RESEARCH

Predicting acupuncture efficacy for functional dyspepsia based on routine clinical features: a machine learning study in the framework of predictive, preventive, and personalized medicine

Tao Yin[1,2] · Hui Zheng[1] · Tingting Ma[3] · Xiaoping Tian[1] · Jing Xu[1] · Ying Li[4] · Lei Lan[1,2] · Mailan Liu[5] · Ruirui Sun[1,2] · Yong Tang[1,6] · Fanrong Liang[1] · Fang Zeng[1,2]

Abstract

Background Acupuncture is safe and effective for functional dyspepsia (FD), while its efficacy varies among individuals. Predicting the response of different FD patients to acupuncture treatment in advance and therefore administering the tailored treatment to the individual is consistent with the principle of predictive, preventive, and personalized medicine (PPPM/3PM). In the current study, the individual efficacy prediction models were developed based on the support vector machine (SVM) algorithm and routine clinical features, aiming to predict the efficacy of acupuncture in treating FD and identify the FD patients who were appropriate to acupuncture treatment.

Methods A total of 745 FD patients were collected from two clinical trials. All the patients received a 4-week acupuncture treatment. Based on the demographic and baseline clinical features of 80% of patients in trial 1, the SVM models were established to predict the acupuncture response and improvements of symptoms and quality of life (QoL) at the end of treatment. Then, the left 20% of patients in trial 1 and 193 patients in trial 2 were respectively applied to evaluate the internal and external generalizations of these models.

Results These models could predict the efficacy of acupuncture successfully. In the internal test set, models achieved an accuracy of 0.773 in predicting acupuncture response and an R^2 of 0.446 and 0.413 in the prediction of QoL and symptoms improvements, respectively. Additionally, these models had well generalization in the independent validation set and could also predict, to a certain extent, the long-term efficacy of acupuncture at the 12-week follow-up. The gender, subtype of disease, and education level were finally identified as the critical predicting features.

Conclusion Based on the SVM algorithm and routine clinical features, this study established the models to predict acupuncture efficacy for FD patients. The prediction models developed accordingly are promising to assist doctors in judging patients' responses to acupuncture in advance, so that they could tailor and adjust acupuncture treatment plans for different patients in a prospective rather than the reactive manner, which could greatly improve the clinical efficacy of acupuncture treatment for FD and save medical expenditures.

Keywords Predictive preventive personalized medicine (PPPM/3PM) · Healthcare · Acupuncture · Functional dyspepsia · Efficacy prediction · Precision medicine · Machine learning · Artificial intelligence · Support vector machine

Abbreviations

AUC	Area under the receiver operating characteristic curve
CAM	Complementary and alternative medicine
EPS	Epigastric pain syndrome
FD	Functional dyspepsia
GLM	Generalized linear model
ML	Machine learning
MSE	Mean squared error
NDSI	Nepean Dyspepsia Symptom Index
NDLQI	Nepean Dyspepsia Life Quality Index

Tao Yin, Hui Zheng and Tingting Ma contributed equally.

✉ Fanrong Liang
 lfr@cdutcm.edu.cn

✉ Fang Zeng
 zengfang@cdutcm.edu.cn

Extended author information available on the last page of the article

PDS	Postprandial distress syndrome
PPPM/3PM	Predictive, preventive, and personalized medicine
QoL	Quality of life
R^2	Coefficient of determination
SVC	Support vector classification
SVM	Support vector machine
SVR	Support vector regression

Introduction

Functional dyspepsia is a worldwide health issue affecting the quality of life of patients

Functional dyspepsia (FD, DD90.3) is a common functional gastrointestinal disease with recurrent symptoms of epigastric pain, epigastric burning, postprandial fullness, and early satiation [1]. According to the main symptoms, FD is sorted into two subtypes, the postprandial distress syndrome (PDS) and the epigastric pain syndrome (EPS) [1]. The PDS is mainly characterized by bothersome postprandial fullness and early satiation, and EPS is dominated by epigastric pain and burning sensation. It was reported that approximately 6.9–17.6% of the population suffered from dyspepsia around the world [2], and 80% of which were unexplained after routine clinical evaluations [3]. As a worldwide health issue, FD significantly affects the psychophysiological health and quality of life (QoL) of patients [4] and brings heavy economic and social burdens [5]. Therefore, searching for safe and efficient therapeutics for FD has been a consensus among patients and gastroenterologists.

Acupuncture is safe and effective for functional dyspepsia, while its efficacy varies among individuals

Acupuncture is a typical individualized therapy [6]. In clinical practice, doctors usually flexibly adjust the acupuncture prescriptions according to the patients' conditions even for the same diseases. As the most widely available complementary and alternative medicine (CAM) therapy, acupuncture was applied to regulate gastrointestinal function for thousands of years in China. Its therapeutic efficacy for FD has also been validated in several high-quality clinical trials [7–10]. For example, our two randomized controlled trials indicated that acupuncture could not only significantly relieve the symptoms but also improve the QoL of FD patients [8, 9], and its efficacy in improving QoL was more pronounced than *Itopride*, a recommended medicine for FD treatment [8]. However, the efficacy of acupuncture was influenced by several endogenous factors, such as subtype of disease [11], and varied among different FD patients.

Besides, a recent study also detected that the baseline sleep condition, anxiety, medications taken, and complications were associated with the relief of symptoms in FD patients after interventions [12]. These findings suggested the significant influences of baseline condition on acupuncture efficacy for FD and implicated two meaningful questions: What kinds of FD patients are more sensitive to acupuncture treatment? And whether the efficacy of acupuncture could be predicted based on patients' baseline characteristics.

Machine learning is a promising approach for acupuncture efficacy predictions in the PPPM/3PM framework

Machine learning (ML) holds the promise to address these above two questions [13] and has been regarded as a powerful tool for predictive, preventive, and personalized medicine (PPPM/3PM) [14]. The PPPM/3PM is an advanced philosophy in healthcare and disease care sectors that enables to predict individual predisposition before the onset of the disease, provide targeted preventive measures, and develop personalized treatment strategies tailored to the individual [15, 16]. In the PPPM/3PM framework, ML algorithms were utilized to select the predictive biomarkers, enabling predictive diagnostics, targeted prevention, and personalization of medical services [17–19]. Among kinds of ML algorithms, the support vector machine (SVM) is the most used in efficacy prediction due to its outstanding performance and strong generalization [20, 21]. Based on the SVM algorithm and baseline functional neuroimaging features, our recent study constructed a prediction model and forecasted the efficacy of acupuncture on FD successfully [22]. However, due to the difficulties in neuroimaging data acquisition and processing, the translational potential of this prediction model was substantially limited. Therefore, it is of great practical importance and value to explore a more accessible approach for the prediction of acupuncture efficacy for FD.

Aims of the study

In the framework of PPPM/3PM, the current study developed three prediction models based on the SVM algorithm and routine clinical features, aiming to predict (1) the treatment response, (2) the improvements of clinical symptoms, and (3) the improvements of QoL in FD patients at the end of treatment and 12-week follow-up. The developed models were expected to assist doctors to forecast patients' responses to acupuncture in advance, thereby preemptively stratifying patients, identifying those appropriate for acupuncture treatment, and ultimately developing personalized treatment tailored to the individual.

Methods

This was a priori planned secondary analysis. Data was from two multi-center randomized controlled trials aiming to determine the efficacy of acupuncture for FD patients [8, 9]. Patients in the first trial were applied to train and internally test the prediction models. Patients in the second trial were served as an independent set to validate the external generalization of the trained models.

Ethical approval and trial registration.

These two trials were approved by the Ethics Committee of Hospital of Chengdu University of Traditional Chinese Medicine and were registered at http://www.clinicaltrial.gov (registered number: NCT00599677 and NCT01671670). The implementation of studies followed the principles of the Declaration of Helsinki.

Participants and sample size

These FD patients in trial 1 were recruited from 3 centers in China during April 2008 and October 2009. FD patients in trial 2 were recruited from 3 hospitals in China from January 2012 to September 2014. All these patients were diagnosed based on Rome III criteria and met the following inclusion criteria: (1) were 18–65 years old, (2) had no structural or biochemical abnormalities to explain their dyspeptic symptoms, and (3) not accompanied with other functional gastrointestinal diseases. All the patients provided written informed consents before entering the trials.

In the first trial [8], 720 FD patients were recruited and randomly assigned into five acupuncture treatment groups and an *Itopride* group. A total of 552 and 542 FD patients in the acupuncture treatment groups were retained at the end of treatment and 12-week follow-up, respectively. In the second trial [9], 200 FD patients were randomly assigned into two acupuncture treatment groups. Finally, 193 and 192 FD patients were retained at the end of treatment and 12-week follow-up, respectively. The number of participants who survived in each stage of these two trials is shown in Fig. 1.

Intervention

All the FD patients received 20 sessions of acupuncture treatment in 4 weeks. The details of acupuncture prescriptions and manipulations could be found in Supplementary 1 and the published articles [8, 9].

Outcome measures

All the FD patients were required to fill in the Nepean Dyspepsia Symptom Index (NDSI) and the Nepean Dyspepsia Life Quality Index (NDLQI) [23] at the baseline, the end of treatment, and the 12-week follow-up, to evaluate the severity of clinical symptoms and QoL of FD patients, respectively. The 15-item NDSI measured the frequency, intensity, and level of discomfort of 15 upper gastrointestinal symptoms over the prior 14 days. The 25-item NDLQI measured the dyspepsia-specific health-related QoL of FD patients from the following four domains: the interference

Fig. 1 The number of participants who survived in each stage
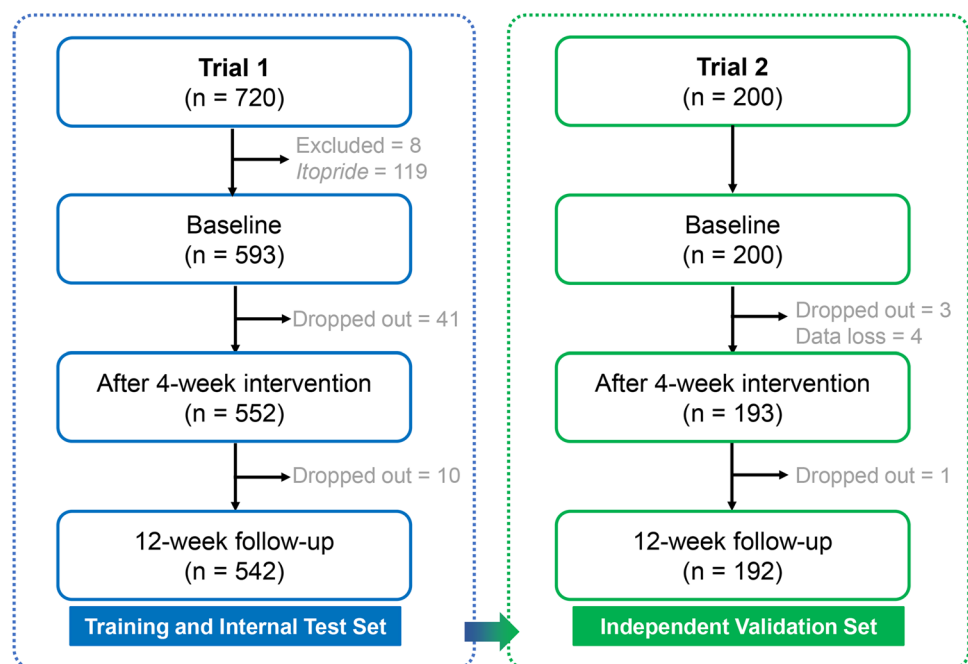
(13 items), the knowledge/control (7 items), the eat/drink (3 items), and the sleep/disturb (2 items).

Prediction of acupuncture efficacy

Machine learning algorithm

The SVM algorithm was utilized to construct the acupuncture efficacy prediction models. Namely, the support vector classification (SVC) was applied to predict responders or non-responders, whereas the support vector regression (SVR) was utilized for the prediction of NDSI and NDLQI improvements. All the SVM analyses were conducted with the LIBSVM toolbox [24] (https://www.csie.ntu.edu.tw/~cjlin/Libsvm) at MATLAB 2017b.

Feature

A total of 28 features were used to construct the prediction models. These features included 5 demographic characteristics (the gender, age, body mass index, education level, and marriage status) and 23 baseline clinical symptoms indicators (the subtype of disease, duration (by month), 15 sub-scores and the total score of the baseline NDSI, and 4 dimensional scores and the mean score of the NDLQI). The education level contained five grades from 1 to 5, indicating the illiterate, primary school, middle school, high school, and college and above. The marriage status included married (indicated with "1") and unmarried (indicated with "0"). The subtype of disease included the PDS (indicated with "1") and the EPS (indicated with "2"), which were determined by the most complained symptoms of patients.

Label

In the SVC models, the labels were set to "1" and "−1". The "1" represented the acupuncture responders, who had an improvement of NDLQI score more than 10 points (the minimally clinically important difference) [25] compared to the baseline. The "−1" represented the acupuncture non-responders, whose NDLQI score was improved less than 10 points. In the SVR model, the improvements of NDSI and NDLQI compared to the baseline were applied as labels.

Prediction models construction and performance evaluation

The prediction models were constructed and internally tested based on the 552 FD patients in trial 1. First, these patients were randomly divided into two groups. Group one was the training set, which contained 80% of patients and was applied to construct the prediction models. Group two was the test set, which included 20% of patients and was used

to evaluate the performance of models. Based on the training set, one linear kernel SVC model and 2 linear kernel SVR models were constructed with the default parameters. Then, these established models were applied to predict the response as well as the improvements of NDSI and NDLQI of patients in the test set. The accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC) were applied to evaluate the performance of the SVC model. The coefficient of determination (R^2) and mean squared error (MSE) were applied to assess the performance of the SVR models. The statistical significance of the prediction results (accuracy and AUC in the SVC model, R^2 and MSE in SVR models) was assessed with the permutation tests [26]. The number of permutations was set to 1000 (statistical efficiency = 0.001).

Independent validation of the prediction models

The 193 patients in trial 2 were applied as an independent validation set to evaluate the external generalization of the prediction models. Namely, the models obtained above were utilized to predict the acupuncture response and the improvements of NDSI and NDLQI at the end of treatment. The performance of prediction and its statistical significance was evaluated in the same way as the internal test.

Long-term efficacy prediction with the models

This study further explored the potential of applying these established models to predict the long-term efficacy of FD patients at the 12-week follow-up in the independent validation. The performance of models and their statistical significance was evaluated in the same way as described above.

Weight calculation and predicting features identification

The weight of the feature reflects the magnitude of each feature's contribution to determining the support vector classification hyperplane [27]. The weight of features was calculated with the formula of *model.SVs' * model.sv_coef* which was introduced in the manual of LIBSVM. The features whose weight exceed the mean + standard deviation were regarded as the predicting features [28].

Results

A total of 552 FD patients in trial 1 and 193 patients in trial 2 were included in the efficacy prediction analyses. The demographic and baseline clinical characteristics of these patients are shown in **Table 1**.

Performance of models in the internal test set

In the internal test set, the established SVC model had an accuracy of 0.773, a specificity of 0.833, a sensitivity of 0.714, and an AUC of 0.804 in predicting acupuncture response at the end of treatment (Fig. 2A-1, 2A-2). Results of the permutation test showed that the accuracy and AUC were both significant (p_ accuracy < 0.001, p_ AUC < 0.001) (Fig. 2A-3, 2A-4).

The SVR models achieved an R^2 of 0.446 and an MSE of 60.110 in the prediction of NDLQI improvement (Fig. 2B-1, 2B-2), an R^2 of 0.413, and an MSE of 108.357 in the prediction of NDSI improvement (Fig. 2C-1, 2C-2) at the end of treatment. Permutation test demonstrated that both the predicting results of NDLQI and NDSI improvements were statistically significant (p_ $R^2 < 0.001$, p_ MSE < 0.001) (Fig. 2B-3, 2B-4, 2C-3, 2C-4).

Predicting features identification

As shown in Fig. 3A, the education level (weight $= 0.617$), gender (weight $= 0.321$), and subtype of disease (weight $= 0.306$) were the most weighted features for the prediction of acupuncture response. Similarly, the education level, subtype of disease, and gender were also critical for the prediction of NDLQI and NDSI improvements. The weights of these features for predicting NDLQI improvement were 2.241, 1.495, and 1.394 (Fig. 3B) and for predicting NDSI improvement were 1.794, 1.294, and 1.154 (Fig. 3C).
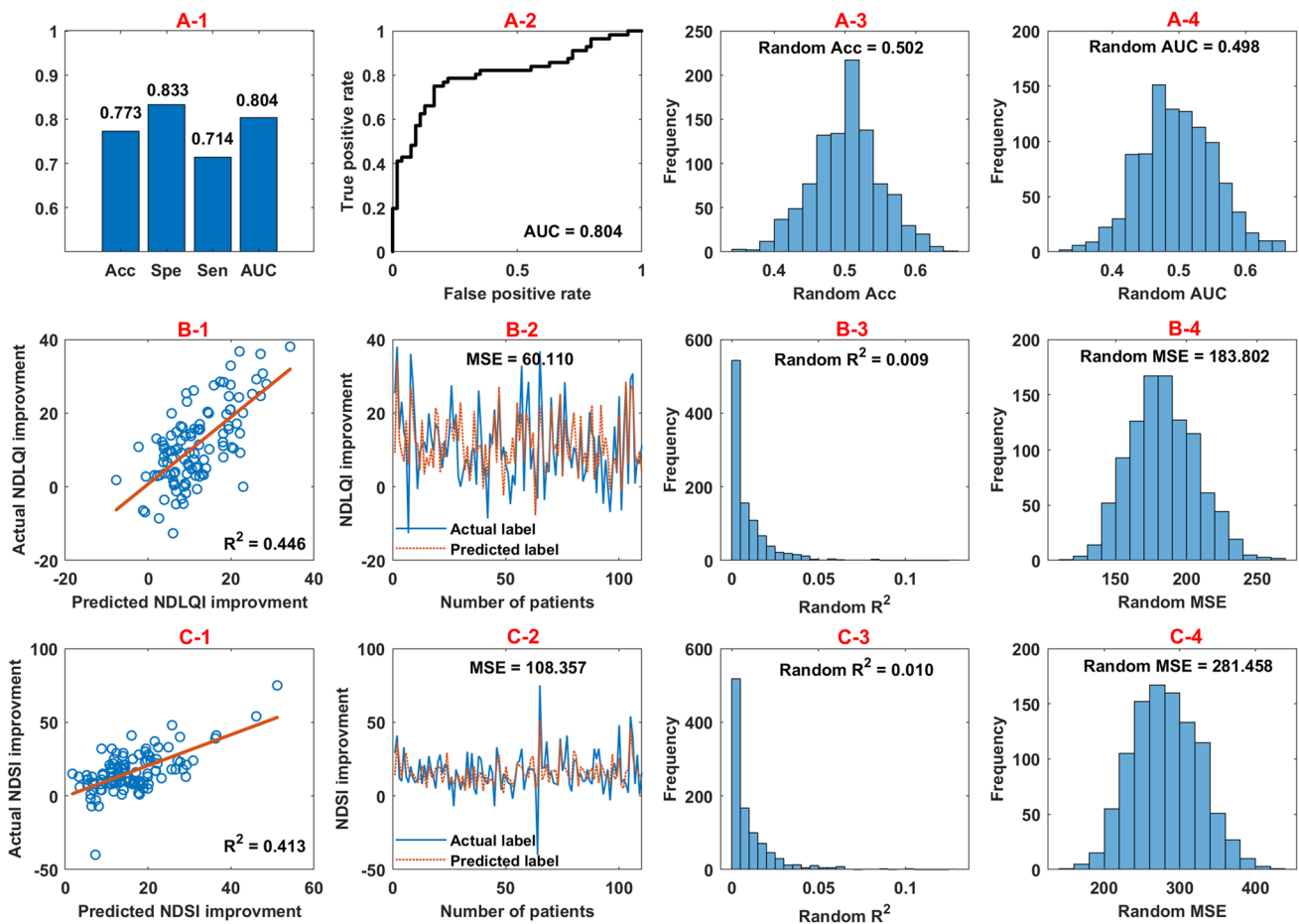

Fig. 2 The performance of prediction models in the internal test set. **(A-1)** illustrates the overall performance of the SVC model in predicting acupuncture response at the end of treatment. **(A-2)** illustrates the receiver operating characteristic curve corresponding to the AUC. **(B-1)** and **(C-1)** show the scatter plots between the actual and the predicted improvements of NDLQI and NDSI of the internal test samples, respectively. **(B-2)** and **(C-2)** illustrate the predicted labels and their corresponding actual labels of these test samples. **(A/B/C-3, A/B/C-4)** displays the results of permutation tests for these predication analyses. **Abbreviation:** Acc, accuracy; Sen, sensitivity; Spe, specificity; AUC, area under the receiver operating characteristic curve; NDSI, Nepean Dyspepsia Symptom Index; NDLQI, Nepean Dyspepsia Life Quality Index; R^2, coefficient of determination; MSE, mean squared error

Table 1 The demographic and clinical characteristics of FD patients

	Trial 1	Trial 2
Gender (male/female)	164/388	63/130
Age (year)	37.01 ± 13.48	39.55 ± 15.19
Body mass index (kg/m^2)	20.82 ± 2.43	20.86 ± 2.48
Education level (1/2/3/4/5)	1/33/199/286/33	4/16/43/107/23
Marriage status (unmarried/married)	184/368	74/119
Duration (month)	67.71 ± 62.73	94.91 ± 93.40
Subtype of disease (PDS/EPS)	355/198	111/82
Epigastric pain score	4.47 ± 3.62	4.23 ± 3.00
Epigastric discomfort score	5.40 ± 3.51	5.85 ± 2.86
Epigastric burning score	1.55 ± 2.65	1.45 ± 2.37
Chest burning score	0.81 ± 2.14	0.94 ± 2.20
Epigastric spasmodic pain score	0.76 ± 1.97	0.95 ± 2.30
Chest pain score	0.61 ± 1.82	0.78 ± 1.94
Irregular meals score	3.53 ± 3.47	2.66 ± 3.40
Reflux of acid/bitterness score	2.91 ± 3.35	2.89 ± 3.06
Epigastric fullness/slow digestion score	6.23 ± 3.48	6.36 ± 3.12
Epigastric pressure score	1.11 ± 2.36	1.62 ± 2.90
Epigastric bloating score	5.59 ± 3.57	5.35 ± 3.32
Nausea score	1.57 ± 2.74	2.30 ± 3.05
Belching score	4.01 ± 3.41	3.46 ± 3.37
Vomiting score	0.58 ± 1.87	0.77 ± 2.13
Halitosis score	0.95 ± 2.52	1.98 ± 3.08
NDSI score	40.08 ± 16.66	41.56 ± 19.97
Interference domain score	76.84 ± 11.12	80.17 ± 16.38
Knowledge/control domain score	72.67 ± 11.06	75.19 ± 15.98
Eat/drink domain score	71.93 ± 12.90	69.64 ± 17.95
Sleep/disturb domain score	75.45 ± 16.10	77.62 ± 19.83
NDLQI score	74.22 ± 10.60	75.65 ± 14.28

The NDSI score contains 15 sub-scores, which measured the severity of 15 common symptoms in patients with functional dyspepsia. The NDLQI score contains 4 dimensions of quality of life, including interference, knowledge/control, eat/drink and sleep/disturb. **Abbreviation:** *1/2/3/4/5* illiterate/primary school/middle school/high school/college and above, *PDS/EPS* postprandial distress syndrome/epigastric pain syndrome, *NDSI* Nepean Dyspepsia Symptom Index, *NDLQI* Nepean Dyspepsia Life Quality Index

Performance of models in the independent validation set

In the independent validation set, the SVC model had an accuracy of 0.720, a sensitivity of 0.716, a specificity of 0.723, and an AUC of 0.782 in predicting acupuncture response. The SVR models obtained an R^2 of 0.327 and an MSE of 116.851 in predicting the improvement of NDLQI and an R^2 of 0.152 and an MSE of 290.836 in predicting the improvement of NDSI. The results of the permutation tests for the predicted values were all significant (Supplementary 2).

Performance of models in long-term efficacy prediction

As shown in Supplementary 3, these models could also predict, to a certain extent, the acupuncture response as well as improvements of NDLQI and NDSI of FD patients at the 12-week follow-up. Namely, the model for the prediction of acupuncture response achieved an accuracy of 0.630, a sensitivity of 0.677, a specificity of 0.581, and an AUC of 0.669. The model for the prediction of NDLQI improvement got an R^2 of 0.346 and an MSE of 149.572. The model for the prediction of NDSI improvement gets an R^2 of 0.145 and an MSE of 374.000.

Performance comparison between linear SVM and other ML algorithms

Based on the machine learning toolbox of Matlab 2017b, this study further compared the performance among the linear SVM and other five common ML algorithms in predicting acupuncture efficacy for FD. The results demonstrated that the linear SVC had better performance than decision tree, Gaussian SVM, boosted tree, K-nearest neighbor, and logistic regression in predicting acupuncture response in both internal test and independent validation sets (Fig. 4A). Linear SVR outperformed the decision tree, Gaussian SVM, boosted tree, and Gaussian process regression, while similar to multiple linear regression in the prediction of NDLQI improvement in these two sets (Fig. 4B). Linear SVR had better performance than the decision tree, Gaussian SVM and boosted tree, and the similar performance to the multiple linear regression and Gaussian process regression in the prediction of NDSI improvement in the internal test set (Fig. 4C).

Discussion

Combining the SVM algorithms and 28 routine clinical features, the current study successfully predicted the acupuncture response as well as the improvements of clinical symptoms and QoL of FD patients at the end of treatment and follow-up. The gender, subtype of disease, and education level were identified as the critical features for the successful predictions. These current findings provided a reference for the application of PPPM/3PM in CAM and promoted the paradigm change of CAM from reactive medicine to PPPM/3PM.

SVM is well-suited for the development of efficacy prediction models

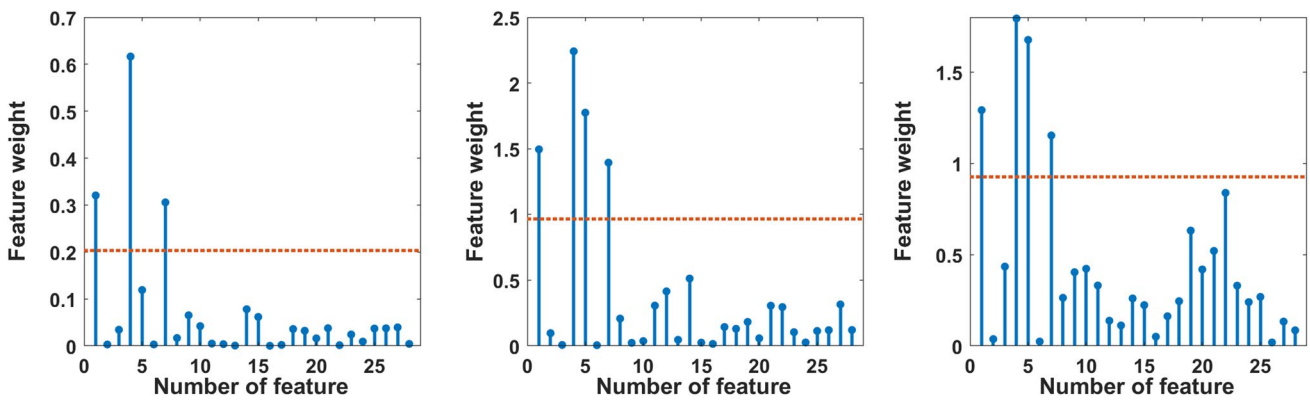The paradigm changes from delayed intervention to the PPPM/3PM have been identified as an essential

Fig. 3 The weight of each feature in predicting acupuncture efficacy. **(A)** illustrates the weight of each feature in predicting acupuncture response; **(B)** and **(C)** show the weight of each feature in predicting NDSI and NDLQI improvements. The ranking order of the features is consistent with Table 1. The red dotted lines represent the mean + standard deviation of features' weights

Fig. 4 Performance comparison among linear SVM and other ML algorithms in prediction acupuncture efficacy. **(A)**, **(B)**, and **(C)** illustrate the results of performance comparison in predicting acupuncture response, NDLQI improvements, and NDSI improvements among these six algorithms in the internal validation set (the left panels) and independent validation set (the right panels).
Abbreviation: Acc, accuracy; Sen, sensitivity; Spe, specificity; NDSI, Nepean Dyspepsia Symptom Index; NDLQI, Nepean Dyspepsia Life Quality Index; R^2, coefficient of determination; MSE, mean squared error; SVC, support vector classification; DT, decision tree; LR, logistic regression; rbf SVM, radial basis function support vector classification; BT, boosted tree; KNN, K-nearest neighbor; SVR, support vector regression; MLR, multiple linear regression; GPR, Gaussian process regression
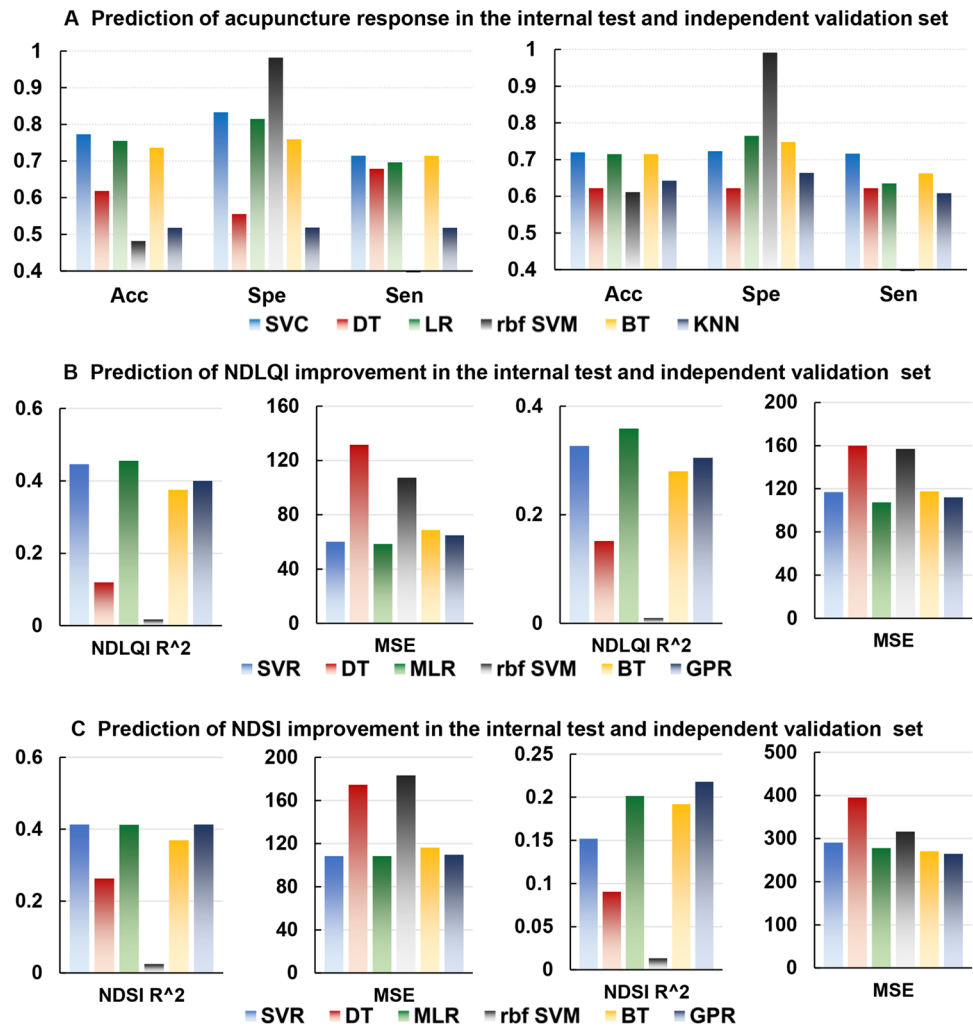
transformation of the overall healthcare approach to benefit the patients and society at large. Different from the conventional medical framework where treatment strategies could only be determined based on the standardized guideline recommendations and adjusted reactively based on the feedback of patients, the ML-based prediction models in

the PPPM/3PM framework could assist doctors to forecast patients' responses to acupuncture in advance, so as to preemptively stratify patients and develop personalized treatment tailored to the individual.

In recent years, a growing number of studies applied the generalized linear models (GLM, e.g., logistic regression and multiple linear regression) to explore the potential associations between the baseline clinical data and acupuncture efficacy [29, 30]. For example, Plunkett A et al. [29] applied the logistic regression algorithm to determine the predictors of acupuncture treatment for chronic pain and found that the higher baseline pain scores were associated with the positive treatment outcomes. These findings provided a valuable reference for developing rational medical strategies and enabling individualized and precise treatment. However, the GLM-based responders screening methods are sensitive to the outliers, and may oversimplify the complex relationships among features, which inevitably lead to the loss of implicit predictive features [21]. The SVM has manifested the strength in discovering the underlying patterns from complex data in a large number of studies [31, 32]. SVM could provide the non-linear classification by mapping the input space into high-dimensional feature space using kernel functions, which outperforms the conventional GML methods when dealing with classification and prediction issues with intricate data [33]. In the current study, we compared the performance of several common ML algorithms and detected that the prediction models constructed based on the SVM algorithm had better performance than others. This finding was consistent with previous studies [34, 35] and illustrated the outstanding performance of SVM in prediction analysis.

The prediction models developed with routine clinical data have better translational potential

Based on the SVM algorithm, recent studies have attempted to apply the novel biomarkers, such as high-throughput neuroimaging metrics, to predict acupuncture efficacy and yielded encouraging results [36, 37]. For example, Tu and colleagues detected with the SVR algorithm that the baseline functional brain connectivity patterns could predict symptom relief in patients with chronic low back pain [36]. However, neuroimaging data, especially the high-field magnetic resonance data, was still not easy to acquire in primary healthcare, which severely restrained the clinical applications of these novel prediction models and biomarkers. In contrast, the prediction models constructed with features collected in routine practice may have greater translational potential. Therefore, the current study developed clinically accessible models to predict the acupuncture efficacy for FD, based on the

demographic characteristics and clinical symptoms metrics. These established models have been encapsulated into a transferable MATLAB code (https://github.com/YinTao0828/FD_ACU_Predicition.git), which could be directly used for the future prediction of prospective data.

The gender, subtype of disease, and education level are the considerable features for the precise predictions

The current study indicated that gender, subtype of disease, and education level were the most weighted features that played important roles in the construction of prediction models not only for acupuncture response, but also for the improvements of NDSI and NDLQI. The significant gender differences of FD patients in prevalence and disease severity had been widely recognized [38]. While in terms of treatment efficacy, the response of female FD patients was also lower than that of males [39]. Therefore, gender could be served as a robust predictor of FD symptoms and QoL improvements, which is similar to the results of Lee AA et al. [40]. Other than gender, the subtype of disease was also found to be another key predictor of acupuncture efficacy for FD. Compared to PDS, patients with EPS often had more severe symptoms and less efficacy in treatment [41]. Our previous study also detected that PDS patients responded better to acupuncture therapy than EPS patients [11]. These studies, together with our current finding, supported the opinion that subtype was an important factor affecting the prognosis of FD. In addition, the education level was also found to be predictive to the efficacy of acupuncture treatment for FD, which was consistent with a previous study [42]. The possible explanation for this finding was that higher education was associated with a greater positive attitude to participate in acupuncture trials, which might enhance treatment response.

The results also demonstrated that these 22 disease-related features including duration, baseline NDSI, and baseline NDLQI score contributed less to the prediction outcomes. It was inconsistent with a recent multivariate linear regression study [12]. This could be attributed to one of the major drawbacks of SVM, known as the black-box effects [43], which implied the opacity of the data processing and the subsequent difficulty in offering explanations for the judgments. In Supplementary 4, we use only these 3 highlighted features to construct the prediction models and find that the performance of the models decreased significantly. This finding indicated that the prediction performance of SVM models was not only influenced by the high weighted features, but also closely related to the complex associations among multiple features.

Strength and limitations

Based on the routine clinical features and SVM algorithm, the current study established the prediction models for the efficacy of acupuncture in the treatment of FD. The prediction models could not only improve the selection of acupuncture responders and reduce the non-response rate of acupuncture for FD in the future, but also provide a reference for the application of PPPM/3PM in the field of CAM. Nonetheless, several limitations should also be concerned in this study. Firstly, as the datasets were retrospectively collected, some features (such as the emotional state and medication use [12]) with the predictive value might not be included, which might be the reason why the accuracy of current models was not extremely high. The subsequent study will increase the dimensions of features to further enhance the performance of prediction models. Secondly, the current study did not consider the impacts of different acupuncture prescriptions for treatment efficacy, as the greatest interest of this study was the general influences of patients' baseline conditions on the efficacy of acupuncture. Thirdly, due to the black-box effect, the weight of features obtained with the SVM algorithm was the relative value and therefore could not provide a concrete contribution of the feature to the predicted outcomes.

Conclusions and expert recommendations

In conclusion, this study constructed the SVM prediction models based on the demographic and clinical characteristics, which predicted the acupuncture response and improvement of symptoms and QoL in FD patients successfully not only at the end of treatment but also at the 12-week follow-up. These findings highlighted the promising future of integrating the ML techniques and routine clinical data to predict acupuncture efficacy and provided a novel approach to facilitate the paradigm of CAM changing from reactive to PPPM/3PM.

The prediction models developed accordingly is promising to assist doctors in judging patients' responses to acupuncture in advance, so that they could tailor and adjust acupuncture treatment plans for different patients in a prospective rather than the reactive manner, which could greatly improve the clinical efficacy of acupuncture treatment for FD and save medical expenditures. Specifically, the positive acupuncture treatment should be given to the predicted responders, while for the predicted non-responders, other effective therapies need to be supplemented to improve the clinical efficacy.

Code availability The prediction models (Matlab code) generated in this study can be found on our Github homepage: https://github.com/YinTao0828/FD_ACU_Predicition.git. Reasonable requests for the original data can be sent to the corresponding author.

Declarations

Ethics approval These included two trials were approved by the Ethics Committee of Hospital of Chengdu University of Traditional Chinese Medicine.

Consent to participate All the patients provided written informed consents before entering the trial.

Consent for publication All authors had full access to all the data in the study and accept responsibility to submit for publication.

Conflict of interest The authors declare no competing interests.

Role of the funder The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

References

1. Stanghellini V, Chan FK, Hasler WL, Malagelada JR, Suzuki H, Tack J, et al. Gastroduodenal disorders. Gastroenterology. 2016;150(6):1380–92.
2. Barberio B, Mahadeva S, Black CJ, Savarino EV, Ford AC. Systematic review with meta-analysis: global prevalence of

uninvestigated dyspepsia according to the Rome criteria. Aliment Pharmacol Ther. 2020;52(5):762–73.

3. Ford AC, Marwaha A, Lim A, Moayyedi P. What is the prevalence of clinically significant endoscopic findings in subjects with dyspepsia? Systematic review and meta-analysis. Clin Gastroenterol Hepatol. 2010;8(10):830–7.

4. Ford AC, Mahadeva S, Carbone MF, Lacy BE, Talley NJ. Functional dyspepsia. Lancet. 2020;396(10263):1689–702.

5. Brook RA, Kleinman NL, Choung RS, Melkonian AK, Smeeding JE, Talley NJ. Functional dyspepsia impacts absenteeism and direct and indirect costs. Clin Gastroenterol Hepatol. 2010;8(6):498–503.

6. Wang W, Russell A, Yan Y. Traditional Chinese medicine and new concepts of predictive, preventive and personalized medicine in diagnosis and treatment of suboptimal health. EPMA J. 2014;5(1):4.

7. Yang JW, Wang LQ, Zou X, Yan SY, Wang Y, Zhao JJ, et al. Effect of acupuncture for postprandial distress syndrome: a randomized clinical trial. Ann Intern Med. 2020;172(12):777–85.

8. Ma TT, Yu SY, Li Y, Liang FR, Tian XP, Zheng H, et al. Randomised clinical trial: an assessment of acupuncture on specific meridian or specific acupoint vs. sham acupuncture for treating functional dyspepsia. Aliment Pharmacol Ther. 2012;35(5):552–61.

9. Zheng H, Xu J, Sun X, Zeng F, Li Y, Wu X, et al. Electroacupuncture for patients with refractory functional dyspepsia: a randomized controlled trial. Neurogastroenterol Motil. 2018;30(7):e13316.

10. Zeng F, Qin W, Ma T, Sun J, Tang Y, Yuan K, et al. Influence of acupuncture treatment on cerebral activity in functional dyspepsia patients and its relationship with efficacy. Am J Gastroenterol. 2012;107(8):1236–47.

11. Rubio A, Van Oudenhove L, Pellissier S, Ly HG, Dupont P, Lafaye de Micheaux H, et al. Uncertainty in anticipation of uncomfortable rectal distension is modulated by the autonomic nervous system--a fMRI study in healthy volunteers. Neuroimage. 2015;107:10–22.

12. Singh P, Ballou S, Rangan V, Katon J, Hassan R, Iturrino J, et al. Clinical and psychological factors predict outcome in patients with functional dyspepsia: a prospective study. Clin Gastroenterol Hepatol. 2021;S1542–3565(21):00825–9.

13. Rafique R, Islam SMR, Kazi JU. Machine learning in the prediction of cancer therapy. Comput Struct Biotechnol J. 2021;19:4003–17.

14. Barrett M, Boyne J, Brandts J, Brunner-La Rocca HP, De Maesschalck L, De Wit K, et al. Artificial intelligence supported patient self-care in chronic heart failure: a paradigm shift from reactive to predictive, preventive and personalised care. EPMA J. 2019;10(4):445–64.

15. Golubnitschaja O, Kinkorova J, Costigliola V. Predictive, preventive and personalised medicine as the hardcore of "Horizon 2020": EPMA position paper. EPMA J. 2014;5(1):6.

16. Golubnitschaja O, Baban B, Boniolo G, Wang W, Bubnov R, Kapalla M, et al. Medicine in the early twenty-first century: paradigm and anticipation - EPMA position paper 2016. EPMA J. 2016;7(1):23.

17. Birkenbihl C, Emon MA, Vrooman H, Westwood S, Lovestone S, Hofmann-Apitius M, et al. Differences in cohort study data affect external validation of artificial intelligence models for predictive diagnostics of dementia - lessons for translation into clinical practice. EPMA J. 2020;11(3):367–76.

18. Garnica O, Gómez D, Ramos V, Hidalgo JI, Ruiz-Giardín JM. Diagnosing hospital bacteraemia in the framework of predictive, preventive and personalised medicine using electronic health records and machine learning classifiers. EPMA J. 2021;12(3):1–17.

19. Kinkorová J, Topolčan O. Biobanks in the era of big data: objectives, challenges, perspectives, and innovations for predictive, preventive, and personalised medicine. EPMA J. 2020;11(3):333–41.

20. Vidyasagar M. Identifying predictive features in drug response using machine learning: opportunities and challenges. Annu Rev Pharmacol Toxicol. 2015;55:15–34.

21. Wu Z, Zhu M, Kang Y, Leung EL, Lei T, Shen C, et al. Do we need different machine learning algorithms for QSAR modeling? A comprehensive assessment of 16 machine learning algorithms on 14 QSAR data sets. Brief Bioinform. 2021;22(4):bbaa321.

22. Tao Y, Sun R, He Z, Ma P, Zeng F. Clinical effects of acupuncture treatment in functional dyspepsia based on resting-state functional brain network. China Journal of Traditional Chinese Medicine and Pharmacy. 2020;35(5):2581–4 (**(In Chinese)**).

23. Talley NJ, Verlinden M, Jones M. Validity of a new quality of life scale for functional dyspepsia: a United States multi-center trial of the Nepean dyspepsia index. Am J Gastroenterol. 1999;94(9):2390–7.

24. Chih-Chung C, Chih-Jen L. Libsvm: a library for support vector machines. ACM Trans Intell Syst Technol. 2011;2(3):1–27.

25. Jones M, Talley NJ. Minimum clinically important difference for the Nepean dyspepsia index, a validated quality of life scale for functional dyspepsia. Am J Gastroenterol. 2009;104(6):1483–8.

26. Golland P, Fischl B. Permutation tests for classification: towards statistical significance in image-based studies. Inf Process Med Imaging. 2003;18:330–41.

27. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20(3):273–97.

28. Tian L, Wang J, Yan C, He Y. Hemisphere- and gender-related differences in small-world brain networks: a resting-state functional MRI study. Neuroimage. 2011;54(1):191–202.

29. Plunkett A, Beltran T, Haley C, Kurihara C, McCoart A, Chen L, et al. Acupuncture for the treatment of chronic pain in the military population: factors associated with treatment outcomes. Clin J Pain. 2017;33(10):939–43.

30. Baeumler PI, Conzen P, Irnich D. High Temporal summation of pain predicts immediate analgesic effect of acupuncture in chronic pain patients-a prospective cohort study. Front Neurosci. 2019;13:498.

31. Liang Y, Ding R, Wang J, Gong X, Yu Z, Pan L, et al. Prediction of response after cardiac resynchronization therapy with machine learning. Int J Cardiol. 2021;doi: https://doi.org/10.1016/j.ijcard.2021.09.049. Online ahead of print.

32. Cao B, Cho RY, Chen D, Xiu M, Wang L, Soares JC, et al. Treatment response prediction and individualized identification of first-episode drug-naïve schizophrenia using brain functional connectivity. Mol Psychiatry. 2020;25(4):906–13.

33. Liu F, Guo W, Fouche JP, Wang Y, Wang W, Ding J, et al. Multivariate classification of social anxiety disorder using whole brain functional connectivity. Brain Struct Funct. 2015;220(1):101–15.

34. Huang WC, Lee PL, Liu YT, Chiang AA, Lai F. Support vector machine prediction of obstructive sleep apnea in a large-scale Chinese clinical sample. Sleep. 2020;43(7):zsz295.

35. Lee HJ, Hwang SI, Han SM, Park SH, Kim SH, Cho JY, et al. Image-based clinical decision support for transrectal ultrasound in the diagnosis of prostate cancer: comparison of multiple logistic regression, artificial neural network, and support vector machine. Eur Radiol. 2010;20(6):1476–84.

36. Tu Y, Ortiz A, Gollub RL, Cao J, Gerber J, Lang C, et al. Multivariate resting-state functional connectivity predicts responses to real and sham acupuncture treatment in chronic low back pain. Neuroimage Clin. 2019;23:101885.

37. Yin T, Sun G, Tian Z, Liu M, Gao Y, Dong M, et al. The spontaneous activity pattern of the middle occipital gyrus predicts the clinical efficacy of acupuncture treatment for migraine without aura. Front Neurol. 2020;11:588207.

38. Narayanan SP, Anderson B, Bharucha AE. Sex- and gender-related differences in common functional gastroenterologic disorders. Mayo Clin Proc. 2021;96(4):1071–89.

39. Kim SE, Kim N, Park SM, Kim WH, Baik GH, Jo Y, et al. Female gender is a poor predictive factor of functional dyspepsia resolution after helicobacter pylori eradication: a prospective, multicenter Korean trial. Korean J Gastroenterol. 2018;72(6):286–94.

40. Lee AA, Rao K, Parkman HP, McCallum RW, Sarosiek I, Nguyen LA, et al. Baseline predictors of longitudinal changes in symptom severity and quality of life in patients with suspected gastroparesis. Clin Gastroenterol Hepatol. 2020;S1542–3565(20):31293–303.

41. Tack J, Talley NJ. Functional dyspepsia–symptoms, definitions and validity of the Rome III criteria. Nat Rev Gastroenterol Hepatol. 2013;10(3):134–41.

42. Yeung WF, Chung KF, Yu YB, Lao L. What predicts a positive response to acupuncture? A secondary analysis of three randomised controlled trials of insomnia. Acupunct Med. 2017;35(1):24–9.

43. Azodi CB, Tang J, Shiu SH. Opening the black box: interpretable machine learning for geneticists. Trends Genet. 2020;36(6):442–55.

Authors and Affiliations

Tao Yin[1,2] · Hui Zheng[1] · Tingting Ma[3] · Xiaoping Tian[1] · Jing Xu[1] · Ying Li[4] · Lei Lan[1,2] · Mailan Liu[5] · Ruirui Sun[1,2] · Yong Tang[1,6] · Fanrong Liang[1] · Fang Zeng[1,2] [ID]

1. Acupuncture and Tuina School/The 3rd Teaching Hospital, Chengdu University of Traditional Chinese Medicine, Chengdu 610075, Sichuan, China

2. Acupuncture-Brain Science Research Center, Chengdu University of Traditional Chinese Medicine, Chengdu 610075, Sichuan, China

3. Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu 610075, Sichuan, China

4. Graduate School, Chengdu University of Traditional Chinese Medicine, Chengdu 610075, Sichuan, China

5. Acupuncture and Tuina School, Hunan University of Chinese Medicine, Changsha 410208, Hunan, China

6. Key Laboratory of Sichuan Province for Acupuncture and Chronobiology, Chengdu 610075, Sichuan, China