ATR-FTIR Spectral Fingerprinting of Serum with Machine Learning Enables Single-Measurement Diagnosis and Prognostic Risk Assessment for Sepsis

Xuejie Wang[1,2], Guoqiang Bao[2,*]
[1]Department of General Surgery, Shaanxi University of Chinese Medicine, Xianyang, China
[2]Department of General Surgery, Tangdu Hospital, The Air Force Military Medical University, Xi'an, China

Abstract: Sepsis necessitates rapid, accurate diagnostic tools, as current biomarkers often lack sufficient speed or specificity. This study explores the novel integration of Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) spectroscopy with machine learning as a promising strategy to address this gap. We analyzed serum samples from 146 subjects (sepsis patients and controls) using ATR-FTIR and compared 16 ML algorithms. The Extra Trees classifier demonstrated superior performance on an independent test set, achieving an area under the curve (AUC) of 0.899, an F1-score of 0.885, and a notably high recall of 0.958, indicating strong potential for minimizing missed diagnoses. For prognostic stratification, a simple clinical score based on NT-PROBNP, Troponin I, and platelet count effectively identified high-risk patients (28-day mortality of 24% vs. 0% in the low-risk group). Furthermore, a prognostic model based solely on spectral features achieved an AUC of 0.716 using a Naive Bayes classifier. Importantly, significant correlations were established between specific spectral features and established clinical biomarkers of infection (PCT, CRP), organ dysfunction (Scr), and age, providing a plausible biological basis for the spectral findings. This proof-of-concept study demonstrates that serum infrared spectral fingerprinting, enhanced by machine learning, can serve as a rapid, adjunctive tool for both diagnosing sepsis and stratifying patient risk. This approach directly responds to the critical care imperative for technologies that are faster, more informative, and capable of improving early clinical decision-making. Further validation in larger, multicenter cohorts is warranted to translate this potential into clinical practice.

Keywords: Sepsis; ATR-FTIR Serum Fingerprinting; Machine Learning; Diagnosis; Prognostic

1. Introduction

Sepsis and septic shock represent a persistent and formidable frontier in critical care medicine, constituting a leading cause of global mortality [1–4]. Despite decades of research, the initial clinical presentation of sepsis is often nonspecific, overlapping with other inflammatory conditions [5,6]. This diagnostic ambiguity leads to critical delays in administering time-sensitive, goal-directed therapy, a primary determinant of patient survival [7,8]. Concurrently, the profound heterogeneity in host response and the dynamic progression of the syndrome render accurate early prognostication exceptionally challenging [8,9]. The current reliance on conventional biomarkers—such as procalcitonin (PCT) and C-reactive protein (CRP)—is hampered by variable sensitivity and specificity, leaving clinicians without a rapid, reliable tool to simultaneously confirm diagnosis and stratify impending risk [10–15]. Consequently, there is an urgent, unmet need for novel diagnostic methodologies that can provide a more comprehensive, real-time molecular snapshot of the host state.

Emerging technologies that leverage multi-analyte profiling and computational analysis offer a promising path forward. Vibrational spectroscopy, particularly Fourier-transform infrared (FTIR) spectroscopy, has garnered interest for its ability to generate a holistic "biochemical fingerprint" from biofluids, capturing the composite signal of molecular fragments (i.e., functional groups) from proteins, lipids, nucleic acids, and carbohydrates in a single, label-free measurement [16–22]. This integrative readout potentially mirrors the complex pathophysiology of sepsis more

faithfully than single-molecule biomarkers [23–25]. However, translating the rich spectral data into clinically actionable insights necessitates advanced computational interpretation [26–28].

Here, we hypothesize that the serum infrared spectral fingerprint, when decoded by machine learning (ML) algorithms, can serve as a rapid and powerful tool to address the dual challenges of sepsis diagnosis and early prognosis. We systematically develop and validate an ML framework based on attenuated total reflection FTIR (ATR-FTIR) spectroscopy of serum. We further explore the biological plausibility of our model by correlating specific spectral signatures with established clinical indicators of infection and organ dysfunction. This work aims to demonstrate a pragmatic, translatable approach that moves beyond conventional biomarker paradigms towards integrated, data-driven decision support in the critical care of sepsis patients.

2. Materials and Methods

2.1 Serum Sample Collection and Processing

Between March 2024 and December 2024, patients who were hospitalized in the Intensive Care Unit (ICU) and general surgery department of Tangdu Hospital were consecutively included in the sepsis group (n = 84), while healthy individuals who underwent routine physical examinations at the same hospital were recruited as the control group (n = 68). For spectral diagnosis, 6 sepsis patients were excluded due to poor spectral quality, resulting in a cohort of 146 subjects (78 sepsis and 68 controls). For prognostic analysis, one sepsis patient was excluded due to missing clinical records, leaving a cohort of 83 patients. Exclusion criteria included patients with neurological diseases, prior antibiotic treatment within two weeks, post-cardiopulmonary resuscitation, and chronic gastrointestinal disorders. The diagnostic criteria for sepsis followed the Third International Consensus Definitions outlined in the Sepsis-3 guidelines by the American Society of Critical Care Medicine, published in 2016. Clinical information and serum validation markers (e.g., PCT and CRP) were collected on the day of serum sampling. Collected serum samples were initially lyophilized for 12 hours at -80°C before Fourier infrared spectroscopy analysis. Informed consent was obtained from all patients to participate in this study. This study was approved by the Medical Ethics Committee of Tangdu Hospital, Fourth Military Medical University (approval number: 202209-01). All procedures were performed in compliance with relevant laws and institutional guidelines. Informed consent was obtained from all individual participants included in the study and/or their legal guardian(s).

2.2 Spectral Acquisition

ATR-FTIR spectral analysis was performed using a Nicolet iS50 Fourier transform infrared spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) equipped with a Built-in diamond ATR accessory. The attenuated total reflection (ATR) module of the spectrometer was employed to analyze the serum samples. Spectral acquisition was conducted with the following parameters: 32 scans were accumulated over the wavenumber range of 4000 to 400 cm^{-1} with a spectral resolution of 4 cm^{-1}. To ensure data robustness, multiple technical replicates (typically 1-5 scans per clinical sample) were acquired. For samples with multiple replicates, the average spectrum was calculated and used for subsequent machine learning analysis to minimize instrumental noise and local sample heterogeneity. Samples had obvious baseline anomalies that could not be corrected were excluded from the spectral dataset. And the generated infrared spectral data were collected and saved as absorbance values. Prior to measurement, the sample stage was cleaned by unidirectionally wiping it with lens paper moistened with 95% ethanol and then air-dried. A sample holder with a fixed-volume sampling well was secured on the stage. The spectrometer and its operating software OMNIC 9.2.86 (Thermo Fisher Scientific, Waltham, MA, USA) were activated, and a background spectrum was collected before placing the freeze-dried serum sample onto the diamond surface.

2.3 Data Preprocessing

All serum spectra preprocessing was performed collaboratively using OMNIC 9.2 and Origin 2021 software. First, rubber band baseline correction was applied to the raw spectra to eliminate interference from environmental light scattering. All spectral data were normalized using the Min-Max method based on the amide I band to correct intensity variations caused by uneven sample deposition. Smoothing was

conducted using the Savitzky–Golay algorithm (2nd order polynomial, 9-point window), followed by calculation of the first derivative. Finally, the integrated areas of six key characteristic regions, as defined in Figure 1e, were computed using Origin 2021 software to construct feature vectors for subsequent input into machine learning algorithms.

2.4 Machine Learning Modeling

A machine learning analysis framework was constructed based on the serum spectral characteristics of the subjects to evaluate its clinical value in sepsis diagnosis and prognosis prediction. First, according to the random splitting strategy shown in Figure 2a, the spectral dataset (n = 146, comprising 78 sepsis and 68 controls) were divided into a training set (n=102) and an independent test set (n=44) at a ratio of 7:3. During the model development phase, using the extracted integrated areas of six spectral features as input variables, sixteen machine learning algorithms, including Extra Trees, Random Forest, XGBoost, Neural Network, and SVM, were constructed and compared in parallel (Table 1).

Clinical data are presented as mean ± standard deviation (mean ± SD). Differences between groups were assessed using the independent samples t-test or the Mann-Whitney U test, based on data distribution characteristics, with statistical significance set at $P < 0.05$. The classification performance of the models was comprehensively evaluated using Accuracy, Precision, Recall, F1-score, and AUC. All statistical analyses and graphical work were completed in the Origin 2021 and Python 3.10.0 environments.

Table 1. Summary of the Machine Learning Algorithms Used in this Study

Category	Full Name	Abbreviation
Linear Models	Logistic Regression	LR
	Linear Discriminant Analysis	LDA
Non-linear Models	Naive Bayes	NB
	Quadratic Discriminant Analysis	QDA
	K-Nearest Neighbors (K=3, 5)	KNN-3, KNN-5
	Support Vector Machine (Linear Kernel)	SVM-Lin
	Support Vector Machine (RBF Kernel)	SVM-RBF
	Neural Network (Multi-layer Perceptron)	MLP
Tree-based Models	Decision Tree	DT
	Random Forest	RF
	Extra Trees	ET
Ensemble Learning	AdaBoost	AB
	Gradient Boosting Decision Tree	GBDT / GB
	XGBoost	XGB
	LightGBM	LGB

2.5 Clinical Prognostic Scoring System Construction

To construct the prognostic assessment model, clinical characteristics were screened and a scoring system was developed based on the cohort of 83 sepsis patients for whom complete 28-day survival outcomes were available. First, based on patients' 28-day survival outcomes (survival/death), univariate analysis of various clinical indicators was performed using ROC curves, and indicators with AUC > 0.7 were selected as candidate prognostic factors. Subsequently, the Youden Index of each candidate factor was calculated to determine the optimal cut-off value. Based on these cut-off values, an additive scoring system was constructed: patients were assigned 1 point if an indicator exceeded (or fell below) the cut-off value, and 0 points otherwise. Finally, based on the total score, patients were classified into a Low-Risk Group and a High-Risk Group, and differences in clinical outcomes between the two groups were statistically analyzed.

2.6 STARD Reporting

We used the STARD reporting guideline to draft this manuscript, and the STARD reporting checklist when editing, included in supplement A.

2.7 Statistics and Reproducibility

No statistical method was used to predetermine sample size. The diagnostic cohort (n=146) and

prognostic cohort (n=83) were determined based on patient availability. As pre-specified, 6 samples were excluded due to spectral anomalies (baseline shifts) and 1 due to missing clinical metadata. To ensure technical reproducibility, each sample was measured in 1–5 technical replicates; the average absorbance spectrum was used for all subsequent analyses. The internal consistency of the results was validated using an independent, "held-out" test set (n=44) that was not involved in model training or hyperparameter tuning. All statistical tests were two-sided. Normality was assessed using the Shapiro-Wilk test. Group differences were evaluated using the independent samples t-test or Mann-Whitney U test as appropriate. Categorical variables were compared using the Chi-square test. Correlations were calculated using Pearson's r or Spearman's ρ. Machine learning models and performance metrics (AUC, F1-score, Recall) were implemented in Python 3.10.0 using the Scikit-learn library. Preprocessing and spectral integration were performed using OMNIC 9.2 and Origin 2021. Statistical significance was defined as $P < 0.05$.

3. Results

3.1 Analysis of Serum Spectral Characteristics in Sepsis Patients

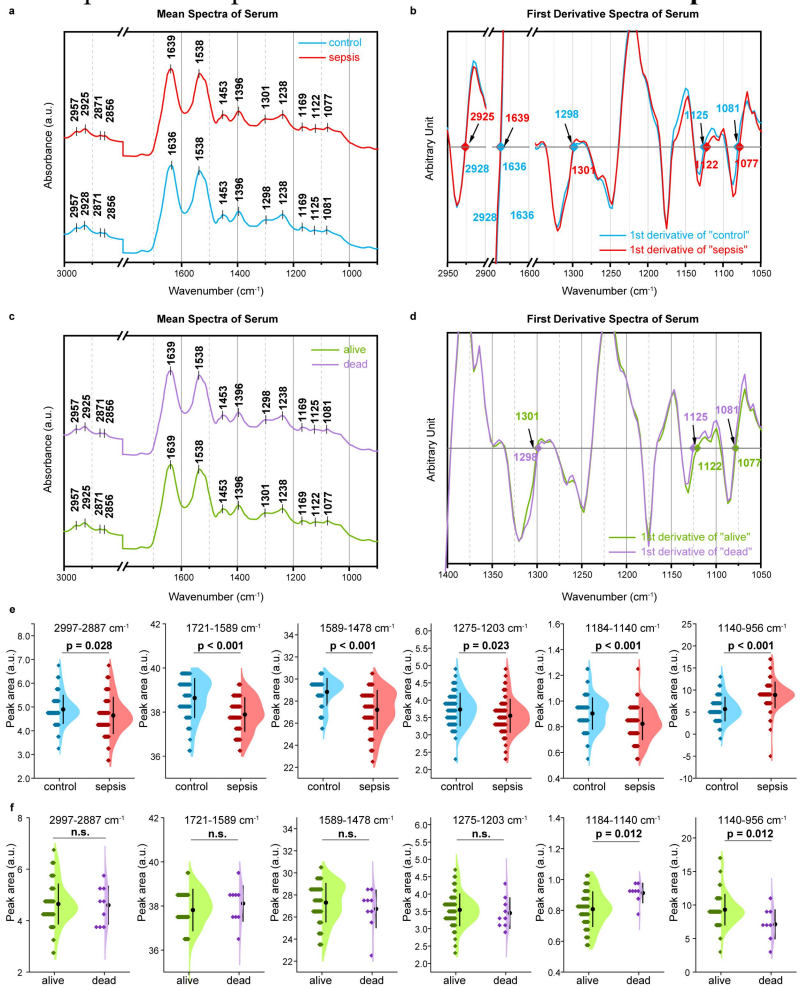

Figure 1. Characterization of Serum ATR-FTIR Spectral Fingerprints and Identification of Differential Spectral Features. (a) Comparison of mean raw infrared absorbance spectra between the sepsis group (red) and the control group (blue) in the 3000–400 cm⁻¹ rang. (b) First-derivative transformation of the spectra shown in (a). (c) Comparison of mean raw spectra between sepsis survivors (green) and non-survivors (purple). (d) First-derivative spectra corresponding to the patient groups in (C). (e) Violin plots comparing the integration areas of six identified key spectral regions (2997-2887, 1721-1589, 1589-1478, 1275-1203, 1184-1140, and 1140-956 cm⁻¹) between control and sepsis groups. (f) Violin plots comparing the integration areas of the same six spectral regions between sepsis survivors and non-survivors. ATR-FTIR, Attenuated Total Reflection Fourier Transform Infrared; a.u.

Table 2. Comparison of Baseline Characteristics between the Sepsis Group and the Control Group (N = 146)

Characteristic	Overall N = 146	control N = 68	sepsis N = 78	p-value[1]
gender, n (%)				<0.001
male	87 (60)	30 (44)	57 (73)	
female	59 (40)	38 (56)	21 (27)	
Age, Mean (SD)	51.08 (15.05)	48.28 (12.31)	53.55 (16.80)	0.038
[1]Pearson's Chi-squared test; Wilcoxon rank sum test				

ATR-FTIR spectra of serum samples from sepsis patients and the control group (Table 2) were initially collected. Comparative analyses of spectral features were conducted between non-sepsis and sepsis patients, as well as between patients with favorable and poor prognoses. ATR-FTIR spectra within the wavelength range of 2.5-25 μm, corresponding to the wavenumber range of 4000-400 cm^{-1}, were selected for analysis. The lipid region (3000-2800 cm^{-1}) and the fingerprint region (1800-900 cm^{-1}) of the spectra reflect the characteristics of major biological macromolecules such as lipids, proteins, nucleic acids, and carbohydrates in biological samples[1–3] (Figure 1a, c). Within the 3000-2800 cm^{-1} region, distinct absorption peaks were observed near 2957 cm^{-1}, 2928 cm^{-1}, 2871 cm^{-1}, and 2856 cm^{-1}. In the 1800-900 cm^{-1} region, characteristic peaks were identified near 1636 cm^{-1}, 1538 cm^{-1}, 1453 cm^{-1}, 1396 cm^{-1}, 1298 cm^{-1}, 1238 cm^{-1}, 1169 cm^{-1}, 1125 cm^{-1}, and 1081 cm^{-1}, with detailed functional group assignments provided for the main bands (Table 3 [24,29–31]). To capture more subtle spectral differences, corresponding first- transformed spectra were plotted (Figure 1b, d). Compared to the control group, the characteristic serum spectral peaks of the sepsis group at 2928 cm^{-1}, 1636 cm^{-1}, 1298 cm^{-1}, 1125 cm^{-1}, and 1081 cm^{-1} shifted to 2925 cm^{-1}, 1639 cm^{-1}, 1301 cm^{-1}, 1122 cm^{-1}, and 1077 cm^{-1}, respectively (Figure 1b). Compared to patients with a poor prognosis, sepsis patients with a favorable prognosis showed shifts in serum spectral peaks from 1298 cm^{-1}, 1125 cm^{-1}, and 1081 cm^{-1} to 1301 cm^{-1}, 1122 cm^{-1}, and 1077 cm^{-1}, respectively (Figure 1d).

Table 3. The Main Absorption Bands and their Corresponding Assignments [24,29–31]

Wavenumber(cm^{-1})	Assignments
3285	Amide A band (N-H stretching vibration)
3078	Amide B band (N-H stretching)
2956 / 2871	Methyl group vibrations ($\nu_{as}(CH_3)$ and $\nu_s(CH_3)$)
2928 / 2856	Methylene stretching ($\nu_{as}(CH_2)$ and $\nu_s(CH_2)$)
1639	Amide I (attributed to C=O stretching)
1537	Amide II (associated with C-N stretch and N-H bend)
1452	CH_2 scissoring bending in lipid chains
1398	Symmetric deformation of CH_3 groups
1306	Amide III vibration mode
1242	PO_2^- asymmetric stretching in phosphodiester
1169-1159	C-O stretching (Protein/Carbohydrate side chains)
1140-956	Skeletal vibrations (C-O, C-C) of carbohydrates and DNA/RNA
ν_s: Symmetric stretching vibration; ν_{as}: Asymmetric stretching; PO_2^-: Phosphodiester group.	

derivative

Based on the above spectral features, we selected six key characteristic regions corresponding to these peaks for integrated area calculations: 2997-2887 cm^{-1}, 1721-1589 cm^{-1}, 1589-1478 cm^{-1}, 1275-1203 cm^{-1}, 1184-1140 cm^{-1}, and 1140-956 cm^{-1}. Differential analysis revealed that the integrated areas of these six characteristic regions all showed significant differences between the sepsis group and the control group ($p < 0.05$, Figure 1e). Compared to the poor prognosis group, sepsis patients in the favorable prognosis group exhibited a significantly decreased integrated area in the 1184-1140 cm^{-1} region ($p = 0.012$) and a significantly increased integrated area in the 1140-956 cm^{-1} region ($p = 0.012$, Figure 1f).

3.2 Machine Learning Diagnostic Model Based on Spectral Features

To evaluate the potential clinical value of serum spectral features for sepsis diagnosis, this study constructed a classification model based on the serum infrared spectral data of 146 subjects (Figure 2a). First, the enrolled samples were randomly split into a training/validation set (N=102) and an independent test set (N=44) at a ratio of 7:3. The training set was used for 10-

fold cross-validation and hyperparameter optimization, while the independent test set was used to assess the generalizability of the final model. Regarding feature engineering, integrated areas of six key characteristic spectral bands were extracted to construct feature vectors. Sixteen classical machine learning algorithms were introduced for comparative performance screening (Figure 2b).

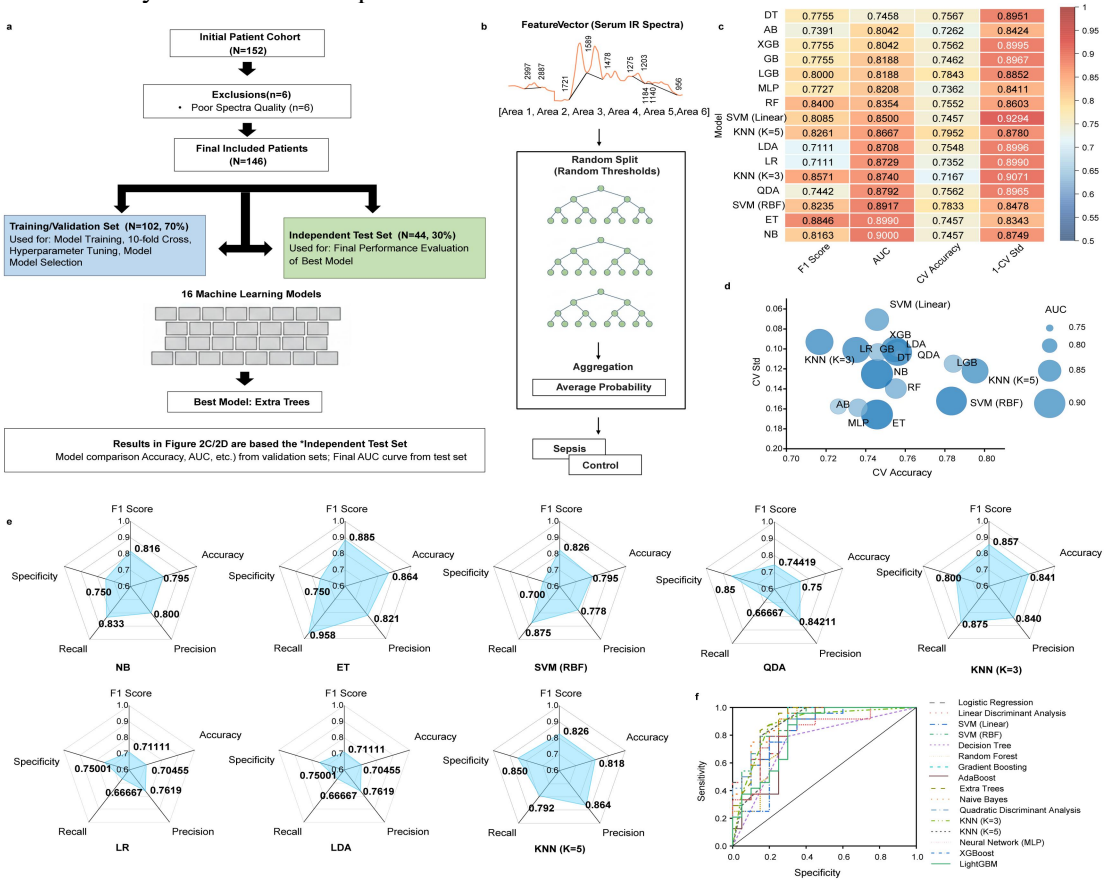

Figure 2. Construction and Performance Evaluation of Machine Learning Models for Sepsis Diagnosis. (a) Flowchart of the patient recruitment, exclusion criteria, dataset partitioning (70% training/validation, 30% independent test), and model selection process. (b) Schematic representation of the machine learning workflow, utilizing the six extracted spectral feature areas as input vectors for classification (Sepsis vs. Control). (c) Heatmap displaying the performance metrics (F1 Score, AUC, CV Accuracy, and 1-CV Std) of 16 different machine learning algorithms on the validation set. Redder colors indicate better performance. (d) Machine learning model performance comparison. Bubble plot visualizing the trade-off between model stability (CV Std) and accuracy (CV Accuracy). Bubble size and color gradient represent the AUC value. (e) Radar charts illustrating the multi-dimensional performance (Specificity, Accuracy, Precision, Recall, F1-Score) of the top-performing models. (f) Receiver Operating Characteristic (ROC) curves of various algorithms on the independent test set, demonstrating the diagnostic capability of spectral features. AUC, Area Under the Curve; CV, Cross-Validation; Std, Standard Deviation; ROC, Receiver Operating Characteristic; ET, Extra Trees; RF, Random Forest; XGB, XGBoost; GB, Gradient Boosting; LGB, LightGBM; MLP, Multilayer Perceptron; SVM, Support Vector Machine; RBF, Radial Basis Function; KNN, K-Nearest Neighbors; LDA, Linear Discriminant Analysis; LR, Logistic Regression; QDA, Quadratic Discriminant Analysis; NB, Naive Bayes; AB, AdaBoost; DT, Decision Tree.

To comprehensively evaluate the performance of each algorithm, multiple metrics were adopted, including F1 Score, AUC, and CV Accuracy. Additionally, to balance both prediction accuracy and robustness, 1-CV Std was introduced as a stability assessment metric. Results showed that models based on algorithms such as Extra Trees, NB, and SVM (RBF) all performed excellently on the validation set (all scores > 0.7) (Figure 2c). Bubble chart analysis

of the distribution between CV Accuracy and CV Std visually revealed the trade-offs between accuracy and stability across different models (Figure 2d). After comprehensive evaluation, the Extra Trees model demonstrated the most outstanding performance on the independent test set, achieving an F1 Score of 0.8846 and an AUC value of 0.8990 (Figure 2d). For algorithms with excellent AUC performance, their characteristics were multi-dimensionally profiled via radar charts, collectively illustrating their performance spectrum across five dimensions: F1 Score, Accuracy, Precision, Recall, and Specificity (Figure 2e). The results indicated that the Extra Trees model achieved a Recall of 0.9580, suggesting its high clinical potential for reducing missed diagnoses. ROC curve analysis (Figure 2f) further confirmed that the serum spectral machine learning model can distinguish sepsis patients from healthy controls with high specificity and sensitivity.

3.3 Predicting Clinical Outcomes in Sepsis

Table 4. Comparison of Baseline Characteristics between the Alive and the Dead Sepsis Patiets (N = 83)

Characteristic	Overall N = 83	alive N = 73	dead N = 10	p-value[1]
gender, n (%)				>0.99
female	23 (28)	20 (27)	3 (30)	
male	60 (72)	53 (73)	7 (70)	
Age, Mean (SD)	54.82 (17.04)	53.92 (16.49)	61.40 (20.37)	0.17
Systolic Pressure, Median (IQR)	131.00 (120.00 – 140.00)	131.00 (121.00 – 139.00)	130.00 (116.00 – 141.00)	0.96
Diastolic Pressure, Median (IQR)	68.00 (64.00 – 74.00)	68.00 (64.00 – 74.00)	69.00 (64.00 – 71.00)	0.53
MAP, Median (IQR)	90.33 (84.00 – 94.33)	90.67 (84.33 – 94.33)	88.00 (83.33 – 93.67)	0.66
Respiratory Rate, Median (IQR)	21.00 (20.00 – 22.00)	20.00 (20.00 – 22.00)	21.50 (20.00 – 23.00)	0.44
Unknown	2	2	0	
D-Dimer, Median (IQR)	6.63 (3.02 – 12.56)	6.63 (3.20 – 12.32)	7.15 (2.40 – 13.12)	0.79
Unknown	1	1	0	
PLT, Median (IQR)	176.00 (109.00 – 291.00)	198.00 (128.00 – 295.00)	108.00 (53.00 – 134.00)	0.015
HCT, Mean (SD)	27.96 (5.67)	27.89 (5.62)	28.42 (6.36)	0.86
Bilirubin, Median (IQR)	16.99 (12.56 – 26.43)	18.35 (13.26 – 26.43)	12.36 (10.34 – 19.00)	0.15
Creatinine, Median (IQR)	68.40 (47.70 – 127.10)	67.30 (47.60 – 118.40)	99.15 (54.00 – 147.20)	0.31
Heart Rate, Mean (SD)	92.59 (16.15)	92.79 (16.51)	91.10 (13.90)	0.70
Troponin I, Median (IQR)	0.02 (0.01 – 0.09)	0.02 (0.01 – 0.09)	0.08 (0.03 – 0.18)	0.016
Unknown	14	14	0	
CK-MB, Median (IQR)	0.93 (0.37 – 1.93)	0.92 (0.35 – 1.91)	1.18 (0.51 – 11.20)	0.23
Unknown	13	13	0	
Myoglobin, Median (IQR)	135.20 (77.99 – 249.50)	121.50 (62.39 – 215.40)	240.25 (136.90 – 420.10)	0.067
Unknown	14	14	0	
NT-proBNP, Median (IQR)	2,100.00 (607.00 – 5,290.00)	1,900.00 (434.00 – 4,670.00)	4,045.00 (1,880.00 – 11,000.00)	0.042
Unknown	14	14	0	
CRP, Median (IQR)	60.86 (21.21 – 114.57)	61.96 (22.54 – 112.54)	37.09 (8.19 – 116.60)	0.64
Unknown	7	6	1	
WBC, Median (IQR)	10.33 (7.66 – 13.64)	10.55 (7.88 – 13.52)	8.53 (6.73 – 14.97)	0.40
Temperature, Median (IQR)	36.90 (36.70 – 37.30)	36.90 (36.70 – 37.30)	36.85 (36.50 – 37.20)	0.45
PCT, Median (IQR)	0.94 (0.29 – 3.89)	0.94 (0.32 – 3.05)	1.21 (0.28 – 12.16)	0.67
IL-6, Median (IQR)	53.08 (26.89 – 108.98)	53.07 (28.42 – 93.60)	108.08 (18.51 – 293.52)	0.40
Unknown	4	4	0	
[1] Fisher's exact test; Wilcoxon rank sum test				

To assess the prognostic risk of sepsis patients, this study first constructed a clinical feature-based scoring system (Figure 3a). Through ROC analysis of patient clinical indicators (Table 4), key prognostic factors with AUC > 0.7 were identified, including NT-proBNP (AUC = 0.741), Troponin I (AUC = 0.745), and Platelets (AUC = 0.738) (Figure 3b). Using the Youden Index, cut-off values for each indicator were determined, and an additive scoring model was established: high TnI, high NT-proBNP, and low PLT were each assigned 1 point, resulting in a total score range of 0–3. Patients were subsequently categorized into a low-risk group (Score 0–1) and a high-risk group (Score 2–3). Survival analysis revealed a significant difference in 28-day clinical outcomes between the two groups (Figure 3c). The survival rate in the low-risk group was 100%, while the mortality rate in the high-risk group reached 24%, indicating that this clinical scoring system possesses effective risk-stratification capability.

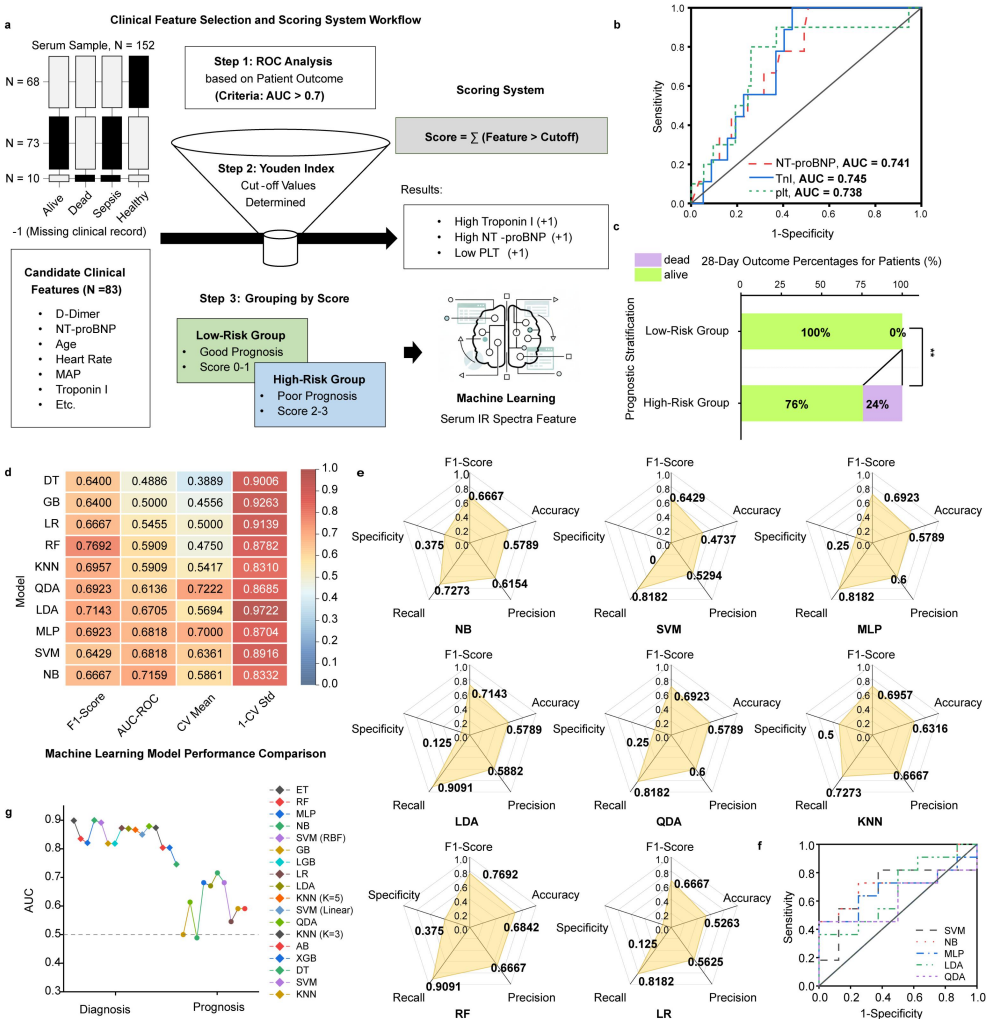

Figure 3. Development of a Clinical Prognostic Scoring System and Spectral-Based Prognostic Prediction. (a) Workflow for the clinical feature selection and prognostic scoring system construction. Clinical features with an AUC > 0.7 were selected, and cut-off values were determined by the Youden Index. (b) ROC curves for the three selected clinical features: NT-proBNP, Troponin I (TnI), and Platelets (PLT), for predicting patient mortality. (c) Stacked bar chart showing the significant difference in 28-day mortality rates between the Low-Risk Group (Score 0-1) and High-Risk Group (Score 2-3) stratified by the clinical scoring system (p < 0.01). (d) Heatmap of performance metrics (F1-Score, AUC, CV Mean, CV Std) for 16 machine learning models trained on serum spectral features to predict patient prognosis. (e) Radar charts depicting the performance profiles of algorithms with an AUC > 0.5. (f) ROC curves of the machine learning models for prognostic prediction. (g) Comparison of AUC values for all algorithms between the diagnostic task (Diagnosis) and the prognostic task (Prognosis), highlighting the performance gap between the two applications. NT-proBNP, N-terminal pro-brain natriuretic peptide; TnI, Troponin I; PLT, Platelets; AUC, Area Under the Curve; ROC, Receiver Operating Characteristic; CV, Cross-Validation. (See Figure 2 for Machine Learning model abbreviations).**

Further investigation was conducted to evaluate the predictive ability of spectral features for clinical outcomes (favorable vs. poor prognosis) in sepsis patients. Ten machine learning algorithms were trained and evaluated based on six characteristic spectral regions (Figure 3d). Results showed that the Naive Bayes model performed relatively well, achieving an AUC of 0.7159 and an F1-score of 0.6667. Radar charts further illustrated the comprehensive performance of algorithms with AUC > 0.5 across dimensions including specificity and accuracy (Figure 3e). Among them, the LDA algorithm achieved a Recall of 0.9091, though its overall balance was inferior to that of the NB model. ROC curve analysis (Figure 3f) further

confirmed that the serum spectral machine learning model holds considerable potential for predicting prognosis and severity in sepsis patients. Figure 3g summarizes and compares the AUC distributions for the diagnostic task versus the prognostic task. In the current dataset, individual spectral features generally demonstrated superior performance in the diagnostic task compared to the prognostic task, which may be attributed to the relatively smaller sample size available for prognostic evaluation.

3.4 Correlation of Spectral Features with Clinical Biomarkers

To explore the clinical significance of changes in serum spectral features, we performed a correlation analysis between the integrated areas of the six extracted spectral regions and patients' clinical characteristics (Figure 4a). The results revealed significant correlations between the spectral features and multiple indicators reflecting infection severity, organ function, and baseline constitution.

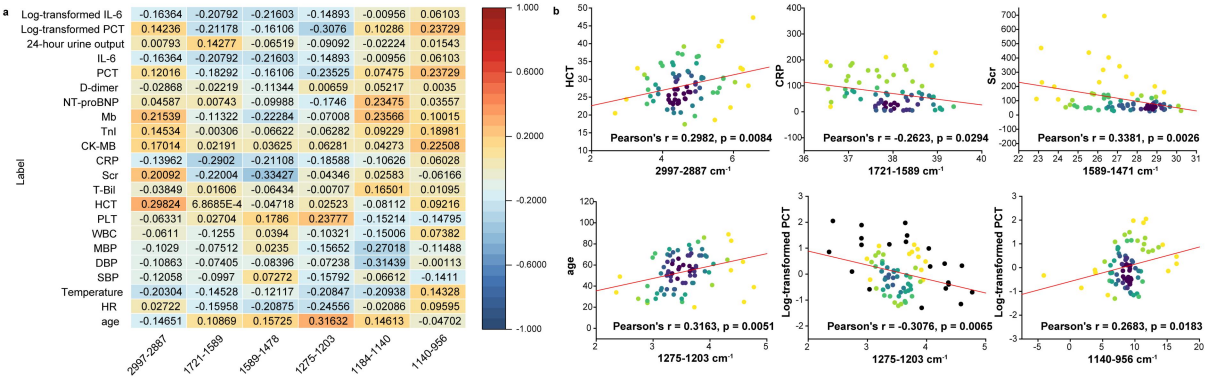

Figure 4. Biological Interpretability and Correlation Analysis between Spectral Features and Clinical Biomarkers. (a) Heatmap showing the correlation between spectral peak areas and clinical parameters. Colors indicate the strength of the correlation. (b) Scatter plots with linear regression fitting illustrating significant correlations between specific spectral regions and clinical markers, including: HCT vs. 2997-2887 cm⁻¹ (Lipids), CRP vs. 1721-1589 cm⁻¹, Scr vs. 1589-1478 cm⁻¹, Age vs. 1275-1203 cm⁻¹, Log-transformed PCT vs. 1275-1203 cm⁻¹, and Log-transformed PCT vs. 1140-956 cm⁻¹. IL-6, Interleukin-6; PCT, Procalcitonin; CRP, C-reactive protein; Scr, Serum creatinine; T-Bil, Total Bilirubin; HCT, Hematocrit; PLT, Platelets; WBC, White Blood Cell count; MBP, Mean Arterial Pressure; DBP, Diastolic Blood Pressure; SBP, Systolic Blood Pressure; HR, Heart Rate; NT-proBNP, N-terminal pro-brain natriuretic peptide; Mb, Myoglobin; TnI, Troponin I; CK-MB, Creatine Kinase-MB isoenzyme.

Regarding infection and inflammatory indicators, the integrated area of the 1275-1203 cm^{-1} region (asymmetric stretching vibration of phosphodiester bonds in nucleic acids and phospholipids) showed a significant negative correlation with log-transformed PCT ($r = -0.3076$, $p = 0.0065$), whereas the 1140-956 cm^{-1} region (nucleic acid/carbohydrate vibrations) exhibited a significant positive correlation with it ($r = 0.2683$, $p = 0.0183$). Simultaneously, the 1721-1589 cm^{-1} region (amide I band) demonstrated a significant negative correlation with CRP ($r = -0.2623$, $p = 0.0294$) (Figure 4b). The 2997-2887 cm^{-1} region (lipid region) was significantly positively correlated with Scr ($r = 0.2982$, $p = 0.0084$). Furthermore, the 1275-1203 cm^{-1} region showed a positive correlation with patient age ($r = 0.3373$, $p = 0.027$).

4. Discussion

The present study demonstrates that serum ATR-FTIR spectroscopy, coupled with machine learning, can generate clinically relevant diagnostic and prognostic information in sepsis. This approach directly addresses two persistent and intertwined challenges in critical care: the need for rapid, accurate diagnosis at the point of clinical suspicion, and the ability to perform early, reliable risk stratification in a highly heterogeneous patient population [32].

Our findings confirm the potential of vibrational spectroscopy as a source of high-dimensional biological data. The superior diagnostic performance of the Extra Trees model (AUC 0.899, Recall 0.958) is particularly noteworthy. The high recall rate suggests this approach could effectively reduce missed diagnoses—a critical advantage given the fatal consequences of delayed antimicrobial therapy. This spectral signature appears to capture a systemic

biochemical shift distinct from the non-septic inflammatory state. The construction of a simple clinical score (based on NT-proBNP, Troponin I, and platelets) and the separate development of a spectral-only prognostic model (AUC 0.716) highlight complementary strategies. The clinical score offers immediate, interpretable risk stratification using common laboratory tests, while the spectral model points toward a future where a single, rapid test could simultaneously inform both diagnosis and prognosis.

The observed correlations between specific spectral regions and validated clinical biomarkers (PCT, CRP, Scr) provide crucial biological plausibility. For instance, the negative correlation between the amide I band (1721-1589 cm^{-1}) and CRP aligns with the known consumption or conformational changes of acute-phase proteins during severe inflammation. These links ground the "black box" of machine learning output in established pathophysiology, strengthening the case for clinical translation.

Several limitations must be acknowledged. The single-center design and moderate sample size necessitate validation in larger, multicenter, prospective cohorts. The prognostic model, while promising, performed less robustly than the diagnostic model, likely reflecting the greater complexity of predicting outcome and the smaller sample size for this analysis. Future work must focus on standardizing pre-analytical procedures and integrating spectral data with electronic health records to build dynamic, multi-modal prediction models.

5. Conclusion

In conclusion, this study moves beyond proof-of-concept to present a pragmatic framework for leveraging biochemical fingerprinting in sepsis. By providing a rapid, composite readout of host response, this technology could evolve into a valuable adjunctive tool at the bedside. It holds the promise of transforming sepsis management from a reactive paradigm to a more precise, pre-emptive strategy—ultimately contributing to the urgent goal of improving outcomes in this deadly syndrome.

References

[1] Fleischmann C, Scherag A, Adhikari N K J, et al. Assessment of Global Incidence and Mortality of Hospital-treated Sepsis. AMERICAN JOURNAL OF RESPIRATORY AND CRITICAL CARE MEDICINE, 2016, 193(3): 259-272.

[2] Rudd K E, Johnson S C, Agesa K M, et al. Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the Global Burden of Disease Study. LANCET, 2020, 395(10219): 200-211.

[3] Fleischmann-Struzek C, Mellhammar L, Rose N, et al. Incidence and mortality of hospital- and ICU-treated sepsis: results from an updated and expanded systematic review and meta-analysis. Intensive Care Medicine, 2020, 46(8): 1552-1562.

[4] Weng L, Zeng X Y, Yin P, et al. Sepsis-related mortality in China: a descriptive analysis. Intensive Care Medicine, 2018, 44(7): 1071-1080.

[5] Evans L, Rhodes A, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. Intensive Care Medicine, 2021, 47(11): 1181-1247.

[6] Cecconi M, Evans L, Levy M, et al. Sepsis and septic shock. LANCET, 2018, 392(10141): 75-87.

[7] Seymour C W, Gesten F, Prescott H C, et al. Time to Treatment and Mortality during Mandated Emergency Care for Sepsis. NEW ENGLAND JOURNAL OF MEDICINE, 2017, 376(23): 2235-2244.

[8] Weng L, Xu Y, Yin P, et al. National incidence and mortality of hospitalized sepsis in China. Critical Care (London, England), 2023, 27(1): 84.

[9] Bolanaki M, Winning J, Slagman A, et al. Biomarkers Improve Diagnostics of Sepsis in Adult Patients With Suspected Organ Dysfunction Based on the Quick Sepsis-Related Organ Failure Assessment (qSOFA) Score in the Emergency Department. Critical Care Medicine, 2024, 52(6): 887-899.

[10] Wang Q, Peng G, Gan L, et al. The Value of Interleukin-10 in the Early Diagnosis of Neonatal Sepsis: A Meta-Analysis. Pediatric Critical Care Medicine: A Journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies, 2021, 22(9): e492-e501.

[11] Zaki H A, Bensliman S, Bashir K, et al. Accuracy of procalcitonin for diagnosing sepsis in adult patients admitted to the emergency department: a systematic review and meta-analysis. Systematic Reviews, 2024, 13(1): 37.

[12] Dhudasia M B, Benitz W E, Flannery D D, et al. Diagnostic Performance and Patient Outcomes With C-Reactive Protein Use in Early-Onset Sepsis Evaluations. The Journal of Pediatrics, 2023, 256: 98-104.e6.

[13] Zhang W, Wang W, Hou W, et al. The diagnostic utility of IL-10, IL-17, and PCT in patients with sepsis infection. Frontiers in Public Health, 2022, 10: 923457.

[14] Friedman N, Yochpaz S, Zirkin S, et al. C-reactive protein and the neonatal early-onset sepsis calculator for the diagnosis of neonatal sepsis. European Journal of Clinical Microbiology & Infectious Diseases: Official Publication of the European Society of Clinical Microbiology, 2021, 40(6): 1227-1234.

[15] Cong S, Ma T, Di X, et al. Diagnostic value of neutrophil CD64, procalcitonin, and interleukin-6 in sepsis: a meta-analysis. BMC infectious diseases, 2021, 21(1): 384.

[16] Baker M J, Faulds K. Fundamental developments in clinical infrared and Raman spectroscopy. Chemical Society Reviews, 2016, 45(7): 1792-1793.

[17] Kumar S, Verma T, Mukherjee R, et al. Raman and infra-red microspectroscopy: towards quantitative evaluation for clinical research by ratiometric analysis. Chemical Society Reviews, 2016, 45(7): 1879-1900.

[18] Budidha K, Mamouei M, Baishya N, et al. Identification and Quantitative Determination of Lactate Using Optical Spectroscopy-Towards a Noninvasive Tool for Early Recognition of Sepsis. Sensors (Basel, Switzerland), 2020, 20(18): 5402.

[19] Montes-Narváez O, García-Juárez M, Beltrán-Pérez G, et al. ATR-FTIR spectroscopy to evaluate serum protein expression in a murine cerebral ischemia model. Spectrochimica Acta. Part A, Molecular and Biomolecular Spectroscopy, 2025, 326: 125261.

[20] Ainsworth C. Infrared spectrometry guides cancer treatment. Nature, 2024.

[21] Lilo T, Morais C L M, Shenton C, et al. Revising Fourier-transform infrared (FT-IR) and Raman spectroscopy towards brain cancer detection. Photodiagnosis and Photodynamic Therapy, 2022, 38: 102785.

[22] Su K Y, Lee W L. Fourier Transform Infrared Spectroscopy as a Cancer Screening and Diagnostic Tool: A Review and Prospects. Cancers, 2020, 12(1): 115.

[23] Wang X, Shen X, Sheng D, et al. FTIR spectroscopic comparison of serum from lung cancer patients and healthy persons. Spectrochimica Acta. Part A, Molecular and Biomolecular Spectroscopy, 2014, 122: 193-197.

[24] Tian Y, Fan X, Chen K, et al. Optical biomarker analysis for renal cell carcinoma obtained from preoperative and postoperative patients using ATR-FTIR spectroscopy. Spectrochimica Acta. Part A, Molecular and Biomolecular Spectroscopy, 2024, 318: 124426.

[25] Eissa T, Leonardo C, Kepesidis K V, et al. Plasma infrared fingerprinting with machine learning enables single-measurement multi-phenotype health screening. Cell Reports. Medicine, 2024, 5(7): 101625.

[26] Chechekina O G, Tropina E V, Fatkhutdinova L I, et al. Machine learning assisted rapid approach for quantitative prediction of biochemical parameters of blood serum with FTIR spectroscopy. Spectrochimica Acta. Part A, Molecular and Biomolecular Spectroscopy, 2025, 326: 125283.

[27] Eshel Y D, Sharaha U, Beck G, et al. Monitoring the efficacy of antibiotic therapy in febrile pediatric oncology patients with bacteremia using infrared spectroscopy of white blood cells-based machine learning. Talanta, 2024, 270: 125619.

[28] Korb E, Bağcıoğlu M, Garner-Spitzer E, et al. Machine Learning-Empowered FTIR Spectroscopy Serum Analysis Stratifies Healthy, Allergic, and SIT-Treated Mice and Humans. Biomolecules, 2020, 10(7): 1058.

[29] Ghimire H, Venkataramani M, Bian Z, et al. ATR-FTIR spectral discrimination between normal and tumorous mouse models of lymphoma and melanoma from serum samples. Scientific Reports, 2017, 7(1): 16993.

[30] Ghimire H, Garlapati C, Janssen E A M, et al. Protein Conformational Changes in Breast Cancer Sera Using Infrared Spectroscopic Analysis. Cancers, 2020, 12(7): 1708.

[31] Baker M J, Hussain S R, Lovergne L, et al. Developing and understanding biofluid vibrational spectroscopy: a critical review. Chemical Society Reviews, 2016, 45(7): 1803-1818.

[32] Wu M, Pu K, Jiang T, et al. Early label-free analysis of mitochondrial redox states by Raman spectroscopy predicts septic outcomes. Journal of Advanced Research, 2021, 28: 209-219.

