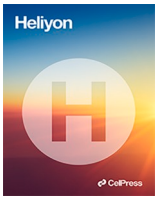
Heliyon 10 (2024) e29578

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Identifying necrotizing soft tissue infection using infectious fluid analysis and clinical parameters based on machine learning algorithms

Chia-Peng Chang [a,b], Chung-Jen Lin [a], Wen-Chih Fann [a], Chiao-Hsuan Hsieh [a,*]

[a] Department of Emergency Medicine, Chang Gung Memorial Hospital, No. 6, W. Sec., Jiapu Rd., Puzih City, Chiayi County, 613, Taiwan
[b] Department of Nursing, Chang Gung University of Science and Technology, Chiayi Campus, No.2, Sec. W., Jiapu Rd., Puzi City, Chiayi County, 613, Taiwan

ARTICLE INFO

Keywords:
Necrotizing soft tissue infection
Artificial intelligence
Machine learning
Infectious fluid

ABSTRACT

Background: Determining the presence of necrotizing soft tissue infection (NSTI) poses a significant hurdle. As of late, there has been a notable increase in the application of artificial intelligence (AI) machine learning techniques in identifying diseases, a shift that can be attributed to their exceptional efficiency, unbiased nature, and high precision.

Methods: Information was gathered from a cohort of 13 patients suffering from NSTI, alongside 12 patients with cellulitis. The construction of NSTI diagnostic machine learning models utilized four different algorithms, specifically random forest, k-nearest neighbors (KNN), support vector machine (SVM), and logistic regression. These models were constructed based on 28 distinctive attributes identified through statistical examination. Following this, the diagnostic efficiency of each algorithms was evaluated. A novel random forest model, streamlined for clinical use, was later developed by focusing on 6 attributes that had the most pronounced influence on the accuracy of our initial random forest model.

Results: The following data was noted regarding the sensitivity and specificity of the four NSTI diagnostic models:logistic regression displayed 78.2 % and 83.7 %, KNN presented 79.1 % and 87.1 %, SVM showed 83.5 % and 86.3 %, and random forest exhibited 89.6 % and 92.9 %, respectively. In comparison, lactate levels in fluid demonstrated 100 % sensitivity and 76.9 % specificity at an optimal cut-off point of 69.6 mg/dL. Among all four machine learning models, random forest outperformed the others and also showed better results than fluid lactate. A newly constructed random forest model, created using 6 of the 13 identified features, displayed promising results in diagnosing NSTI, having a sensitivity and specificity of 90.2 % and 92.2 %, respectively.

Conclusions: Developing a diagnostic model for NSTI employing the random forest algorithm has resulted in a diagnostic technique that is more efficient, cost-effective, and expedient. This

Abbreviations: AI, Artificial intelligence; AUC, Area under the ROC curve; BUN, Blood urea nitrogen; CRP, C-reactive protein; CT, Computed tomography; DBP, Diastolic blood pressure; ED, Emergency department; ICU, intensive care unit; INR, International normalized ratio; KNN, k-nearest neighbor; LDH, Lactate dehydrogenase; LRINEC, Laboratory Risk Indicator for Necrotizing Fasciitis; MLA, Machine learning algorithms; MRI, Magnetic resonance image; NLR, negative likelihood ratio; NPV, negative predictive value; NSTI, Necrotizing soft tissue infection; PLR, positive likelihood ratio; PPV, positive predictive value; RF, random forest; ROC, Receiver–operator characteristic; SBP, Systolic blood pressure; SD, Standard deviation; SVM, support vector machine; TP, total protein.

* Corresponding author.
E-mail address: g167111010@tmu.edu.tw (C.-H. Hsieh).

https://doi.org/10.1016/j.heliyon.2024.e29578
Received 15 December 2023; Received in revised form 8 April 2024; Accepted 10 April 2024
Available online 25 April 2024

approach could provide healthcare practitioners with the tools to identify and manage NSTI with greater efficacy.

1. Background

Necrotizing soft tissue infection (NSTI) refers to a severe soft tissue infection that results in rapid and extensive damage to muscle fascia and adjacent tissues. This condition is life-threatening, and its morbidity and mortality rates are exacerbated by late detection and treatment, including delays in administering comprehensive antibiotics and surgical intervention [1–4]. The incidence of NSTI is estimated to be between 0.4 and 1 case per 100,000 population annually, although this rate may vary by region and population. The mortality rate associated with NSTI ranges from 10 % to 36 %, with higher mortality observed when treatment is delayed [5]. The swift progression of NSTI underscores the importance of early diagnosis and treatment. Distinguishing NSTI from cellulitis based solely on clinical signs and symptoms is often challenging [5]. Various diagnostic tools have been developed to accurately identify NSTI, including laboratory tests, the Laboratory Risk Indicator for Necrotizing Fasciitis (LRINEC) scoring system, ultrasound imaging, computed tomography (CT) scans, magnetic resonance imaging (MRI), and fascia biopsies [4,6]. However, the lack of a single test that can quickly and precisely differentiate NSTI from other conditions continues to be a significant challenge in prompt diagnosis.

Recent advancements in artificial intelligence (AI) have led to a surge in its application within medical science. Machine learning, a subset of AI, enables computers to learn without being programmed for specific tasks. Utilizing machine learning algorithms such as support vector machines (SVM), k-nearest neighbors (KNN), and random forests (RF) allows for the creation of highly efficient, unbiased, and accurate models for disease detection. Chatterjee et al. successfully applied the random forest technique to develop a prognostic model for lymph node spread in head and neck squamous cell carcinoma [7], achieving an impressive diagnostic accuracy of 88 %. Another study by Cho et al. [8] employed decision trees and other machine learning methods to develop glaucoma diagnosis models, which exhibited sensitivities, specificities, and accuracies exceeding 95 %.

SVM is adept at handling high-dimensional data and modeling complex class boundaries, making it suitable for distinguishing NSTI from other infections. Random forests can identify intricate patterns without explicit feature engineering and highlight critical variables for NSTI diagnosis. Logistic regression is effective when there is a linear relationship between predictors and the outcome's log-odds, while KNN assumes that similar cases will have similar outcomes. Despite these potential applications, the use of machine learning algorithms in NSTI diagnosis is uncommon, and no comparative study has yet been conducted on the diagnostic capabilities of different models. In our research, we evaluated SVM, RF, logistic regression, and KNN to determine the most effective model for NSTI diagnosis, focusing on their diagnostic efficiency.

2. Methods

2.1. Study design

This study was approved by the Institutional Review Board of Chiayi Chang Gung.

Memorial Hospital (No.: 201900447B0C601). The study was also performed in accordance with the Declaration of Helsinki. We included adult patients who visited the Emergency Department (ED) between March 2019 and November 2020, adhering to three selection criteria: (1) the presence of serious soft tissue infection in a limb, with a suspicion of NSTI as per the supervising emergency doctors, (2) the identification of an infectious fluid accumulation along the deep fascia at the infection location via ultrasound, and (3) successful ultrasound-guided aspiration of the said fluid, with the results available for analysis. Those with a surgical history, chronic osteomyelitis, or persistent or recurrent soft tissue infections (including NSTI, cellulitis, or a soft tissue abscess) at the infected site were not included. Patients exhibiting skin abnormalities such as tumors or significant trauma, or those who had already started antibiotic therapy, were also left out. Infections of the soft tissue accompanied by abscess formation, for instance, cellulitis with pus and pyomyositis, were likewise excluded given the clear diagnosis from the appearance of pus-like infectious fluid. The usual course of action for patients exhibiting symptoms suggestive of NSTI upon arrival at the ED included the provision of broad-spectrum antibiotics and immediate consultation with an orthopedic surgeon to evaluate the need for surgical intervention. Ultrasound serves as a supplementary diagnostic tool that may aid in identifying NSTI; the disease is suspected if fluid build-up exceeds 2 mm in depth along the fascia [9,10]. The definitive clinical diagnosis was arrived at by the ED physicians and surgeons in the ED. Subsequently, urgent surgical debridement, including fasciotomy, was executed for patients diagnosed with NSTI.

2.2. Data collection

We examined electronic medical records and documented variables such as age, gender, pre-existing health conditions, Exposure history to seawater or contaminated agricultural water, past instances of animal bites, vital signs upon ED admission, laboratory findings from blood and infected soft tissue fluid, ultrasound images of the infected area, and surgical documents (such as records of fasciotomy or amputation). All blood tests were conducted within an hour of ED arrival, those infectious fluid sampled prior to the injection of antibiotics. Patients were categorized into either the NSTI group or the cellulitis group, in line with their discharge diagnoses. The discharge final diagnosis of NSTI was validated based on the surgical pathology report. Those patients lacking pathology reports to substantiate an NSTI diagnosis, or those who did not undergo surgical treatment, were assigned to the cellulitis group.

2.3. Statistical analysis and machine learning model evaluation

Data processing was conducted using SPSS version 22.0 software. Categorical variables were presented in terms of numbers and percentages, while continuous variables were represented as median values and their ranges. Variations in categorical variables between patients diagnosed with NSTI and those without were assessed using the Chi-square test. The Mann-Whitney *U* Test was used to evaluate differences in the continuous variables. A P-value less than 0.05 was considered statistically significant, and variables of statistical significance were incorporated into the diagnostic models.

In the context of this research, we utilized random forest, logistic regression, KNN, SVM to generate models aimed at diagnosing NSTI. During model creation, logistic regression, KNN, and SVM necessitated preliminary data adjustments. This step entailed normalizing the data to eliminate constraints related to the units of measure, all while preserving the original data distribution. For any missing values, we used the normalized mean values as replacements. On the other hand, the random forest algorithm did not require any data preprocessing. Therefore, we directly used the raw data in the segmentation process.

The parameters for each algorithm were configured in the following manner: (1) For logistic regression, the iteration limit was set to 60, the regularization coefficient was 1, and the minimum convergence error was defined as 0.00001. (2) In the case of KNN, we used the 5 closest neighbors. (3) For SVM, both positive and negative penalty factors were set to 1.0, with a convergence coefficient of 0.001. (4) Lastly, in the random forest model, the forest was composed of 60 trees, the minimal leaf node data size was 2, the minimum ratio of leaf node data to parent node data was 0, there was no restriction on the maximum tree depth, and each tree accepted 100,000 random data inputs.

We evaluated the diagnostic effectiveness of the four algorithmic models, namely their sensitivity, specificity, and precision, using a confusion matrix that contained both predicted and actual classification data. Given that the dataset was divided randomly into a training set comprising 80 % of the data and a testing set containing the remaining 20 %, we carried out ten rounds of model construction tests for each algorithm. The aim was to establish the average sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (PLR), negative likelihood ratio (NLR), and precision for each model—these measurements represented the final outcomes. Among the four models, the random forest model exhibited the most robust diagnostic performance. This relative influence of each feature on the random forest model's precision was also ranked. For the random forest model, features leading to higher average reductions in the Gini index were considered more impactful on the model's classification accuracy.

Table 1
Comparison of clinical parameters and infectious fluid analysis between NSTI group and cellulitis group.

Variable	NSTI			Cellulitis			p-value
	N	Mean	SD	N	Mean	SD	
Age (years)	13	72	12.08	12	65.25	13.8	0.21
Body temperature at triage (°C)	13	36.87	1.35	12	36.76	0.76	0.81
Heart rate at triage	13	91.85	17.17	12	97.25	26.31	0.55
Respiration rate at triage	13	19.31	1.7	12	18.75	0.87	0.32
SBP at triage (mmHg)	13	105.08	29.64	12	146.58	22.6	<0.001
DBP at triage (mmHg)	13	64.46	17.75	12	87.67	11.69	<0.001
In Bloodrowhead							
WBC (10^3/μL)	13	15.98	6.55	12	9.91	3.16	0.01
Segment (%)	13	75.53	14.98	12	72.03	11.98	0.53
Hemoglobin (g/dL)	13	12.75	2.14	12	12.93	1.95	0.84
Platelet (10^3/μL)	13	202.15	93.17	12	222.67	84.46	0.57
INR	13	1.14	0.15	8	1.03	0.07	0.08
Sodium (mEq/L)	13	134.77	3.24	12	135.33	2.74	0.64
Glucose (serum) (mg/dL)	13	181.38	91.14	11	175.73	64.93	0.87
C-reactive protein (mg/L)	11	200.36	125.79	11	71.76	63.2	0.01
Creatinine (mg/dL)	13	1.65	0.71	12	1.13	0.27	0.02
BUN (mg/dL)	11	28.66	9.95	9	14.67	6.55	<0.001
Potassium (mEq/L)	13	3.86	0.42	12	3.72	0.45	0.41
Alanine transaminase (U/L)	13	37.38	22.12	12	37.5	27.38	0.99
Albumin (serum) (g/dL)	13	3.75	0.49	12	3.91	0.56	0.47
TP (serum) (g/dL)	13	6.43	0.74	9	6.77	0.59	0.27
LDH (serum) (U/L)	13	194.17	59.64	12	187.29	59.94	0.78
Lactate (serum) (mg/dL)	13	31.56	24.5	11	14.46	7.72	0.04
Infectious fluid analysisrowhead							
Albumin (fluid) (g/dL)	13	2.19	0.96	12	1.33	0.72	0.02
LDH (fluid) (U/L)	13	5023.66	6083.34	12	315.32	307.07	0.01
Glucose (fluid) (mg/dL)	13	82.23	68.43	11	146	48.66	0.02
TP (fluid) (g/dL)	12	3.68	1.48	12	1.87	1.07	<0.001
Lactate (fluid) (mg/dL)	13	121.98	75.63	11	24.65	13.72	<0.001
pH (fluid)	11	8.05	0.45	8	8.6	0.32	0.01

DBP: diastolic blood pressure, INR: international normalized ratio, LDH: lactate dehydrogenase, TP: total protein, SBP: systolic blood pressure, SD: standard deviation, WBC: White blood cells.

3. Results

3.1. Patient characteristics

The inclusion criteria was satisfied by 31 patients in total. However, six of these individuals were disqualified because of the discovery of infectious fluid akin to pus following aspiration. The ultimate diagnosis for these excluded six individuals was pyomyositis for one patient and soft tissue infections with abscess formation for the remaining five. Consequently, this study comprised of 25 participants, who were then divided into two groups: the NSTI group, consisting of 13 patients, and the cellulitis group with 12 patients. The clinical traits and laboratory results upon their initial visit to the Emergency Department were then assessed and contrasted among the patients belonging to cellulitis or NSTI group.

3.2. Features derived from clinical parameters and infectious fluid analysis

Among those 28 features, 13 features were significantly different between NSTI and cellulitis patients (Table 1). Those significant 13 features were throwed into the model: C-reactive protein (CRP), systolic blood pressure (SBP) at triage; diastolic blood pressure (DBP) at triage; blood urea nitrogen (BUN), white blood cell count (WBC), Creatinine and lactate in blood; albumin, lactate dehydrogenase (LDH), glucose, total protein, lactate, and pH value in infectious fluid.

3.3. Diagnostic performances among the four algorithms

The diagnostic capabilities of the four machine learning algorithms for NSTI are illustrated in Table 2 and Figs. 1 and 2. Among these models, the random forest model stands out as the most effective one for diagnosing NSTI. It boasts higher levels of sensitivity, specificity, PPV, NPV, PLR, and accuracy when compared to logistic regression, KNN, and SVM models. Meanwhile, its NLR is observed to be lower than that of the other three models.

3.4. Evaluation performance of new derived RF model with best 6 features derived from gini index

In order to streamline the new random forest model for clinical use, we chose six features, which were determined based on the Gini index (as depicted in Fig. 3). These selected features formed the basis for the creation of the new random forest model. The effectiveness of this newly established random forest model in diagnosing NSTI is demonstrated in Table 3.

4. Discussion

Significant 13 features were including CRP, SBP at triage; DBP at triage; BUN, WBC in blood, Creatinine and lactate in blood; albumin, LDH, glucose, total protein, lactate, and pH value in infectious fluid. Septic shock is a common complication of NSTI. To protect the heart, brain, and other important organs, renal perfusion pressure decompensation may occur in the early stage. In addition, blood transfusion and infusion therapy can further aggravate the renal ischemia-reperfusion injury, resulting in short-term increased creatinine levels. Moreover, renal insufficiency also limits the use of most antibiotics and can lead to inadequate anti-infective treatment, which can increase mortality. Low blood pressure, elevated BUN and creatinine values were also significant contributing features in our predictive model. CRP, which is produced mainly in the liver, has long been used as a marker of infection. High CRP levels may indicate severe infection or inflammation caused by soft tissue infections. In predicting bacterial infection in patients with decompensated cirrhosis, the CRP level was superior to the procalcitonin level and white blood cell count [11]. In patients visiting the ED, there was an association between bacterial infection and the CRP level [12]. However, a single measurement of CRP is nonspecific and a variation in the CRP level is a more specific predictive factor for mortality. Accordingly, if the first CRP measurement is high, additional tests after hospitalization are required [13]. The clinical significance of serum lactate was established over the last decade. Trends in lactate concentration over time reflect the clinical response of patients to resuscitation and surgical intervention. Elevated lactate was shown to be independently associated with mortality rate in critical ill patients. Even intermediate initial serum lactate was an indicator of mortality of organ dysfunction and shock in severe sepsis patients in the emergency department. In hospitalized patients, increased lactate indicated high mortality, mechanical ventilation, vasopressor requirement [14,

Table 2
Performance of the four MLAs and lactate in fluid for diagnosing NSTI.

	AUC	Sensitivity	Specificity	PPV	NPV	PLR	NLR	Accuracy
Logistic regression	0.834	78.2 %	83.7 %	80.1 %	84.9 %	4.80	0.26	82.7 %
KNN	0.885	79.1 %	87.1 %	82.5 %	83.8 %	6.13	0.23	83.9 %
SVM	0.923	83.5 %	86.3 %	82.6 %	86.9 %	6.09	0.19	81.5 %
RF	0.958	89.6 %	92.9 %	90.8 %	91.4 %	12.6	0.11	90.5 %
Lactate in fluid	0.937	100 %	76.9 %	80.6 %	100 %	4.33	0	84.3 %

AUC: Area under the ROC curve; KNN: k-nearest neighbor; NLR: negative likelihood ratio; NPV: negative predictive value; PLR: positive likelihood ratio; PPV: positive predictive value; RF: random forest; SVM: support vector machine.

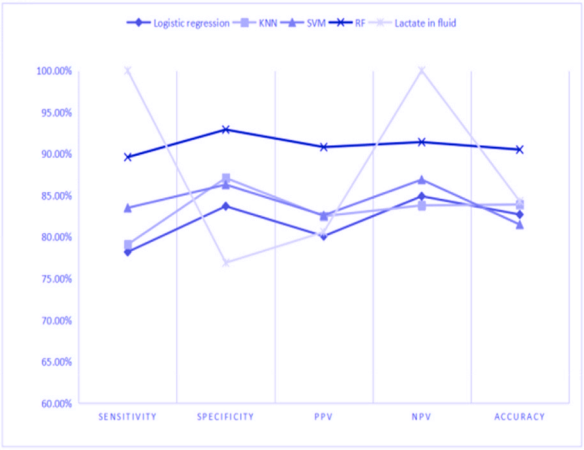


Fig. 1. Sensitivity, specificity, accuracy, PPV, NPV curves for lactate in fluid and the four algorithmic models.

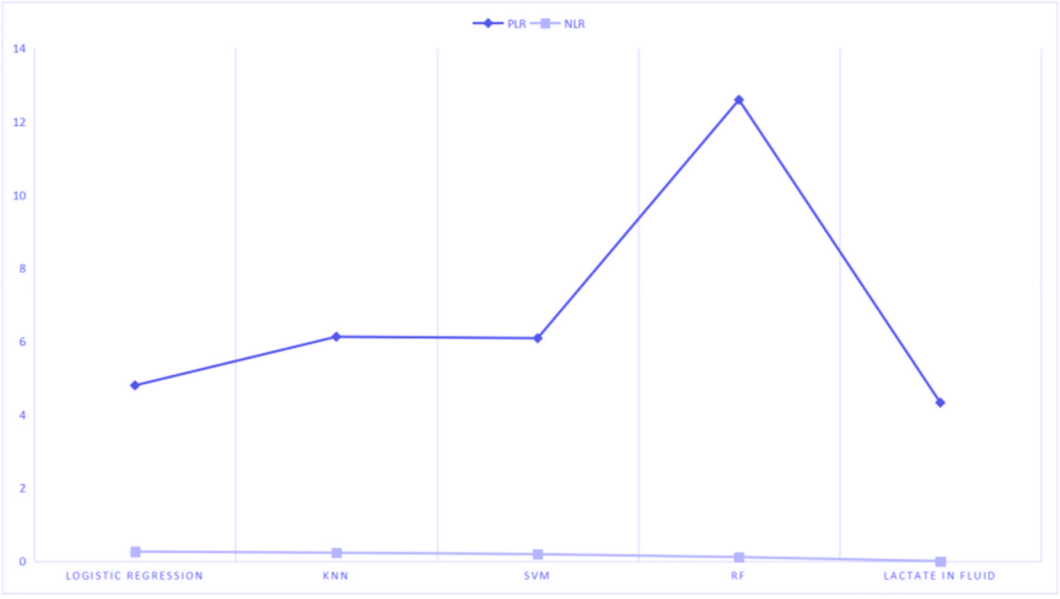


Fig. 2. The PLR, NLR curves for lactate in fluid and the four algorithmic models.

15].

Synovial fluid lactate is an important biomarker that can be used in the diagnostic and prognostic evaluation of various joint disorders. It is usually present at lower levels in healthy joints, but its concentration tends to increase in response to various pathological conditions. The rationale behind this is tied to the body's metabolic responses to injury or disease. After orthopaedic surgery or trauma, an increase in synovial lactate concentration may indicate the presence of complications such as infection or tissue necrosis. LDH levels in fluid can be influenced by various factors, including inflammation, tissue damage, and certain diseases. An elevation in synovial LDH levels may suggest increased cellular turnover and metabolic activity within the joint, which could be associated with conditions such as rheumatoid arthritis, osteoarthritis, or joint infections [16,17].Total protein in fluid can help differentiate between inflammatory and non-inflammatory joint conditions. Inflammatory conditions typically exhibit higher total protein levels compared to non-inflammatory conditions, such as osteoarthritis. Serial measurements of synovial fluid total protein can be useful in monitoring the response to treatment in various joint disorders. Decreases in protein levels over time may indicate a positive response to therapy, while persistently high or increasing levels may suggest ongoing inflammation or treatment resistance. Synovial albumin levels can reflect the integrity of the synovial membrane, which lines the joint cavity. Damage or disruption of the synovial membrane can result in increased leakage of albumin into the synovial fluid. It may be correlated with disease severity and prognosis in certain joint conditions [16–21]. Despite the diagnostic value of soft tissue fluid, seldom researches were mentioned those parameters in NSTI. In our knowledge, we are the first research to establish the importance of infectious soft tissue fluid analysis in diagnosing necrotizing soft

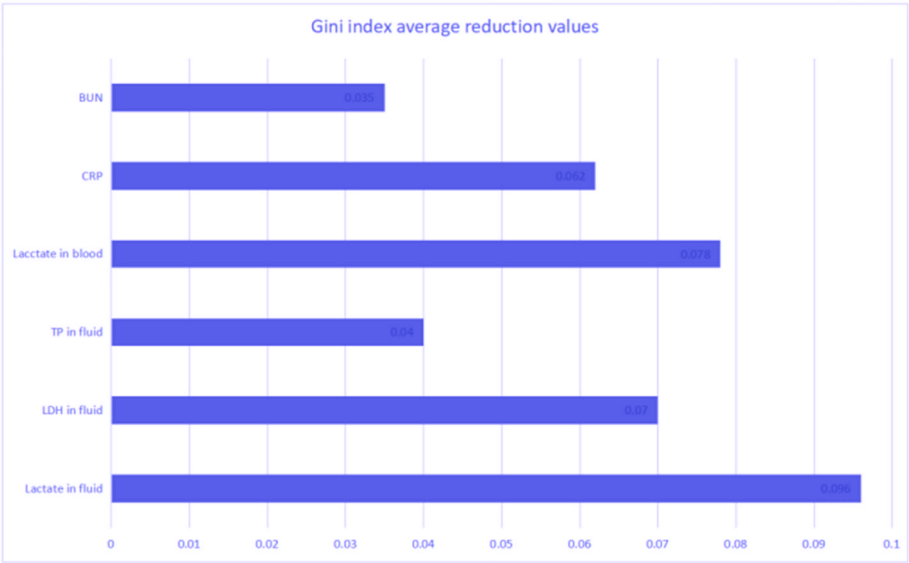


Fig. 3. The impacts of the first 6 features on the accuracy of the RF model.
BUN: Blood urea nitrogen; CRP: C-reactive protein; LDH: Lactate dehydrogenase; TP: total protein.

Table 3
Performance of new RF model for diagnosing NSTI.

	AUC	Sensitivity	Specificity	PPV	NPV	PLR	NLR	Accuracy
RF	0.962	90.4 %	92.2 %	85.8 %	93.1 %	11.6	0.1	91.8 %

AUC: Area under the ROC curve; NLR: negative likelihood ratio; NPV: negative predictive value; PLR: positive likelihood ratio; PPV: positive predictive value; RF: random forest.

tissue infection.

AI has been the harbinger of transformations across various facets of society in the past few years, and investigations into its probable applications in the realm of medicine are also on the rise. When it comes to diagnosing diseases, AI's machine learning algorithms are capable of handling copious amounts of data. Moreover, they can optimally utilize established information to devise incredibly predictive models for disease diagnosis. As of now, medical diagnostics employ various machine learning algorithms such as random forest, logistic regression, KNN, SVM, among others. Logistic regression is a traditional statistical technique. The outcome produced by the logistic regression model is a probability ranging from 0 to 1, where a score of 0.5 signifies a binary classification. It is commonly employed in the field of medical diagnostics. The KNN method operates through a majority voting mechanism. When an example is introduced for testing, a voting procedure takes place based on the classifications of the k-closest training samples. As a result, the test sample is allocated to the category that accumulates the maximum number of votes. SVM is an alternative machine learning algorithm that can be employed for classification. Essentially, SVM operates by building a hyperplane that optimizes the distance between two different sample types and this hyperplane. Random forest, an alternate classification approach, encompasses numerous decision trees which amalgamate input data using If-Then rules to form a tree classifier. Based on the model thus created, fresh input data is routed to the leaf nodes starting from the root node, and the ultimate level output symbolizes the final categorization outcome. Nonetheless, over-fitting could potentially happen and lead to erroneous outcomes if the decision tree is overly complex. To avert over-fitting, random forest employs the principle of integrated learning, which consolidates the classification outputs of each decision tree, thereby producing results that are more precise and steady.

Xiao and his team employed random forest technique for building a diagnostic model for prostate cancer that achieved an accuracy rate of 83.1 %, a sensitivity of 65.6 %, and a specificity of 93.8 % [22]. On a similar note, Casanova and his associates evaluated the diagnostic efficacy of both logistic regression and RF in diagnosing diabetic retinopathy [23]. The findings of their study underscored that RF provides superior classification accuracy.

Currently, there is a dearth of studies focused on the application of AI machine learning algorithms in diagnosing NSTI. In our research, we created diagnostic models for NSTI utilizing four machine learning algorithms (KNN, SVM, logistic regression, and random forest). We then compared their diagnostic capabilities of these four artificial intelligence models to identify the most effective one for distinguishing NSTI. We incorporated 13 features that exhibited statistical variances between the NSTI group and the cellulitis group into the model. These included blood pressures, blood metrics, and infectious fluid assessments. Given that all these measures are extensively utilized in clinical settings, they can be quickly and easily procured without requiring any specialized apparatus. As a result, the models illustrated in this paper hold broad applicability.

The findings indicate that among the four algorithms, random forest displays the most superior diagnostic performance with a sensitivity, specificity, accuracy, and AUC of 89.6, 92.9, 90.5 %, and 0.958 respectively. The diagnostic performances of SVM, KNN, and logistic regression were found to be alike. Prior research has shown that RF is remarkably efficient in classifying various diseases. For instance, Huang et al. utilized four machine learning models (SVM, naive Bayes, KNN, and random forest) to create decision-support systems for diagnosing liver fibrosis [24]. The findings indicated that random forest demonstrated superior precision compared to the other three machine learning algorithms. In a further comparison carried out by Chicco et al., they assessed the effectiveness of various classifiers, including probabilistic neural networks, neural networks based on perceptrons, random forest, One Rule (OneR), and decision tree classifiers, particularly in the forecasted diagnosis of pleural mesothelioma [25]. Their findings revealed that random forest superseded all other MLA models. Thus, random forest model evidently holds an advantage when used for diagnosing diseases.

To enhance its clinical usability, we picked the six most influential features on initial random forest model's accuracy to establish a new random forest model. This finding demonstrates that the diagnostic performance of this freshly developed model is akin to the original random forest model that was built using 13 features. Minimizing the number of features in the model holds substantial importance as it could potentially lower medical costs and offer more convenience in clinical settings. There are certain limitations to this research, including the fact that the data was gathered from a single-center cohort and the number of participants was somewhat limited. For future research, implementing a multi-center prospective study with a more extensive sample could prove advantageous in constructing a more accurate diagnostic model for NSTI.

5. Conclusion

Implementing AI machine learning algorithms and ultrasound assisted aspiration of infectious fluid to construct a diagnostic model for NSTI could potentially enhance diagnostic capabilities. In this context, random forest proves to be more effective than logistic regression, KNN, and SVM. Developing a diagnostic model for NSTI using random forest algorithms could lead to a more efficient, cost-effective, and rapid diagnostic procedure based on common clinical data. This would aid clinicians in making superior diagnostic and therapeutic choices.

Funding

The Chang Gung Medical Research Program provided the funding for this study. (CMRPG6J0111 and CMRPG6J0112).

Ethics approval and consent to participate

The Institutional Review Board of Chang Gung Memorial Hospital approved this retrospective study (No. 201900447B0C601 and 201900692B0). Informed consent is waived by Institutional Review Board of Chang Gung Memorial Hospital that approved the need for informed consent.

Data availability statement

Data will be made available on request.

CRediT authorship contribution statement

Chia-Peng Chang: Writing – original draft, Data curation, Conceptualization. **Chung-Jen Lin:** Investigation, Data curation. **Wen-Chih Fann:** Validation, Funding acquisition, Formal analysis. **Chiao-Hsuan Hsieh:** Writing – review & editing, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] D.L. Stevens, A.E. Bryant, Necrotizing soft-tissue infections, N. Engl. J. Med. 377 (23) (2017) 2253–2265.
[2] G.M. Howell, M.R. Rosengart, Necrotizing soft tissue infections, Surg. Infect. 12 (3) (2011) 185–190.
[3] P.H. Wu, K.H. Wu, C.T. Hsiao, S.R. Wu, C.P. Chang, Utility of modified laboratory risk indicator for necrotizing fasciitis (MLRINEC) score in distinguishing necrotizing from non-necrotizing soft tissue infections, World J. Emerg. Surg. 16 (1) (2021) 26.
[4] M. Sartelli, X. Guirao, T.C. Hardcastle, Y. Kluger, M.A. Boermeester, K. Raşa, et al., WSES/SIS-E consensus conference: recommendations for the management of skin and soft-tissue infections, World J. Emerg. Surg. 13 (1) (2018) 58, 2018.
[5] K.-H. Wu, C.-P. Chang, Differentiating lower extremity necrotizing soft tissue infection from severe cellulitis by laboratory parameters and relevant history points, Infect. Drug Resist. 14 (2021) 3563–3569.
[6] C.H. Wong, L.W. Khin, K.S. Heng, K.C. Tan, C.O. Low, The LRINEC (laboratory risk indicator for necrotizing fasciitis) score: a tool for distinguishing necrotizing fasciitis from other soft tissue infections, Crit. Care Med. 32 (7) (2004) 1535–1541.

[7] R. Forghani, A. Chatterjee, C. Reinhold, et al., Head and neck squamous cell carcinoma: prediction of cervical lymph node metastasis by dual-energy CT texture analysis with machine learning, Eur. Radiol. 29 (11) (2019) 6172–6181.
[8] S.J. Kim, K.J. Cho, S. Oh, Development of machine learning models for diagnosis of glaucoma, PLoS One 12 (5) (2017) e0177726.
[9] C.N. Lin, C.T. Hsiao, C.P. Chang, T.Y. Huang, K.Y. Hsiao, Y.C. Chen, et al., The relationship between fluid accumulation in ultrasonography and the diagnosis and prognosis of patients with necrotizing fasciitis, Ultrasound Med. Biol. 45 (7) (2019) 1545–1550.
[10] K.H. Wu, P.H. Wu, C.Y. Chang, et al., Differentiating necrotizing soft tissue infections from cellulitis by soft tissue infectious fluid analysis: a pilot study, World J. Emerg. Surg. 17 (1) (2022) 1, https://doi.org/10.1186/s13017-022-00404-4. Published 2022 Jan 8.
[11] S. Khedher, N. Fouthaili, A. Maoui, et al., The diagnostic and prognostic values of C-reactive protein and procalcitonin during bacterial infections in decompensated cirrhosis, Gastroenterol. Res. Pract. 2018 (2018) 5915947.
[12] R.O. Ximenes, A.Q. Farias, A.S. Neto, et al., Patients with cirrhosis in the ED: early predictors of infection and mortality, Am. J. Emerg. Med. 34 (2016) 25–29.
[13] J.-P. Cervoni, A. Amorós, R. Bañares, et al., Prognostic value of C-reactive protein in cirrhosis: external validation from the CANONIC cohort, Eur. J. Gastroenterol. Hepatol. 28 (2016) 1028–1034.
[14] H. Khosravani, R. Shahpori, H.T. Stelfox, et al., Occurrence and adverse effect on outcome of hyperlactatemia in the critically ill, Crit. Care 13 (2009) R90.
[15] M.E. Mikkelsen, A.N. Miltiades, D.F. Gaieski, et al., Serum lactate is associated with mortality in severe sepsis independent of organ failure and shock, Crit. Care Med. 37 (2009) 1670–1677.
[16] G. Balato, T. Ascione, D. Rosa, P. Pagliano, G. Solarino, B. Moretti, M. Mariconda, Synovial fluid lactate levels in common arthropathies: a promising tool for differential diagnosis, J. Int. Med. Res. 45 (3) (2017) 1106–1112.
[17] M.E. Margaretten, J. Kohlwes, D. Moore, S. Bent, Does this adult patient have septic arthritis? JAMA 297 (13) (2007) 1478–1488.
[18] M. Lenski, M.A. Scherer, Diagnostic potential of inflammatory markers in septic arthritis and periprosthetic joint infections: a clinical study with 719 patients, Inf. Disp. 47 (6) (2015) 399–409.
[19] E. Shu, L. Farshidpour, M. Young, M. Darracq, C. Ives Tallman, Utility of point-of-care synovial lactate to identify septic arthritis in the emergency department, Am. J. Emerg. Med. 37 (3) (2019) 502–505.
[20] J.M. Porcel, R.W. Light, Diagnostic approach to pleural effusion in adults, Am. Fam. Physician 73 (7) (2006) 1211–1220.
[21] M.E. Wilcox, C.A.K.Y. Chong, M.B. Stanbrook, A.C. Tricco, C. Wong, S.E. Straus, Does this patient have an exudative pleural effusion?: the rational clinical examination systematic review, JAMA 311 (23) (2014) 2422–2431.
[22] L.H. Xiao, P.R. Chen, Z.P. Gou, et al., Prostate cancer prediction using the random forest algorithm that takes into account transrectal ultrasound findings, age, and serum levels of prostate-specific antigen, Asian J. Androl. 19 (5) (2017) 586–590.
[23] R. Casanova, S. Saldana, E.Y. Chew, et al., Application of random forests methods to diabetic retinopathy classification analyses, PLoS One 9 (6) (2014) e98587.
[24] Y. Chen, Y. Luo, W. Huang, et al., Machine-learning-based classification of realtime tissue elastography for hepatic fibrosis in patients with chronic hepatitis B, Comput. Biol. Med. 89 (2017) 18–23.
[25] D. Chicco, C. Rovelli, Computational prediction of diagnosis and feature selection on mesothelioma patient health records, PLoS One 14 (1) (2019) e0208737.