/1412.6980> [accessed 2021-04-24]
49. Reddi SJ, Kale S, Kumar S. On the convergence of Adam and beyond. Cornell University. 2019. URL: <https://arxiv.org/abs/1904.09237> [accessed 2021-04-24]
50. Manning CD, Raghavan P, Schütze H. Natural language engineering. In: Introduction to information retrieval. Cambridge, MA: Cambridge University Press; 2010:100-103.

51. Narasimhan H, Pan W, Kar P, Protopapas P, Ramaswamy HG. Optimizing the multiclass F-measure via biconcave programming. 2017 Presented at: IEEE 16th International Conference on Data Mining (ICDM); December 12-15, 2016; Barcelona, Spain. [doi: [10.1109/icdm.2016.0143](https://doi.org/10.1109/icdm.2016.0143)]
52. Rassweiler JJ, Knoll T, Köhrmann KU, McAteer JA, Lingeman JE, Cleveland RO, et al. Shock wave technology and application: an update. *Eur Urol* 2011 May;59(5):784-796 [FREE Full text] [doi: [10.1016/j.eururo.2011.02.033](https://doi.org/10.1016/j.eururo.2011.02.033)] [Medline: [21354696](https://pubmed.ncbi.nlm.nih.gov/21354696/)]
53. Türk C, Petřík A, Sarica K, Seitz C, Skolarikos A, Straub M, et al. EAU Guidelines on Interventional Treatment for Urolithiasis. *Eur Urol* 2016 Mar;69(3):475-482. [doi: [10.1016/j.eururo.2015.07.041](https://doi.org/10.1016/j.eururo.2015.07.041)] [Medline: [26344917](https://pubmed.ncbi.nlm.nih.gov/26344917/)]

Abbreviations

LSTM: long short-term memory
MAE: mean absolute error
MSE: mean squared error
PPC: preoperative patient characteristics
ReLU: rectifier linear unit
RFC: random forest classifier
RFR: random forest regression
RMSE: root mean squared error
RNN: recurrent neural network
SVC: support vector classifier
SVR: support vector regression
SWL: shock wave lithotripsy

Edited by G Eysenbach; submitted 02.10.20; peer-reviewed by JA Benítez-Andrades, G Lim; comments to author 06.11.20; revised version received 31.12.20; accepted 11.04.21; published 11.05.21

Please cite as:

Chen Z, Zeng DD, Seltzer RGN, Hamilton BD

Automated Generation of Personalized Shock Wave Lithotripsy Protocols: Treatment Planning Using Deep Learning

JMIR Med Inform 2021;9(5):e24721

URL: <https://medinform.jmir.org/2021/5/e24721>

doi: [10.2196/24721](https://doi.org/10.2196/24721)

PMID:

©Zhipeng Chen, Daniel D Zeng, Ryan G N Seltzer, Blake D Hamilton. Originally published in JMIR Medical Informatics (<https://medinform.jmir.org>), 11.05.2021. 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 JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.