Predicting Timing of Surgical Intervention Using Recurrent Neural Network for Necrotizing Pancreatitis

JIAWEI LUO¹, LAN LAN², LI PENG¹, MENGJIAO LI¹, HUIMIN LU², DUJIANG YANG³, XIAOYAN YANG¹, AND XIAOBO ZHOU⁴

¹West China Biomedical Big Data Center, West China Hospital/West China School of Medicine, Sichuan University, Chengdu 610041, China
²Department of Pancreatic Surgery, West China Hospital/West China School of Medicine, Sichuan University, Chengdu 610041, China
³Department of Gastrointestinal Surgery, West China Hospital/West China School of Medicine, Sichuan University, Chengdu 610041, China
⁴School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA

Corresponding author: Lan Lan (lanl@scu.edu.cn)

This work was supported by the 1·3·5 Project for Disciplines of Excellence, West China Hospital, Sichuan University, under Grant ZYJC18010, the 1·3·5 Project for Disciplines of Excellence—Clinical Research Incubation Project, West China Hospital, Sichuan University, under Grant 2019HXFH022, the Postdoctoral Research Project, West China Hospital, Sichuan University, under Grant 2019HXBH039, and the Center of Excellence-International Collaboration Initiative Grant, under Grant 139170052.

ABSTRACT

Previous research focused on the qualitative discussion of the correlation between surgical time and mortality in acute pancreatitis (AP). Recommendations for surgical timing of necrotizing pancreatitis are delayed as far as possible, without recommendations for individuals. The aim of this article is to predict timing of surgical intervention in necrotizing pancreatitis with recurrent neural network (RNN). Time series data in AP were retrospectively extracted from a hospital in China (n = 15,813) to develop model, and the Cerner Health Facts database in the United States (n = 142,650) was used to externally validate. The developed model, time-aware Phased-Decay long short-term memory (LSTM), was used to predict timing of surgical intervention and critical clinical features in necrotizing pancreatitis based on laboratory tests, compared to other machine learning models. Area under the ROC curve (AUC) of RNN-based models was more than 0.70. The AUC of time-aware Phased-Decay LSTM (0.75) with more explanatory was similar to that (0.76) of Phased-Decay LSTM to predict surgical timing. The developed model visualized the specific surgical process and laboratory indicators changes for the patients with AP from the onset to discharge. Different clinical features had different contribution modes with time to the specific surgical event in AP. Heart function contributed the most to the prediction of surgical intervention at the onset and in the first week. Our developed model could monitor the specific surgical process from the onset of AP to discharge and extract the contribution modes of clinical features with time in necrotizing pancreatitis.

INDEX TERMS

Recurrent neural network, surgical timing, necrotizing pancreatitis, electronic medical record.

I. INTRODUCTION

Acute pancreatitis (AP), an inflammatory disorder of pancreas, is the leading cause of admission to hospital for gastrointestinal disorders in the United States and many other countries [1], [2], and cost an estimated 2.6 billion dollars per year in hospitalization [3], with an annual incidence ranging from 13 to 45/100,000 people [4]–[6]. The inpatient mortality of AP is 0.8% [7], and in severe cases, it could even reach 30% [8] where presenting many therapeutic challenges [9], [10]. Severe AP is characterized by the presence of either infected (peri)pancreatic necrosis or persistent organ failure (POF) [11]. The current management guideline for necrotizing pancreatitis from IAP/APA [8] recommends delaying the timing of surgical intervention until four or more weeks. In addition to IAP/APA guideline, recommendations for surgical timing of necrotizing pancreatitis in the United States, United Kingdom, Italy, and Japan are also delayed as far as possible, without recommendations for individuals [12]–[15]. A prospective study observed that POF in the first week was more likely to determine mortality than infection in patients with necrotizing pancreatitis [16]. While a
prospective cohort study from the Netherlands showed that there were no associations between infection, onset of organ failure, duration of organ failure and mortality in patients with necrotizing pancreatitis [17]. Additionally, current studies cannot explain the relationship between the suggested surgical indications of necrotizing pancreatitis and surgical timing [18]–[21]. A randomized controlled trial (RCT) which was established to optimize timings of surgery following percutaneous catheter drainage (PCD) in patients with infected pancreatic necrosis was forced to stop early due to practical difficulties [22]. The surgical timing problem has not been resolved by RCT.

Therefore, in order to predict the timing of surgical intervention for necrotizing pancreatitis individually, a novel recurrent neural network (RNN) based deep learning model, namely time-aware Phased-Decay LSTM, was developed and applied. There are three contributions in this study: 1) we developed a novel deep learning model, which performed better than conservative machine learning approaches; 2) the developed model was applied to predict the timing of the specific surgical intervention and important laboratory indicators in necrotizing pancreatitis for the first time; 3) patients with necrotizing pancreatitis could be monitored from the onset of AP to discharge by using our developed model.

II. RELATED WORK

In order to solve specific medical problems, many machine learning and deep learning algorithms are improved to be effectively applied and solved on the medical problems [23]–[28]. Clinical studies generally have a small sample size, so data enhancement or transfer learning is needed to help train the model [29]–[32]. However, the sample size of this study is relatively large in clinical practice, so data enhancement and transfer learning can be avoided.

Additionally, predicting the timing of surgical intervention in patients with necrotizing pancreatitis is a very meaningful task. However, previous research just focused on qualitative discussion of the correlation between surgical time and mortality rate [33]. Even though our previous study has proved that there is a L-shaped relationship between timing of surgical intervention and risk of death in necrotizing pancreatitis patients [34], it cannot explain the relationship between the timing of surgical intervention and risk of death personally. The methods they used were mainly based on hypothesis testing and regression models. Machine learning methods were also used to classify infected necrotizing pancreatitis for surgery within or beyond four weeks [35]. However, modeling temporal information is common in the field of clinical research. Wu et al. made a comprehensive summary of the characteristics of multivariate time series [36]. The bin method is often adopted because time stamp is unified and the sequences are changed from asynchronous to synchronous [37], [38]. The typical approach is to fill the missing values in some way such as mean value filling, last observation filling, generating missing mask information, filling with zero etc., and then to predict by using classifiers [37]–[39]. A good approach is to integrate missing value filling and modeling prediction. A representative work is applying decay based on gated recurrent unit (GRU), a variant of RNN, named GRU-D [40]. The GRU-D model introduces the concept of decay, which makes full use of the missing mask information of the variables and the interval time information between the monitoring points to fill the missing values. The filling process is guided by the prediction results. Such an architecture can relatively give a more reasonable padding value. However, GRU-D may fill in some values that should not exist, thereby losing the information of whether the patient was actually followed up, which can be solved by phase gate of the phased long short-term memory (Phased-LSTM) [41], [42], another variant of RNN. Recently, with respect to modeling for clinical time series data, Choi et al. used the attention mechanism, which weighted two dimensions of time and variables and achieved good accuracy and interpretability [43]. Neil et al. extended the gate mechanism of RNN to automatically choose which moments of information are valuable for hidden state updates. As a result, not only the model training was accelerated, but better accuracy also was achieved [42]. It was reported that in the time series prediction task, the performance of the RNN-derived model is better than that of the Hidden Markov model (HMM) [44].

III. TIME-AWARE PHASED-DECAY LSTM

A. MOTIVATION

It is very challenging to fuse and model based on the asynchronous, unequal length, and heterogeneity data. Therefore, laboratory tests only were used to predict. According to our data, laboratory tests included 1 ∼ 33 measurements, and the average number of measurements was 5.36 in AP. Those sequences were characterized by unequal spacing, unequal length, and asynchronous. RNN and its derivative models can handle the problem of unequal length for temporal sequences, which cannot be handled by HMM. Unequal spacing and asynchronous can result in missing values during temporal sequences analysis. We adopted decay to improve the missing issue in this study. The introduction of phase gate enables RNN to automatically learn the weights in the time dimension and control the hidden layer update. At the same time, time is also mapped into a vector of the same length as the input vector and multiplied by it to achieve the effect of filtering redundant information from the variable dimension. Missing values at each time point in sequences are different, and decay can fill in sequences. However, decay is not able to directly solve the problem of rationality of filling by using phase gate. Therefore, we introduced weight gate, which can automatically learn different time points and give their weights, so as to make up for the problem of filling rationality. On the other hand, as AP is a sudden-onset and severe disease, time is very important for the treatment process, and the variables are different for the importance of surgical intervention at different time points. The time-aware mechanism proposed in this study can maximize the usability of time information and
provide important clinical features for surgical intervention of necrotizing pancreatitis at different time points. This is the motivation for developing time-aware Phased-Decay LSTM.

\[ \hat{x}_t = m_t x_t + (1-m_t) (\gamma_t x_{t-1} + (1-\gamma_t) \tilde{x}) \]  \hspace{1cm} (11)

Time weight is determined by Eq. (13), which is developed to update the input for LSTM as shown in Eq. (14). Time weight represents time-aware mechanism in this work.

\[ w_t = \sigma_v (t \ast W_w + b_w) \]  \hspace{1cm} (13)
\[ x_t = \tilde{x}_t \odot w_t \]  \hspace{1cm} (14)

Phase gate function \( k_t \) of Phased-LSTM [42] is used to achieve more efficient training over the irregularly sampled data. To reduce the calculation burden and achieve the same effect as the previous segmented function, we use \( \sin \) function to get \( k_t \) formulating as

\[ k_t = \sin (t \ast W_h + s) + c \]  \hspace{1cm} (15)

where \( s \) represents phase position and \( c \) represents adjustment of the vertical position of the periodic function through updating:

\[ \tilde{c}_t = f_t \odot c_{t-1} + i_t \odot \sigma_c (x_t W_{sc} + \tilde{h}_{t-1} W_{hc} + b_c) \]  \hspace{1cm} (16)
\[ \tilde{h}_t = o_t \odot \tanh (\tilde{c}_t) \]  \hspace{1cm} (17)
\[ h_t = k_t \odot \tilde{h}_t + (1-k_t) \odot \tilde{h}_{t-1} \]  \hspace{1cm} (18)

The prediction values can be calculated by:

\[ \hat{\psi}_t = \text{softmax} (\text{linear}(h_t)) \]  \hspace{1cm} (19)
\[ \hat{\psi}_t = \text{identity}(\text{linear}(h_t)) \]  \hspace{1cm} (20)

where \( \text{softmax} \) function is for classification output and \( \text{identity} \) function is for continuous output, and \( W, s, c, b \) are parameters. We got parameters by minimizing the loss functions defined below.

**C. PREDICTING THE TIMING OF SPECIFIC SURGICAL INTERVENTION**

We defined our loss function for the single task of predicting the timing of specific surgical intervention in AP as:

\[ \text{Loss}_1 = \frac{1}{\text{batch}} \ast \frac{1}{T} \sum_{1}^{\text{batch}} \sum_{1}^{T} -\varphi_t (\log(\hat{\psi}_t)) \]
\[ - (1-\varphi_t) - \log(1-\hat{\psi}_t) \]  \hspace{1cm} (21)

where \( \text{batch} \) represents the number of patients in an iteration, \( T \) represents the number of measurement of features for a patient in an encounter, \( \varphi_t \) indicates the real label of surgical intervention and \( \hat{\psi}_t \) represents the prediction value.
D. MULTI-TASK PREDICTION

We predicted critical laboratory indicators for AP such as C-reactive protein, white blood cell count (WBC), creatinine, and calcium ion, as well as predicting the timing of specific surgical intervention using the development cohort. These critical laboratory indicators were output as continuous values. The loss functions for multi-task were defined as:

\[
Loss_2 = Loss_1 + \sum_{e=1}^{E} (\psi_e - \hat{\psi}_e)^2 \tag{22}
\]

where \(\psi_e\) is the real label, \(\hat{\psi}_e\) is the predicting value of critical laboratory indicators and \(E\) is the number of critical laboratory indicators.

IV. EXPERIMENTS

A. PATIENTS

For necrotizing pancreatitis, surgical treatment can be classified into three categories: drainage, pancreatectomy, and pancreatic necrotic tissue removal plus extensive drainage [45]. Therefore, in this study, patients who underwent any of the three surgical interventions described above were defined as patients with necrotizing pancreatitis to develop our model. We collected data on patients admitted to West China Hospital of Sichuan University from 2010 to 2018. The inclusion criteria for the study patients were: 1) diagnosis of AP on admission based on International Classification of Diseases (ICD) code; 2) having as least one surgical intervention experience including pancreatic drainage, pancreatic debridement, or pancreatectomy in an encounter. 15,813 patients diagnosed with AP were initially identified. We further screened patients who underwent the specific surgery and survived after surgery. Finally, 1,112 patients were included in our development cohort (Fig. 2).

B. TREATMENT PROTOCOL AND DATA COLLECTION

After admission, all patients diagnosed with AP from West China Hospital of Sichuan University initially received traditional treatment. The etiology for patients were main biliary, alcohol abuse and others. The indications for the specific surgery were infectious necrotizing pancreatitis or severe complications such as POF. When surgical indications occurred, experienced surgeons discussed the case with radiologist to decide the type and time for surgical intervention, which delayed as much as possible after four weeks. When patients had persistent clinical manifestations of sepsis, prompt surgical intervention was considered. Patients’ data were extracted from the Electronic Medical Record (EMR) of West China Hospital of Sichuan University to develop model, including diagnosis, laboratory tests, vital signs, surgery information, etc. This study retrospectively collected data of patients with AP and followed the STROBE guidelines [46] for observational studies. The research protocol was approved by the ethics review board of West China Hospital of Sichuan University, and the need for informed consent was waived owing to the retrospective nature of the study. The Cerner Health Facts database was also collected to validate the performance of our developed model externally. The Cerner Health Facts database comprises de-identified Electronic Health Records (EHR) data from over 600 participating Cerner client hospitals and clinics in the United States and represents over 68 million unique patients from 2000 to 2017.

C. TRAINING AND TESTING

Given a time point \(t\), if the time of any specific surgical intervention was included in the range of \([t, t + 7]\) days, the label of \(y\) at \(t\) was 1, otherwise it was 0 in the development cohort. All laboratory tests information from admission to \(t\) was regarded as input, and then features with a missing rate of less than 85% were selected. The missing rate was calculated by the formula:

\[
\frac{\sum_{j=1}^{R} \sum_{i=1}^{l_{ij}} I_{ij}}{\sum_{i=1}^{R} \sum_{j=1}^{l_{ij}} 1}
\]

where \(I_{ij}\) represents whether it is missing and \(n\) is the number of patients. The length of the time point was taken as 15. The original length was kept if the length was less than 15, and the last 15 time points were taken if the length was over 15. We used the union of the monitoring points of each sequence, which preserves the original time information to the greatest extent.

The raw dataset after pre-processing in the development cohort was further divided into training dataset, validation dataset and testing dataset according to the proportions of 70%, 20% and 10%. In order to compare with traditional classifiers that perform well, such as support vector machine (SVM) and random forest (RF), SVM and RF with decay were adopted to predict the timing of the specific surgical intervention. We also compared standard LSTM and GRU, time-aware LSTM and GRU, and Phased-LSTM, where decay was used to fill missing values.

The definition of the label in the single task of prediction is whether to perform target surgery in a certain time window in the future. Obviously, the length of the time window is
TABLE 1. Performance of the development cohort.

| Model          | Accuracy Validation | Accuracy Test | AUC Validation | AUC Test | Sensitivity Validation | Sensitivity Test | Specificity Validation | Specificity Test |
|----------------|---------------------|---------------|----------------|----------|------------------------|-----------------|------------------------|-----------------|
| D-SVM          | 0.65±0.01           | 0.65±0.01     | 0.64±0.01      | 0.76±0.01| 0.74±0.01              | 0.45±0.01       | 0.49±0.01              |
| D-RF           | 0.67±0.01           | 0.67±0.01     | 0.70±0.01      | 0.75±0.02| 0.75±0.02              | 0.31±0.01       | 0.48±0.01              |
| D-LSTM         | 0.66±0.01           | 0.65±0.02     | 0.71±0.02      | 0.75±0.02| 0.75±0.02              | 0.61±0.01       | 0.61±0.02              |
| D-GRU          | 0.68±0.02           | 0.68±0.02     | 0.72±0.02      | 0.76±0.02| 0.75±0.02              | 0.63±0.02       | 0.63±0.02              |
| T-D-LSTM       | 0.71±0.01           | 0.72±0.01     | 0.74±0.02      | 0.77±0.02| 0.77±0.01              | 0.65±0.01       | 0.65±0.01              |
| T-D-GRU        | 0.73±0.02           | 0.73±0.01     | 0.75±0.02      | 0.76±0.02| 0.78±0.01              | 0.65±0.02       | 0.66±0.01              |
| P-D-LSTM       | 0.72±0.02           | 0.72±0.02     | 0.75±0.02      | 0.77±0.02| 0.78±0.02              | 0.65±0.02       | 0.64±0.02              |
| T-P-D-LSTM     | 0.74±0.01           | 0.74±0.01     | 0.76±0.01      | 0.79±0.01| 0.80±0.02              | 0.66±0.01       | 0.65±0.01              |
| T-P-D-LSTM (Multi-task) | 0.69±0.02 | 0.67±0.01 | 0.74±0.01 | 0.75±0.01 | 0.76±0.02 | 0.64±0.02 | 0.63±0.01 |

D: decay; T: time-aware; P: phased.

a hyperparameter. It is very important for AP. Based on the performance of the validation dataset, the size of time window was tuned to determine the specific period of time for the target surgery. The time window ranging from 3 to 10 days were tested based on the best-performing model in the development cohort. Fig. 3 shows the accuracy and AUC from the best-performing model by tuning the time window from 3 to 10 days. When the time window was set as seven-day, the model performed best based on accuracy or AUC.

The number of hidden layer neurons was set to 200. The batch size was set to 350 and the learning rate was set to 0.001. Models were optimized using a gradient descent approach. The performance of models was an average of 100 epochs. We added dropout layer to avoid overfitting when training [47].

FIGURE 3. Exploratory analysis for the time window.

Preprocessing of data, SVM and RF were done with R (version 3.6.0), and RNN-based models were done with Python (version 3.5).

D. TEST RESULTS

1) PERFORMANCE OF THE DEVELOPMENT COHORT

In a single prediction task, standard LSTM and GRU obtained higher area under the ROC curve (AUC) in testing than SVM and RF. With the addition of time-aware, the AUC of testing increased compared to standard LSTM and GRU. Phased-Decay LSTM (P-D-LSTM) had the highest AUC of testing at 0.76. The AUC of testing from multi-task prediction also reached 0.72. Time-aware Phased-Decay-LSTM had the best sensitivity and second-best specificity of testing following T-D-GRU. Table 1 also shows the standard deviation of performance.

2) VISUALIZATION OF AP FROM THE ONSET TO DISCHARGE

We visualized the changes in the patient’s status since the onset of AP. A patient case was taken as an example. Fig. 4 shows the predicted possibility of surgery for the AP patient from onset to discharge based on the laboratory indicators using time-aware Phased-Decay LSTM. The estimated probability for patients who didn’t have surgery should be less than 0.5, while the estimated probability for patients who have undergone surgery should be higher than 0.5. As seen in Fig. 4, our developed model learned the changes in the laboratory indicators to accurately predict the timing of the specific surgical intervention.

3) TIME-AWARE

Fig. 5 shows the weight mode / contribution mode of clinical features over time from time-aware Phased-Decay LSTM in
FIGURE 5. Contribution mode of clinical features with time, where $a_1 \sim a_{11}$ represent liver function, pancreatic function, infection index, pH balance, coagulation, tumor index, kidney function, anemia index, thrombosis index, intestinal function and heart function, respectively.

E. EXTERNAL VALIDATION

We used the Cerner Health Facts database to predict the timing of any surgical intervention in AP to validate the performance of our developed model externally. At first, based on the ICD code, patients diagnosed with AP on admission (142,650 patients) were extracted. Due to the differences in surgical records between EHR of the Cerner Health Facts database and EMR of West China Hospital of Sichuan University, we extracted the patients with AP who had any surgical intervention experience in the Cerner Health Facts database. A total of 1,597 patients were included in the external validation dataset. The data processing in the external validation dataset was the same with that in the development dataset.

By using our developed model, the AUC can reach 0.71 to predict the timing of any surgical intervention in AP in the external validation cohort, close to the performance of the development cohort, where the performance also got about 0.70 based on the data from Northeast, Midwest, South and West region of the Cerner Health Facts database (Table 2).

TABLE 2. Performance of the external validation cohort.

| T-P-D-LSTM | Accuracy Validation | Test | AUC Validation | Test |
|------------|---------------------|------|---------------|------|
| All        | 0.69±0.01           | 0.67±0.02 | 0.73±0.02      | 0.71±0.02 |
| Northeast  | 0.69±0.02           | 0.68±0.02 | 0.72±0.02      | 0.72±0.02 |
| Midwest    | 0.69±0.03           | 0.67±0.02 | 0.73±0.02      | 0.71±0.03 |
| South      | 0.69±0.03           | 0.66±0.05 | 0.73±0.03      | 0.69±0.06 |
| West       | 0.69±0.02           | 0.67±0.02 | 0.73±0.02      | 0.71±0.02 |

T: time-aware; P: phased; D: decay.

V. DISCUSSION

This study developed a novel model, time-aware Phased-Decay LSTM, which could give the importance of clinical features with time in acute disease. For the first time, the novel model based on RNN was applied to the timing prediction of the specific surgical intervention in AP. And by using the model, the specific surgical process of AP from the onset to discharge could be monitored.

Because LSTM can forget and update the long-term state for better prediction, while SVM and RF contain redundant information, which affects performance. LSTM performed better than SVM and RF, which illustrates the advantages of time series models in processing time series data. After adding time-aware, the performance of the model was better compared to standard LSTM and GRU because AP is time-sensitive. In addition to predicting the timing of surgical intervention, the novel model can also predict clinical features that surgeons believe are important.
The performance of our developed model was not as good as that of Phased-Decay LSTM when predicting timing, but it was close. It shows that phase gate function of Phased-Decay LSTM can effectively use time information, however, it was hard to explain. By using time-aware, we got the importance of clinical features in the time dimension for the specific surgical intervention. It indicates that in the developing course of AP, at different time points, we can focus on the more important features. For example, heart function contributed the most at the onset and in the first week.

Based on the data, we found the seven-day-time window was the best prediction interval because if the frequency of laboratory tests monitoring was low in EMR, it cannot be accurately used to predict in a narrow window, and some laboratory indicators may have a large change of value after one week which makes it difficult to predict accurately in an excessively long window. The specific surgical process of patients with AP from the onset to discharge was monitored dynamically for the first time through seven-day-time window.

The prediction results for some patients were not accurate. Due to the asynchronous, unequal length, and heterogeneity of the data, it is very challenging to monitor the patients with AP from the onset to complications to surgical intervention to discharge in real time based on EMR, which takes time. Thus, the possible reason for the failure is the inclusion of laboratory information only. To some extent, it illustrates that our developed model is useful for predicting the timing of the specific surgical intervention in AP. By using this patient-specific model, whether a patient with AP needs a specific surgical intervention in the next week can be determined in advance. It would be a great help when making surgical decisions in AP. The Cerner Health Facts database validated the generalization ability of our developed model.

There are some limitations in this work. We collected patients’ data for necrotizing pancreatitis retrospectively, rather than prospectively. There may be some misclassification. It is the first time to use RNN-based model to predict the timing of surgical intervention in necrotizing pancreatitis and there were some failure prediction cases, which may be related to the lack of including information (laboratory information only). Next, we will collect more multimodal data of AP, like clinical notes, images and EMR to improve the performance of our model.

VI. CONCLUSION
In this study, the developed novel model, time-aware Phased-Decay LSTM, could extract the contribution modes of clinical features with time that was important for surgical intervention and monitor the surgical process from the onset of AP to discharge in necrotizing pancreatitis.

ACKNOWLEDGMENT
The authors would like to thank Hua Xu and Yujia Zhou in the School of Biomedical Informatics, The University of Texas Health Science Center at Houston, for providing support of extracting data from the Cerner Health Facts database.

REFERENCES
[1] D. Yadav and A. B. Lowenfels, “The epidemiology of pancreatitis and pancreatic cancer,” Gastroenterology, vol. 144, no. 6, pp. 1252–1261, May 2013, doi: 10.1053/j.gastro.2013.01.060.
[2] P. G. LankischM Apte and P. A. Banks, “Acute pancreatitis,” Lancet, vol. 386, no. 9988, pp. 85–96, 2015, doi: 10.1016/S0140-6736(14)60649-8.
[3] A. F. Peery, E. S. Dellon, J. Lund, S. D. Crockett, C. E. McGowan, W. J. Bulsiewicz, L. M. Gangarosa, M. T. Thiny, K. Sitzenberg, D. R. Morgan, and Y. Ringel, “Burden of gastrointestinal disease in the United States update,” Gastroenterology, vol. 143, no. 5, pp. 1179–1187, 2012, doi: 10.1053/j.gastro.2012.08.002.
[4] K. Satoh, T. Shimosegawa, A. Masumune, M. Hirota, K. Kikuta, Y. Kihara, S. Kuriyama, I. Tsuji, A. Satoh, and S. Hamada, “Nationwide epidemiological survey of acute pancreatitis in Japan,” Pancreas, vol. 40, no. 4, pp. 503–507, May 2011, doi: 10.1097/MPA.0b013e318214812b.
[5] H.-N. Chen, C.-L. Lu, and C.-Y. Li, “Epidemiology of first-attack acute pancreatitis in Taiwan from 2000 through 2009,” Pancreas, vol. 41, no. 1, pp. 1127–1132, 2012, doi: 10.1097/MPA.0b013e31825bd4d1.
[6] D. Yadav and D. C. Whitcomb, “The role of alcohol and smoking in pancreatitis,” Nature Rev. Gastroenterology Hepatology, vol. 7, no. 3, pp. 131–145, Mar. 2010, doi: 10.1038/nrgastro.2010.6.
[7] A. Unalp-Arida and C. E. Ruhl, “The burden of acute pancreatitis in the United States population,” Gastroenterology, vol. 152, no. 5, 2017, Art. no. 5289, doi: 10.1016/j.gastro.2017.03.055.
[8] R. Pezzilli, A. Zerbi, D. Campra, G. Capurso, R. Golfieri, M. Yokoe et al., “Japanese guidelines for the management of acute pancreatitis,” Pancreatology, vol. 13, no. 4, pp. e1-e15, 2013, doi: 10.1016/j.pan.2013.07.063.
[9] C. E. Forstmark and J. Baillie, “AGA institute technical review on acute pancreatitis,” Gastroenterology, vol. 132, no. 5, pp. 2022–2044, May 2007, doi: 10.1053/j.gastro.2007.03.065.
[10] S. J. Pandol, A. K. Saluja, C. W. Imrie, and P. A. Banks, “Acute pancreatitis: Bench to the bedside,” Gastroenterology, vol. 132, no. 3, pp. 1127–1151, Mar. 2007, doi: 10.1053/j.gastro.2007.01.055.
[11] E. P. Dellinger, C. E. Forstmark, P. Lévy, E. Maravi-Poma, M. S. Petrov, T. Shimosegawa, A. K. Siriwijardena, G. Uno, D. C. Whitcomb, and J. A. Windsor, “Determinant-based classification of acute pancreatitis severity,” Ann. Surgery, vol. 256, no. 6, pp. 875–880, Dec. 2012, doi: 10.1097/SLA.0b013e318256f778.
[12] S. Tenner, J. Baillie, J. DuWitt, and S. V. Sege, “American college of gastroenterology guideline: Management of acute pancreatitis,” Amer. J. Gastroenterology, vol. 108, no. 9, pp. 1400–1415, Sep. 2013, doi: 10.1038/ajg.2013.218.
[13] U.K. Working Party on Acute Pancreatitis, “UK guidelines for the management of acute pancreatitis,” Gut, vol. 54, no. 3, pp. 31–39, May 2005, doi: 10.1136/gut.2004.057036.
[14] R. Pezzilli, A. Zerbi, D. Campra, G. Capurso, R. Golfiere, P. P. Aricidiascono, P. Bills, G. Buttunini, L. Calccoli, R. Cannizzaro, S. Carrara, S. Crippa, R. De Gaudio, P. De Rai, L. Frulloni, E. Mazza, M. Mutignani, N. Pagano, P. Rabitti, and G. Balzano, “Consensus guidelines on severe acute pancreatitis,” Ann. Surg., vol. 256, no. 6, pp. 875–880, Dec. 2012, doi: 10.1097/SLA.0b013e318256f778.
[15] M. Yokoe et al., “Japanese guidelines for the management of acute pancreatitis: Japanese guidelines 2015,” J. Hepato-Biliary-Pancreatic Sc., vol. 22, no. 6, pp. 405–432, Jun. 2015, doi: 10.1002/jbhp.259.
[16] Q. Guo, A. Li, Q. Xia, X. Liu, B. Tian, G. Mai, Z. Huang, G. Chen, W. Tang, X. Jin, and W. Chen, “The role of organ failure and infection in necrotizing pancreatitis: A prospective study,” Ann. Surg., vol. 259, no. 6, pp. 1201–1207, 2013.
[17] N. J. Schepers, O. J. Bakker, M. G. Besselink, U. A. Ali, T. L. Bollen, H. G. Gooszen, H. C. Van Santvoort, and M. J. Bruno, “Impact of characteristics of organ failure and infected necrosis on mortality in necrotising pancreatitis,” Gut, vol. 68, no. 6, pp. 1044–1051, 2018, doi: 10.1136/gutjnl-2017-314657.
[18] I. P. Gomatos, C. M. Halloran, P. Ghanem, M. G. Ratary, F. Polydoros, J. C. Evans, H. L. Smart, R. Yagati-Satchidanand, J. M. Garry, P. A. Whelan, and F. E. Hughes, “Outcomes from minimal access retroperitoneal and open pancreatic necrosectomy in 394 patients with necrotizing pancreatitis,” Ann. Surg., vol. 263, no. 5, pp. 992, 2015.
LI PENG received the B.S. degree in medical imaging technology from Sichuan University, in 2018, where she is currently pursuing the M.S. degree in medical informatics with the West China Hospital/West China School of Medicine. Her research interests include medical image processing, medical data analysis, and medical artificial intelligence.

MENGJIAO LI received the B.S. degree in medicine from Zhengzhou University, in 2013, and the M.S. degree from Sichuan University, in 2016, where she is currently pursuing the Ph.D. degree with the West China Hospital/West China School of Medicine. Her research interests include study of pathogenic factors, law of disease progression of diseases with EMR data using statistics, and machine learning methods.

HUIMIN LU received the bachelor’s degree in clinical medicine, the master’s degree in general surgery, and the Ph.D. degree in general surgery from Sichuan University, in 2004, 2007, and 2010, respectively.

He is currently an Attending Surgeon with the Department of Pancreatic Surgery, West China Hospital/West China School of Medicine, Sichuan University. His research interest includes surgical treatment of pancreatitis.

DUJIANG YANG received the master’s degree in surgery from Sichuan University, in 2019, where he is currently pursuing the Ph.D. degree in surgery with the West China Hospital/West China School of Medicine. His research interests include treatment of acute pancreatitis and pancreatic tumor.

XIAOYAN YANG received the M.S. degree in epidemiology and health statistics from the West China School of Public Health, Sichuan University, Chengdu, China, in 2005, and the Ph.D. degree in evidence-based medicine and clinical epidemiology from the West China School of Medicine, Sichuan University, in 2009.

She carried out Postdoctoral Research in management science and engineering with Sichuan University, where she is currently an Associate Professor and a master’s Tutor with the West China Biomedical Big Data Center, West China Hospital/West China School of Medicine. Her research interests include statistical methodology, evidence-based decision making and management, data mining, and artificial intelligence.

XIAOBO ZHOU received the Ph.D. degree in applied mathematics from Peking University. He is currently a Professor and the Director of the Center for Systems Medicine, School of Biomedical and Bioinformatics, The University of Texas Health Science Center at Houston. His research interests include data sciences, imaging informatics, and clinical informatics.