Infection and Drug Resistance Dovepress

Open Access Full Text Article ORIGINAL RESEARCH

# Machine Learning Approach to Classify *Vibrio vulnificus* Necrotizing Fasciitis, Non-Vibrio Necrotizing Fasciitis and Cellulitis

Chia-Peng Chang[1,2], Kai-Hsiang Wu[1–3]

[1]Department of Emergency Medicine, Chiayi Chang Gung Memorial Hospital, Puzih City, Chiayi County, Taiwan; [2]Department of Nursing, Chang Gung University of Science and Technology, Chiayi Campus, Puzi City, Chiayi County, Taiwan; [3]Graduate Institute of Clinical Medical Sciences, College of Medicine, Chang Gung University, Taoyuan City, Taiwan

Correspondence: Kai-Hsiang Wu, Email dr5853@cgmh.org.tw



**Background:** Recent advancements in artificial intelligence have led to increased adoption of machine learning in disease identification, particularly for challenging diagnoses like necrotizing fasciitis and *Vibrio vulnificus* infections. This shift is driven by the technology's efficiency, objectivity, and accuracy, offering potential solutions to longstanding diagnostic hurdles in clinical practice.

**Methods:** This investigation incorporated 180 inpatients suffering from soft tissue infections. The participants were categorized into groups: cellulitis, non-Vibrio necrotizing fasciitis (NF), or *V. Vulnificus* NF. To predict the three relevant outcomes, we employed Light Gradient Boosting Machine (LightGBM) and 5-fold cross-validation methodologies for the development of a multi-class categorization model. Moreover, we applied the SHapley Additive exPlanations (SHAP) methodology to decipher the model's predictions.

**Results:** The multi-classification model possesses substantial predictive capacity, with a weighted-average AUC of 0.86, sensitivity of 87.2%, specificity of 74.5%, NPV of 81.6%, and PPV of 85.4%. The model's calibration was assessed using the Brier score, yielding a weighted mean of 0.084. This low value demonstrates a strong correlation between predicted probabilities and actual outcomes, indicating high predictive accuracy and reliability in the model's forecasts.

**Conclusions:** We effectively developed a multiclassification model aimed at forecasting the occurrence of cellulitis, non-Vibrio NF, or V. Vulnificus NF in patients suffering from soft tissue infection, and we further described the model's predictions using the SHAP algorithm.

**Keywords:** necrotizing fasciitis, machine learning, cellulitis, *Vibrio vulnificus*

## Background

Necrotizing fasciitis (NF) is a severe soft tissue infection characterized by rapid, progressive destruction of muscle fascia and surrounding tissues.[1] Type 3 NF, caused by Vibrio vulnificus, is particularly aggressive, with research suggesting its progression outpaces even Staphylococcus aureus infections.[2] This infection typically results from wounds acquired during seafood handling, seawater exposure, or consumption of inadequately cooked, contaminated seafood. Immunocompromised individuals face heightened risk.[3–5] Optimal management involves immediate, extensive debridement coupled with broad-spectrum antibiotic therapy.[6] Despite prompt intervention, serious complications including amputation, multiorgan failure, and high mortality rates remain significant concerns.[7–9]

Rapid differentiation between necrotizing fasciitis (NF) and non-necrotizing soft tissue infections, such as cellulitis, is clinically challenging. Early in the disease course, both conditions present with similar skin manifestations, including erythema, swelling, and pain. However, these two diseases have fundamentally different prognoses; NF has a significantly higher mortality rate. Due to the coastal location of our hospital and the large population of fishermen, the incidence of NF caused by marine Vibrio infections is much higher than in other regions. For frontline clinicians without extensive experience, the diagnosis of NF may be delayed, leading to missed opportunities for timely emergency surgery, thus negatively affecting patient outcomes. To further aid in differentiation, diagnostic tools can be employed.

Received: 28 August 2024
Accepted: 6 December 2024
Published: 11 December 2024

Laboratory findings in necrotizing fasciitis often show elevated white blood cell count, elevated creatinine, and a high C-reactive protein (CRP) level. The LRINEC (Laboratory Risk Indicator for Necrotizing Fasciitis) score, which incorporates these lab values along with hemoglobin, sodium, and glucose levels, is a useful tool for assessing the likelihood of NF. Imaging studies can also be informative. Ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI) can help distinguish between cellulitis and necrotizing fasciitis. However, definitive diagnosis of necrotizing fasciitis is typically established through surgical exploration, revealing a lack of resistance to blunt dissection.[9,10] *V. vulnificus* NF is predominantly caused by a Gram-negative, halophilic bacterium. This organism thrives in warm, saline environments and is frequently associated with exposure to seawater or consumption of raw or undercooked seafood. Individuals with underlying conditions like liver disease, hemochromatosis, diabetes, and immunosuppression are particularly susceptible to *V. vulnificus* infections. Non-Vibrio NF, in contrast, is typically caused by other bacteria, such as Group A Streptococcus (GAS), Staphylococcus aureus, or various anaerobic bacteria, often in a polymicrobial fashion. Unlike *V. vulnificus* NF, the non-Vibrio form of the disease does not require a specific environmental exposure and can result from minor skin traumas, surgical incisions, or even non-apparent sources.[11,12] Studies have shown that NF caused by Gram-negative bacteria is associated with a worse prognosis compared to the more common Gram-positive infections. Therefore, we aim to leverage artificial intelligence to assist in differentiating *Vibrio* NF, non-*Vibrio* NF, and cellulitis.

Past research has demonstrated machine learning (ML),which is extensively utilized in diagnosing and prognosticating diseases, exhibiting a performance that exceeds traditional linear models. ML in medical diagnosis refers to the use of algorithms and computational models that improve their performance or "learn" over time through experience. This field is part of the larger realm of artificial intelligence, where intelligent machines are designed to mimic human intelligence. The incorporation of ML in medical diagnosis can improve efficiency, accuracy, and predictability. However, it's important to note that these systems should be used to aid health-care professionals, not replace them. Also, issues related to data privacy, security, and the ethical use of AI in healthcare are significant considerations as this technology continues to advance.[13–15] For example, a particular research leveraged a decision tree algorithm, a basic ML technique, to anticipate the amputation risk in patients with diabetic foot ulcer. Although it yielded certain results, the decision tree algorithm is prone to overfitting and fails to consider the interrelation between the features included, thereby necessitating more sophisticated and intricate models.[16] Nevertheless, these advanced models frequently involve highly complex and computationally demanding processes that may result in a lack of interpretability of the results. Such models are often perceived as "black boxes", leading to their limited acceptance among medical professionals.[17]

In this study, our objective was to formulate a machine learning model capable of forecasting the occurrence of *V. Vulnificus* NF, non-*Vibrio* NF and cellulitis. In addition, we utilized model-agnostic techniques to seek better efficiency of complex, opaque models, aiming to enhance the trust from medical provider in employing these models. This also paves the way for innovative strategies for personalized evaluation of risk elements associated with the progression of NF from soft tissue infection.

# Methods

## Study Design and Participants

This study, a retrospective chart analysis, was conducted at Chiayi Chang Gung Memorial Hospital with approval from the institution's Institutional Review Board (No.: 202302020B0). The research examined patient records from April 2015 to March 2018, focusing on cases meeting specific criteria: (1) Surgically confirmed NF with corresponding discharge diagnosis (2) Primary discharge diagnosis of cellulitis (3) Management initiated in the hospital's emergency department.

From this cohort, we identified three distinct groups: Vibrio vulnificus NF group, Non-Vibrio NF group and Cellulitis group. We employed a matching strategy to compare these group (1)V. vulnificus NF to cellulitis: 1:4 ratio (2) V. vulnificus NF to non-Vibrio NF: 1:1 ratio. Matching criteria included gender, admission year, and age (within a 5-year range). The V. vulnificus NF group was defined by positive V. vulnificus cultures from blood, wound, or both. This methodology allowed for a comprehensive comparison between V. vulnificus NF cases and other related conditions, providing insights into their distinct clinical characteristics and outcomes.

## Data Collection

We reviewed of electronic health records and analyzed various patient data, including: demographic information (age, gender), existing medical conditions, environmental exposure history (seawater, contaminated farm water), animal bite incidents, initial vital signs recorded during ED triage, blood test results obtained within the first hour of ED visitation. The definitive diagnosis of NF was confirmed through surgical pathology reports. Patients were categorized into the cellulitis group if they lacked pathological confirmation of NF or did not undergo surgical intervention. This methodical approach allowed for a rigorous classification of cases, ensuring accurate differentiation between NF and cellulitis for our analysis.

## Statistical Analysis

For normally distributed variables, as determined by the Shapiro–Wilk test, we employed *t*-tests and ANOVA with linear polynomial contrasts. These methods allowed for pairwise comparisons and trend analyses between groups. Variables with non-normal distributions were analyzed using Kruskal–Wallis and Jonckheere-Terpstra tests to assess inter-group differences and trends, respectively. Categorical variables were presented as counts (n) and percentages (%). We utilized Chi-squared tests for group comparisons and linear-by-linear association Chi-squared tests for trend analyses. Statistical significance was set at $p < 0.05$. This comprehensive approach enabled a thorough examination of data distributions, group differences, and trends across variables.

## Model Development

The Light Gradient Boosting Machine (LightGBM) is a gradient boosting framework that uses tree-based learning algorithms. It is developed by Microsoft and is designed to be a fast, scalable and efficient algorithm, particularly for large-scale data. The "Light" in LightGBM refers to its high speed. LightGBM can handle large size data and takes lower memory to run. It is capable of performing equally well on large datasets with a significant reduction in training time compared to other boosting algorithms. The algorithm reduces overfitting through regularization and offers a number of tunable hyperparameters to optimize performance and accuracy. One of the key differentiators of LightGBM is its Exclusive Feature Bundling, a technique for handling sparse data, which is common in real-world datasets. LightGBM offered a powerful and precise approach for managing multi-category classification challenges. This method was applied in our research to build a predictive model for *V. vulnificus* NF, non-*Vibrio* NF, and cellulitis.[18]

The data collection was randomly allocated into two categories: 80% for training and 20% for testing purposes. The main function of the training data was to facilitate the development of the model and the fine-tuning of the model's hyperparameters. To achieve this, we implemented a method of 5-fold cross-validation. The training data were arbitrarily divided into five subsets of equal sizes. Each of these subsets served as the testing set at some point while the remaining subsets were used for model training. An efficient hyperparameter tuning technique known as Bayesian hyperparameter optimization, which utilizes the tree-Parzen estimator that takes into account past tuning data, was employed to identify the best hyperparameter in the process of cross-validation.[15]

## Model Evaluation and Explanation

Model performance was evaluated across all categories using a comprehensive set of metrics: area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV).

These five indicators collectively provided a thorough assessment of the model's discriminative ability, accuracy, and predictive power across various clinical scenarios. The aggregate performance of the model was showcased by the weighted mean of each category's performance. Each category's weight corresponded to its share of the total sample count.

The SHapley Additive exPlanations (SHAP) algorithm is a well known technique in machine learning field. It is a unified measure of feature importance that allocates the contribution of each feature to the prediction for each instance. SHAP values interpret the impact of having a certain value for a given feature in comparison to the prediction we'd make if that feature took some baseline value. SHAP has many practical advantages. It is model-agnostic, meaning it can be

applied to any machine learning model, and it also offers high interpretability. It provides a way to understand the complex decision-making processes of models by attributing the change in the output of the model to the corresponding change in input feature values. Therefore, SHAP is often used to interpret the predictions of complex models, as well as to create interpretable models.[19] In order to gauge the effect of specific patient variables on the predictions, SHAP values were employed. The prediction value for a particular patient is computed by adding the average prediction value from the model's training set to the total of the SHAP values corresponding to that patient's variables. The whole process of model creation, assessment, and interpretation was executed using conventional Python libraries.

# Results

## Statistical Analysis

Based on the inclusion criteria mentioned above, this study initially included 43 cases of Vibrio NF, 61 cases of non-Vibrio NF, and 426 cases of cellulitis. Referring to other literature, many studies use ratios of 1:2 or 1:4 as their standard, so we chose a ratio of 1:4 between NF and cellulitis. In addition, to meet the strict matching strategy and to exclude patients with missing data, a 1:1 ratio was selected between Vibrio NF and non-Vibrio NF. Therefore, the final result included a total of 180 patients. The cohort's mean age was 67 years, with males constituting 63.3% of the sample. Table 1 presents a comparative analysis of clinical features and outcomes across these groups, following 1:1 and 1:4 matching protocols. NF groups demonstrated higher prevalence of liver and kidney comorbidities, fever, hypotension, bacteremia, and history of seawater or seafood exposure. They also exhibited elevated serum lactate and C-reactive protein (CRP) levels. Interestingly, among LRINEC score components, only CRP showed significant differentiation between V. vulnificus NF, non-Vibrio NF, and cellulitis groups. Other LRINEC variables did not demonstrate significant discriminatory power across these groups.

**Table 1** Clinical Characteristics of V. Vulnificus,Non-Vibrio NF, and Cellulitis Groups

| | ***V. Vulnificus* NF (n=30)** | **Non-*Vibrio* NF (n=30)** | **Cellulitis (n=120)** | **p-value** |
|---|---|---|---|---|
| **Age** (years, mean ± SD) | 67 (57–76) | 67 (56–76) | 67 (56–75) | 0.99 |
| **Sex (Male)** | 19 (63.3%) | 19 (63.3%) | 76 (63.3%) | 1.0 |
| **Comorbidities** | | | | |
| Diabetes mellitus | 7 (23.3%) | 9 (30%) | 27 (22.5%) | 0.67 |
| Liver disease | 12 (40.0%) | 10 (33.3%) | 34 (28.3%) | 0.03* |
| Malignancy | 2 (6.7%) | 2 (6.7%) | 10 (8.3%) | 0.89 |
| Kidney disease | 8(26.7%) | 9 (30.0%) | 23 (19.2%) | 0.04* |
| Peripheral vascular disease | 2 (6.7%) | 3 (10.0%) | 6 (5.0%) | 0.12 |
| **Variables on admission** | | | | |
| Temperature > 38.0°C | 7 (23.3%) | 8 (26.7%) | 17 (14.1%) | 0.06 |
| Systolic pressure < 90 mmHg | 7 (23.3%) | 3 (10.0%) | 8 (6.7%) | < 0.01* |
| Sea water or seafood contact | 25 (83.3%) | 4 (13.3%) | 11 (9.2%) | < 0.01* |
| **Presence with bacteremia** | 19 (63.3%) | 6 (20.0%) | 9 (7.5%) | < 0.01* |
| **LRINEC variables**(mean) | | | | |
| CRP (mg/L) | 73.5 (32.6–165.1) | 88.3 (10.2–215.6) | 28.6 (13.7–81.5) | <0.01* |
| WBC (cells/mm$^3$) | 13,574 (6120–22,479) | 15,415 (6153–25,650) | 8995 (7365–13,850) | 0.06 |
| Hemoglobin (g/dL) | 12.5 (11.1–13.3) | 12.6 (9.8–13.9) | 13.1 (10.6–14.8) | 0.82 |
| Blood glucose (mg/dL) | 154 (139–205) | 180 (92–214) | 147 (95–185) | 0.47 |
| Sodium (mmol/L) | 134 (131–137) | 136 (131–142) | 135 (134–139) | 0.53 |
| Creatinine (mg/dL) | 1.42 (0.88–2.24) | 1.35 (0.92–2.88) | 1.15 (0.52–2.24) | 0.74 |
| **Lactate (mg/dL)** | 18.8 (9.2–20.5) | 13.8 (8.3–20.8) | 9.5 (5.6–17.8) | <0.01* |

**Note**: *P < 0.05.
**Abbreviations**: IQR, interquartile ratio; LRINEC, Laboratory Risk Indicator for Necrotising Fasciitis; NF, necrotizing fasciitis; WBC, white blood cell; *Vibrio, Vibrio vulnificus*.

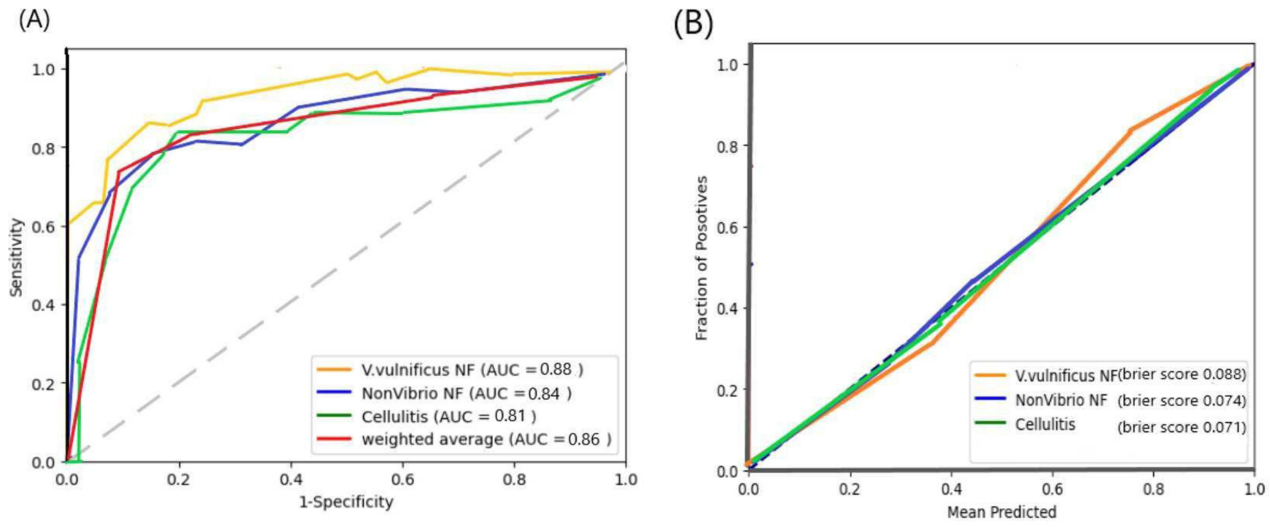


**Figure 1** Discrimination and calibration performance of the multi-class classification model. (**A**) Receiver-operating-characteristic curves (ROC) for each class and their weighted average ROC curve. (**B**) Calibration curves for each class of the model.

## Model Performance

The model that was developed showcased its predictive prowess as determined by the AUC (0.5 = an entirely arbitrary prediction and 1 = flawless distinction) in the testing group, as follows: *V. Vulnificus* NF (0.88), non-*Vibrio* NF (0.84), cellulitis (0.81) (Figure 1A). Further evaluation of the model's efficacy is presented in Table 2. These observed results indicate that the multi-classification model possesses substantial predictive capacity, with a weighted-average AUC of 0.86, 95% CI (0.81–0.90);

Sensitivity of 87.2%,95% CI (80.0%-92.5%); specificity of 74.5%, 95% CI (67.0%-81.5%); NPV of 81.6%, 95% CI (74.0%-88.0%) and PPV of 85.4%, 95% CI (78.0%-91.5%). Furthermore, the calibration curve for every classification aligns closely with the 45-degree line, and the weighted mean of the Brier score stood at 0.084. This indicates that the model's forecasted probability closely matches the actual observed probability (Figure 1B).

## Model Interpretation

The SHAP algorithm effectively demonstrates the method for deriving the predicted probability using the initial risk and the individual patient's features. For example, consider a patient with *V. Vulnificus* NF; the model predicted the likelihood of cellulitis, non-*Vibrio* NF, and *V. Vulnificus* NF as 0.29, 0.23, and 0.48 respectively (Figure 2). The baseline risks related to this patient for cellulitis, non-*Vibrio* NF, and *V. Vulnificus* NF stood at 0.72, 0.13, and 0.15 correspondingly, values arrived at by averaging the prediction outcomes of the model samples from the training set.

Our analysis of predictive factors for *V. Vulnificus* NF revealed several key influences on disease probability:

**Table 2** Performance of the Multi-Classification Model in Test Set

| Metrics | *V. Vulnificus* NF (N = 6) | Non-*Vibrio* NF (N = 6) | Cellulitis (N = 24) | Overall |
|---|---|---|---|---|
| Sensitivity | 35.8% | 61.2% | 93.2% | 87.2% |
| Specificity | 96.5% | 93.7% | 70.1% | 74.5% |
| NPV | 94.7% | 94.1% | 74.3% | 81.6% |
| PPV | 49.9% | 59.8% | 93.5% | 85.4% |

**Abbreviations**: NPV, negative predictive value; PPV, positive predictive value; NF, necrotizing fasciitis; *Vibrio, Vibrio vulnificus.*

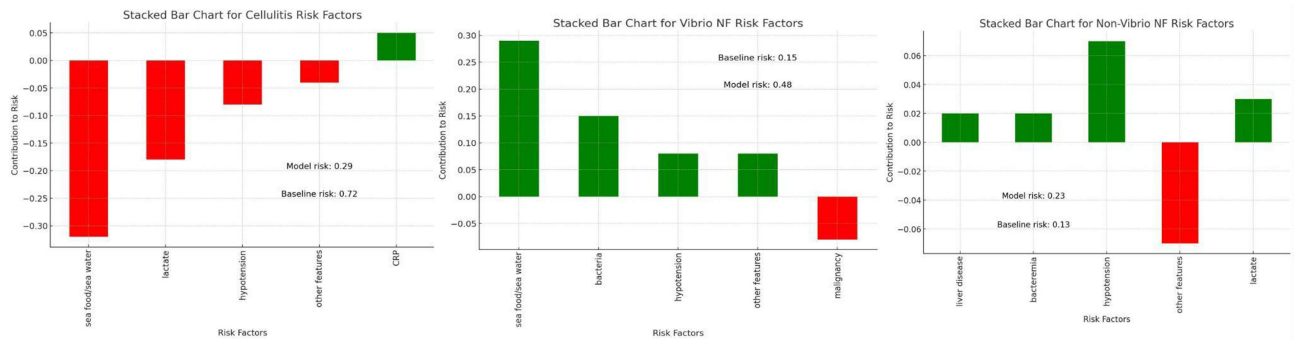


**Figure 2** Illustrative example of SHAP algorithm for interpreting the developed model. Illustrative example of how baseline risk and patient characteristics constitute the risk of cellulitis, non-*Vibrio* NF, or *V. Vulnificus* NF predicted by the model. The baseline risk was obtained by calculating the average value of the prediction of the model in the training set samples.

1. Seawater/seafood exposure increased probability by 0.29.
2. Hypotension contributed an additional 0.08.
3. Malignancy showed a negative association, decreasing probability by 0.08.
4. Bacteremia presence elevated probability by 0.15.
5. Other features collectively added 0.08 to the probability.

This probabilistic model provides insights into the relative importance of various clinical factors in predicting *V. Vulnificus* NF. Similar interpretations can be applied to the probability models for cellulitis and non-Vibrio NF, offering a comparative framework for differential diagnosis.

## Discussion

V. vulnificus infections primarily manifest as gastrointestinal disorders, septicemia, and wound infections, progressing rapidly due to toxins, leading to hemorrhagic bullae and severe skin necrosis. Data from two hospitals in Southern Taiwan showed a 13.3% in-hospital mortality rate for V. vulnificus NF, consistent with previous reports of 10–13%. This study underscores the severity of V. vulnificus infections and the importance of prompt diagnosis and treatment.[20–23]

Namany et al demonstrated that patients with monomicrobial NF had significantly poorer outcomes, including higher 90-day mortality, increased ICU admissions, and greater need for vasopressor support. These findings highlight the virulence of monomicrobial infections, especially gram-negative organisms, underscoring the need for rapid pathogen identification and targeted treatment strategies in managing severe soft tissue infections.[24]

Previous studies have shown a strong link between hyperlactatemia and in-hospital mortality in NF patients, independent of acidosis.[19,20] Our research extends this relationship to V. vulnificus infections. While the LRINEC score is widely studied for NF diagnosis,[10] our findings suggest it is more useful as a prognostic indicator for V. vulnificus NF, with higher scores (>8) correlating with increased mortality. These results highlight the complex nature of V. vulnificus NF and the need for multifaceted approaches in diagnosis and prognosis.

The Brier score is a metric used in statistics and machine learning to evaluate the accuracy of probabilistic predictions by measuring the mean squared difference between predicted probabilities and actual outcomes. It ranges from 0 to 1, with lower values indicating better performance. A Brier score of 0 represents perfect prediction, while 1 indicates complete inaccuracy. It is widely used in fields like weather forecasting, sports analytics, and medical diagnostics. Though useful, the Brier score has limitations, and other metrics should also be considered for comprehensive model evaluation.[24]

In this study, we used LightGBM to construct a multi-class classifier to estimate the likelihood of cellulitis, non-Vibrio NF, or V. vulnificus NF in patients with soft tissue infections. The model's AUC and Brier score were 0.86 and 0.084, indicating strong discrimination and alignment between predicted and observed probabilities. The SHAP algorithm was used to interpret the influence of patient attributes on predictions. Despite various diagnostic aids, such as lab

tests, LRINEC scores, and imaging studies, prompt identification of NF remains challenging due to the lack of a definitive diagnostic test. This highlights the need for a multifaceted diagnostic approach combining clinical, laboratory, and imaging data.

The progression of ML techniques provides new ways to address the limitations of linear models. We used LightGBM to create a multi-category classification model due to its efficiency and scalability. LightGBM employs techniques like Gradient-based One-Side Sampling and Exclusive Feature Bundling, making it faster and better at handling large datasets compared to other gradient boosting algorithms. Its leaf-wise tree growth strategy enhances accuracy, particularly for tabular datasets, where tree-based models generally outperform neural networks. Deep learning models are more suitable for computer vision and natural language processing tasks.[25,26]

Besides prediction, model interpretability is crucial. Despite the good predictive ability of the model, it was labeled as a "black box" due to its complex, high-dimensional calculations, reducing doctors' trust. To address this, we used the SHAP algorithm to provide comprehensible insights into NF risk, enhancing interpretability and physician trust.[27–30] SHAP helped illustrate that factors like exposure to seawater, hypotension, and bacteremia increased NF probability, while reducing cellulitis likelihood. Consistent with previous studies, higher serum lactate and CRP also raised NF risk. Elevated CRP indicates inflammation, and high lactate suggests poor tissue oxygenation, often seen in severe infections like septic shock.[31–34]

The developed explainable ML model yielded strong results. We utilized a tree-based ML technique known for its speed, accuracy, and effective performance. The multi-category prediction model successfully estimated the likelihoods of cellulitis, non-Vibrio NF, and V. vulnificus NF in patients with soft tissue infections.

Despite the retrospective nature of our research, it had limitations. We employed internal cross-validation in the model's creation, yet it lacked external validation groups and clinical substantiation. Small sample size may lead to uncertainties or increased risk of bias in the performance of the machine learning model. However, machine learning may have advantages in handling complex non-linear relationships. We need to collect more data to evaluate the performance differences between machine learning models and traditional linear models in practical applications.

Although our model identified numerous factors related to the clinical outcomes, the causal relationships among these factors remain largely unknown. This is an important limitation, as identifying causal factors could significantly impact clinical management and decision-making. The concept of causal inference, as discussed by Zhang et al[35] provides a framework to better understand these relationships and determine which variables may play a direct role in influencing outcomes. Future research should consider employing methods such as causal inference to validate whether the relationships observed in this study have causal significance, which could lead to more targeted and effective clinical interventions.

Furthermore, while the SHAP algorithm was capable of discerning the impact of patient attributes on the forecasts, it merely established a mapping correlation between the predictive variables and outcomes, not causality. It's important to note that while LightGBM has many advantages, it may not be the best choice for every situation. It also has certain limitations including not suitable for small datasets, sensitive to overfitting with noise, lack of community, requires for tuning. While LightGBM can handle missing values, it does not necessarily handle them in the most sophisticated way. It treats missing values as separate categories, which might not always be the most appropriate way of dealing with them. No single algorithm is the best for every situation. The choice of algorithm depends on the specific requirements of tasks, including the size and nature of datasets. Although its shortcomings, our research marked the firstly development of an interpretable machine learning model that predicted soft tissue infection in patients with adequate precision.

## Conclusion

In sum, we effectively developed a multiclassification model aimed at forecasting the occurrence of cellulitis, non-Vibrio NF, or V. Vulnificus NF in patients suffering from soft tissue infection, and we further described the model's predictions using the SHAP algorithm. Our investigation demonstrated that this interpretable machine learning model not only possessed superior forecasting capability but also introduced a novel framework for personalized assessments of patients' risk factors.

## Abbreviations

AUC, Area under the ROC curve; CRP, C-reactive protein; CT, Computed tomography; ED, Emergency department; LightGBM, Light gradients boosting machine; LRINEC, Laboratory Risk Indicator for Necrotizing Fasciitis; MRI, Magnetic resonance image; NF, Necrotizing fasciitis; PPV, positive predictive value; ROC, Receiver-operator characteristic; SHAP, SHapley Additive exPlanations.

## Data Sharing Statement

Please contact the corresponding author for data requests.

## Ethics Approval and Consent to Participate

This study was conducted in accordance with the principles of the Declaration of Helsinki. Ethical approval was obtained from the Institutional Review Board of Chang Gung Memorial Hospital (No. 202302020B0). The IRB confirm that the data was anonymized or maintained with confidentiality. Consent to participate was not applicable and waived by Institutional review board of Chia-yi Chang Gung Memorial Hospital due to its a low risk retrospective medical chart review study.

## Acknowledgments

The authors would like to thank all the participants in this study and the support from Chiayi Chang Gung Memorial Hospital and National Science and Technology Council.

## Funding

This study was founded by National Science and Technology Council, NSTC 113-2221-E-182A-005 -, NMRPG6P0061.

## Disclosure

The authors declare that they have no competing interests.

## References

1. Wong CH, Chang HC, Pasupathy S, Khin LW, Tan JL, Low CO. Necrotizing fasciitis: clinical presentation, microbiology, and determinants of mortality. *J Bone Joint Surg Am*. 2003;85(8):1454–1460. doi:10.2106/00004623-200308000-00005
2. Tsai YH, Wen-Wei Hsu R, Huang KC, Huang TJ. Comparison of necrotizing fasciitis and sepsis caused by Vibrio vulnificus and Staphylococcus aureus. *J Bone Joint Surg Am*. 2011;93(3):274–284. doi:10.2106/JBJS.I.01679
3. Chang CP, Hsiao CT, Lin CN, Fann WC. Risk factors for mortality in the late amputation of necrotizing fasciitis: a retrospective study. *World J Emerg Surg*. 2018;13(1):45. doi:10.1186/s13017-018-0207-0
4. Hsu JC, Shen SH, Yang TY, Chen PH, Huang KC, Tsai YH. Necrotizing fasciitis and sepsis caused by Vibrio vulnificus and Klebsiella pneumoniae in diabetic patients. *Biomed J*. 2015;38(2):136–142. doi:10.4103/2319-4170.137767
5. Lee YC, Hor LI, Chiu HY, Lee JW, Shieh SJ. Prognostic factor of mortality and its clinical implications in patients with necrotizing fasciitis caused by Vibrio vulnificus. *Eur J Clin Microbiol Infect Dis*. 2014;33(6):1011–1018. doi:10.1007/s10096-013-2039-x
6. Headley AJ. Necrotizing soft tissue infections: a primary care review. *Am Fam Physician*. 2003;68(2):323–328.
7. Ozalay M, Ozkoc G, Akpinar S, Hersekli MA, Tandogan RN. Necrotizing soft-tissue infection of a limb: clinical presentation and factors related to mortality. *Foot Ankle Int*. 2006;27(8):598–605. doi:10.1177/107110070602700806
8. Roje Z, Roje Z, Mati D, Librenjak D, Dokuzovi S, Varvodi J. Necrotizing fasciitis: literature review of contemporary strategies for diagnosing and management with three case reports: torso, abdominal wall, upper and lower limbs. *World J Emerg Surg*. 2011;6(1):46. doi:10.1186/1749-7922-6-46
9. Hakkarainen TW, Kopari NM, Pham TN, Evans HL. Necrotizing soft tissue infections: review and current concepts in treatment, systems of care, and outcomes. *Curr Probl Surg*. 2014;51(8):344–362. doi:10.1067/j.cpsurg.2014.06.001
10. Wong CH, Khin LW, Heng KS, Tan KC, Low CO. The LRINEC (Laboratory Risk Indicator for Necrotizing Fasciitis) score: a tool for distinguishing necrotizing fasciitis from other soft tissue infections. *Crit Care Med*. 2004;32(7):1535–1541. doi:10.1097/01.ccm.0000129486.35458.7d
11. Oliver JD. Wound infections caused by Vibrio vulnificus and other marine bacteria. *Epidemiol Infect*. 2005;133(3):383–391. doi:10.1017/S0950268805003894
12. Sarani B, Strong M, Pascual J, Schwab CW. Necrotizing fasciitis: current concepts and review of the literature. *J Am Coll Surgeons*. 2009;208(2):279–288. doi:10.1016/j.jamcollsurg.2008.10.032
13. Ruan Y, Bellot A, Moysova Z, et al. Predicting the risk of inpatient hypoglycemia with machine learning using electronic health records. *Diabetes Care*. 2020;43(7):1504–1511. doi:10.2337/dc19-1743
14. Aminian A, Zajichek A, Arterburn DE, et al. Predicting 10-year risk of end-organ complications of type 2 diabetes with and without metabolic surgery: a machine learning approach. *Diabetes Care*. 2020;43(4):852–859. doi:10.2337/dc19-2057

15. Armstrong DG, Rowe VL, D'Huyvetter K, Sherman RA. Telehealth-guided home-based maggot debridement therapy for chronic complex wounds: peri-and post-pandemic potential. *Int Wound J*. 2020;17(5):1490–1495. doi:10.1111/iwj.13425
16. Kasbekar PU, Goel P, Jadhav SP. A decision tree analysis of diabetic foot amputation risk in Indian patients. *Front Endocrinol*. 2017;8:25. doi:10.3389/fendo.2017.00025
17. Gastounioti A, Kontos D. Is it time to get rid of black boxes and cultivate trust in AI? *Radiol Artif Intell*. 2020;2(3):e200088. doi:10.1148/ryai.2020200088
18. Bergstra J, Bardenet R, Bengio Y, Kégl B Algorithms for hyperparameter optimization. Paper presented at: Proceedings of the 25th Annual Conference on Neural Information Processing Systems (NIPS 2011). 2011.
19. Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell*. 2020;2(1):56–67. doi:10.1038/s42256-019-0138-9
20. Hsiao CT, Weng HH, Yuan YD, Chen CT, Chen IC. Predictors of mortality in patients with necrotizing fasciitis. *Am J Emerg Med*. 2008;26(2):170–175. doi:10.1016/j.ajem.2007.04.023
21. Naamany E, Shiber S, Duskin-Bitan H, et al. Polymicrobial and monomicrobial necrotizing soft tissue infections: comparison of clinical, laboratory, radiological, and pathological hallmarks and prognosis. A retrospective analysis. *Trauma Surg Acute Care Open*. 2021;6(1):e000745. doi:10.1136/tsaco-2021-000745
22. Bair MJ, Chi H, Wang WS, Hsiao YC, Chiang RA, Chang KY. Necrotizing fasciitis in southeast Taiwan: clinical features, microbiology, and prognosis. *Int J Infect Dis*. 2009;13(2):255–260. doi:10.1016/j.ijid.2008.04.015
23. Chang CP, Fann WC, Wu SR, et al. Lactate on emergency department arrival as a predictor of in-hospital mortality in necrotizing fasciitis: a retrospective study. *J Orthop Surg Res*. 2019;14(1):73. doi:10.1186/s13018-019-1108-y
24. BRIER GW. Verification Of Forecasts Expressed In Terms Of Probability. *Mon Wea Rev*. 1950;78(1):1–3. doi:10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2
25. Chen T, Guestrin C Xgboost: a scalable tree boosting system. Paper presented at: 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining. 2016
26. Deng L, Yu D. Deep learning: methods and applications. *Found Trends Signal Process*. 2014;7(3–4):197–387. doi:10.1561/2000000039
27. Lundberg S, Lee S-I A unified approach to interpreting model predictions. arXiv preprint arXiv:170507874; 2017.
28. Lu S, Chen R, Wei W, Lu X Understanding heart-failure patients EHR clinical features via SHAP interpretation of tree-based machine learning model predictions. arXiv preprint arXiv:210311254; 2021.
29. Ramon Y, Martens D, Provost F, Evgeniou T. A comparison of instance-level counterfactual explanation algorithms for behavioral and textual data: SEDC, LIME-C and SHAP-C. *Adv Data Anal Classif*. 2020;14(4):801–819. doi:10.1007/s11634-020-00418-3
30. Chang CY, Wu KH, Wu PH, et al. In-hospital mortality associated with necrotizing soft tissue infection due to Vibrio vulnificus: a matched-pair cohort study. *World J Emerg Surg*. 2022;17(1):28. doi:10.1186/s13017-022-00433-z
31. Wu KH, Wu PH, Chang CY, et al. Differentiating necrotizing soft tissue infections from cellulitis by soft tissue infectious fluid analysis: a pilot study. *World J Emerg Surg*. 2022;17(1):1. doi:10.1186/s13017-022-00404-4
32. Wu PH, Wu KH, Hsiao CT, et al. Utility of modified Laboratory Risk Indicator for Necrotizing Fasciitis (MLRINEC) score in distinguishing necrotizing from non-necrotizing soft tissue infections. *World J Emerg Surg*. 2021;16(26). doi:10.1186/s13017-021-00373-0.
33. Lau KK, Hsiao CT, Fann WC, Chang CP. Utility of the Lactate/Albumin Ratio as a Predictor for Mortality in Necrotizing Fasciitis Patients. *Emerg Med Int*. 2021;2021:3530298. doi:10.1155/2021/3530298
34. Wu KH, Chang CP. Differentiating Lower Extremity Necrotizing Soft Tissue Infection from Severe Cellulitis by Laboratory Parameters and Relevant History Points. *Infect Drug Resist*. 2021;14:3563–3569. doi:10.2147/IDR.S327880
35. Zhang Z, Jin P, Feng M, et al. Causal inference with marginal structural modeling for longitudinal data in laparoscopic surgery: a technical note. *Laparoscopic Endoscopic Robot Surg*. 2022;5(2022):146–152. doi:10.1016/j.lers.2022.10.002