

# Machine Learning-Based Prediction Model for Multidrug-Resistant Organisms Infections: Performance Evaluation and Interpretability Analysis

Wenting Zhao<sup>1,2,\*</sup>, Pei Sun<sup>1,2,\*</sup>, Wei Li<sup>3</sup>, Linping Shang<sup>2,4</sup>

<sup>1</sup>College of Nursing, Changzhi Medical College, Changzhi, Shanxi, People's Republic of China; <sup>2</sup>College of Nursing, Shanxi Medical University, Taiyuan, Shanxi, People's Republic of China; <sup>3</sup>Infection Management Department, the First Hospital of Shanxi Medical University, Taiyuan, Shanxi, People's Republic of China; <sup>4</sup>Nursing Department, the First Hospital of Shanxi Medical University, Taiyuan, Shanxi, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Linping Shang, College of Nursing, Shanxi Medical University, Taiyuan, Shanxi, People's Republic of China, Email shanglp2002@163.com

**Background:** Multidrug-resistant organism (MDRO) infections pose a significant global health threat, particularly in intensive care units (ICUs), where delayed identification exacerbates clinical outcomes. Although machine learning (ML) holds promise for infection prediction, the opaque nature of complex algorithms impedes clinical adoption. This study evaluated an interpretable machine learning model incorporating SHapley Additive exPlanations (SHAP) to predict MDRO infections in ICU patients.

**Methods:** A retrospective cohort study was conducted on 888 ICU patients (2020–2022) from a tertiary hospital in China. Following TRIPOD guidelines, key predictors were identified using Lasso regression from a comprehensive set of clinical variables, including demographics, treatments, and laboratory data. Six machine learning algorithms—Neural Networks, Random Forests, Support Vector Machines, Logistic Regression, Decision Trees, and Gaussian Naive Bayes—were evaluated based on AUC, accuracy, and calibration curves. SHAP analysis provided both global and local interpretability.

**Results:** Among 825 eligible cases (375 MDRO infections), the Random Forest model exhibited the highest performance (AUC = 0.83, accuracy = 76.7%). SHAP analysis identified urinary catheterization, ventilator use, and prolonged antibiotic exposure as key modifiable risk factors. Case-level interpretation via dynamic force plots illustrated individualized risk stratification. Decision curve analysis indicated clinical utility within probability thresholds of 0.44–0.60.

**Conclusion:** This study establishes an interpretable prediction framework integrating RF algorithms with SHAP explainability, balancing predictive accuracy with clinical transparency. The model's dynamic visualization capabilities support individualized risk assessment and evidence-based antimicrobial stewardship. Integration into hospital information systems with real-time dashboards could enhance early intervention strategies.

**Keywords:** MDRO, machine learning, prediction, intensive care unit

## Introduction

Multidrug-resistant organisms (MDRO) are microorganisms, primarily bacteria, that have acquired resistance to three or more antimicrobial classes through diverse genetic mechanisms, significantly restricting therapeutic options.<sup>1</sup> This classification encompasses both Gram-positive and Gram-negative bacteria. The escalating threat posed by MDRO to public health results in millions of fatalities annually.<sup>2</sup> In recognition of this crisis, the World Health Organization identified antimicrobial resistance as one of the top 10 global health threats in 2019.<sup>3</sup>

MDRO infections in intensive care unit (ICU) patients are particularly concerning due to their profound impact on treatment efficacy.<sup>4</sup> These infections are associated with increased inpatient mortality, higher readmission rates,

prolonged hospital stays, and escalated medical costs.<sup>5–7</sup> A retrospective study on carbapenem-resistant Gram-negative bacilli (CRGNB) infections in Malaysia reported pneumonia (40.7%) and bacteremia (25.5%) as the most prevalent manifestations, with an overall hospital mortality rate reaching 41.4%.<sup>8</sup>

Furthermore, ICU patients with MDRO infections exhibit significantly higher mortality rates than those with non-MDRO infections.<sup>9</sup> The burden of MDRO is not confined to healthcare settings. A study in Shenzhen, China, reported an MDRO carriage rate of 26.7% among individuals with no recent hospitalization or antibiotic use within the preceding six months.<sup>10</sup>

Recent epidemiological studies in respiratory ICUs emphasize the necessity of early detection and rigorous monitoring of MDRO strains to mitigate antimicrobial resistance. Given the substantial morbidity associated with ICU-acquired infections, accessible risk prediction tools are urgently needed.<sup>11</sup> Effective containment strategies are essential to curtail the spread of MDRO within medical institutions and communities. Healthcare workers primarily rely on isolation protocols and stringent hand hygiene to prevent cross-transmission from MDRO-infected patients. However, delays in infection surveillance remain a critical challenge. Consequently, establishing high-performance predictive models is imperative for the early identification of patients at elevated risk of MDRO infections.

The pathogenesis of MDRO infections involves intricate interactions among multiple risk factors, with recent studies highlighting diverse risk profiles across clinical settings. Chen et al identified a broad spectrum of predisposing factors, including prior antimicrobial therapy, inappropriate antimicrobial use, chronic pulmonary and liver diseases, neurological conditions, prior MDRO infections, recent or prolonged hospitalization, tracheotomy, mechanical ventilation, nasogastric feeding, nursing home admission, and higher severity of illness scores.<sup>12</sup> While certain risk factors are consistently observed across healthcare environments, variations also exist. For example, Tseng et al emphasized age, cerebrovascular accidents, and antimicrobial exposure as key determinants of multidrug-resistant Gram-negative bacterial infections.<sup>13</sup>

These factors collectively heighten susceptibility to MDRO infection. Patients with cerebrovascular accidents often experience prolonged immobility, predisposing them to pressure ulcers and chronic wounds that compromise skin integrity and foster biofilm formation, thereby facilitating MDRO persistence.<sup>14</sup> Emerging evidence suggests that extracellular polymeric substance (EPS) production within biofilms enhances microbial adhesion and serves as a reservoir for resistance gene transfer, complicating infection management.<sup>14</sup> Additionally, the Katz Index of Independence in Activities of Daily Living (ADL) quantifies patient mobility and self-care capacity, with lower ADL scores correlating with increased infection risk due to immobility and malnutrition, further predisposing individuals to MDRO colonization.<sup>15</sup> A meta-analysis of 108 studies identified age, recent invasive procedures, antibiotic exposure, and prior hospitalization as high-frequency predictors.<sup>16</sup>

Traditional MDRO infection prediction models have predominantly relied on conventional statistical approaches, which impose limitations on feature selection and lack automation.<sup>17,18</sup> For instance, Song et al employed logistic regression in a retrospective cohort study of 444 ICU patients at a tertiary hospital to identify risk factors for carbapenem-resistant *Enterobacteriaceae* (CRE) infections, developing a clinical risk prediction model to guide infection prevention strategies.<sup>19</sup> With advancements in artificial intelligence, machine learning algorithms are increasingly being integrated into medical applications.<sup>20</sup> However, a systematic review of 15 studies revealed that while most utilized traditional statistical models such as logistic regression, only three incorporated machine learning techniques.<sup>21</sup>

Machine learning algorithms excel in processing large datasets and have demonstrated strong predictive performance in various diseases, including sepsis and stroke.<sup>22</sup> Several studies have explored machine learning-based approaches for MDRO infection prediction.<sup>23</sup> For example, Goodman et al developed a clinical decision tree to predict ESBL-positive *Escherichia coli* or *Klebsiella* bacteremia, identifying indwelling vascular devices, age  $\geq 43$  years, recent hospitalization in high-incidence areas, and antimicrobial exposure  $\geq$  six days within six months as key risk factors.<sup>24</sup> The model achieved high positive and negative predictive values (90.8% and 91.9%, respectively). Similarly, Hartwigsen et al constructed an MRSA infection prediction model using electronic medical records, incorporating demographic data, treatment history, clinical nursing notes, vital signs, medications, and laboratory results.<sup>25</sup> Among logistic regression, support vector machines, and random forests, the latter demonstrated the strongest clinical applicability. Li et al integrated data from the People's Liberation Army General Hospital (PLAGH-ICU) and the MIMIC-IV database to develop a predictive model using Random Forests, XGBoost, and logistic regression, which was subsequently embedded

into the hospital's Electronic Health Record (EHR) system to provide real-time alerts for precision antimicrobial therapy.<sup>26</sup>

Despite their high predictive accuracy, machine learning models face significant adoption barriers in clinical settings due to the inherent “black box” nature of complex algorithms, which obscures the decision-making process and fosters skepticism among healthcare providers.<sup>27</sup> This lack of interpretability limits their practical utility, as clinicians require transparent and explainable outputs to inform patient management.

The introduction of SHapley Additive exPlanations (SHAP) has significantly enhanced machine learning interpretability.<sup>28</sup> By quantifying each clinical feature's contribution to risk prediction, SHAP allows clinicians to visualize how specific factors—such as mechanical ventilation duration or indwelling urinary catheter use—impact MDRO infection risk. For instance, SHAP analysis has demonstrated that prolonged invasive procedures substantially elevate infection risk, offering actionable insights for early intervention. This approach has been successfully applied in sepsis and right ventricular dysfunction prediction in nonischemic cardiomyopathy.<sup>29,30</sup>

This study aimed to develop an machine learning-based early prediction model for MDRO infections in ICU patients, addressing the critical need for timely, data-driven interventions in intensive care. By retrospectively analyzing three years of ICU patient data and evaluating six machine learning algorithms, the study integrated routine clinical indicators with advanced predictive analytics and SHAP interpretability. This dual emphasis on predictive accuracy and explainability is expected to enhance risk stratification, elucidate underlying mechanisms, and facilitate targeted prevention strategies, ultimately mitigating the burden of MDRO infections in high-risk populations.

## Methods

### Study Design

This retrospective modeling study employed a three-step framework encompassing model development, validation, and interpretability analysis. Individuals aged 18 years or older with a documented infection between 2020 and 2022 were included. Six predictive models were constructed using different machine learning algorithms, with the highest-performing model selected for further interpretation via SHAP analysis. Model developed adhered to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) guidelines (Figure 1).<sup>31</sup> Ethical approval was obtained from the Ethics Committee of the First Hospital of Shanxi Medical University (No. 2022-K159), and the study complied with the Declaration of Helsinki. All identifiable patient information was anonymized, retaining only essential analytical data.

### Study Setting

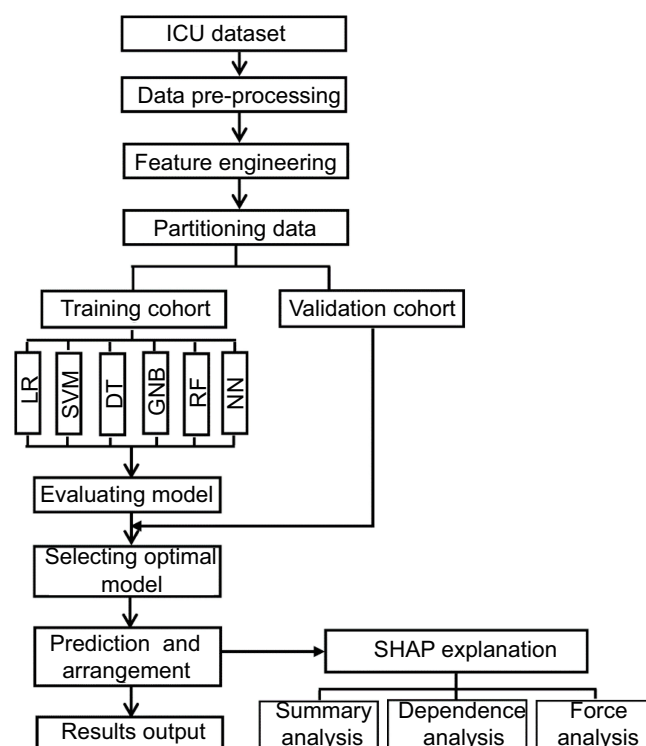
The dataset was derived from a tertiary hospital in Shanxi Province, China, a 2000-bed facility offering comprehensive medical, teaching, rehabilitation, and research services. Participants were recruited from both urban and rural areas across multiple provinces.

### Participants

The study cohort included ICU patients admitted between January 2020 and December 2022, with inclusion criteria requiring a hospital stay of at least 48 hours and an age of 18 years or older. Exclusion criteria encompassed: (a) cases with missing non-critical data exceeding 30% of the total dataset, (b) patients who did not undergo microbial testing during hospitalization, and (c) cases with missing critical data.

### MDRO Definition and Outcome Measurement

MDROs were defined to include carbapenem-resistant *Enterobacteriaceae*, carbapenem-resistant *Acinetobacter baumannii*, carbapenem-resistant *Pseudomonas aeruginosa*, multidrug-resistant *Acinetobacter baumannii*, pan-resistant *Acinetobacter baumannii*, methicillin-resistant *Staphylococcus aureus*, vancomycin-resistant *Enterococcus faecalis*, and multidrug-resistant *Pseudomonas aeruginosa*. The primary outcome was the incidence of MDRO infections.



**Figure 1** Study flowchart.

**Abbreviations:** ICU, Intensive care unit; NN, Neural Network; RF, Random Forest; SVM, Support Vector Machine; LR, Logistic Regression; DT, Decision Tree; GNB, Gaussian Naive Bayes; AUC, Area under the curve; SHAP, SHapley Additive exPlanations.

## Diagnostic Criteria

Diagnosis was based on a combination of clinical manifestations and laboratory findings, including bacterial culture and antimicrobial susceptibility testing performed by the bacteriology laboratory, with confirmation by both clinicians and infection control specialists.

## Variable Selection

To ensure model robustness and mitigate the risks of overfitting or underperformance, variable selection was guided by a comprehensive literature review and in-depth discussions within the research team.<sup>17,32,33</sup> The selected variables were categorized into five domains: (1) demographic characteristics (eg, age, gender, marital status), (2) clinical interventions (eg, invasive procedures, mechanical ventilation, urinary catheterization), (3) clinical nursing assessments recorded within 24 hours of ICU admission (eg, nutritional status, pressure injury risk, mobility scores), (4) antibiotic administration parameters (eg, duration, type, frequency), and (5) comorbidities and primary diagnoses documented at any time during hospitalization (Table S1).

## Model Construction and Validation

Six machine learning binary classifiers were developed to predict MDRO infections: Neural Network (NN), Random Forest (RF), Gaussian Naive Bayes (GNB), Logistic Regression (LR), Support Vector Machine with a radial basis function (SVM), and Decision Tree (DT). These models were trained, tested, and validated using the collected dataset, with performance assessed based on accuracy, AUC, sensitivity, and specificity.

## Sample Size Determination

The sample size was determined using the events per variable (EPV) approach, ensuring statistical robustness.<sup>34</sup> According to EPV guidelines, the minimum required number of events (the lesser of positive or negative cases) must be at least ten times the number of independent variables.

LASSO regression was applied to refine the initial 45 predictors, yielding 31 key variables. The minimum required events were calculated as  $\text{EPV threshold} \times \text{number of variables} = 10 \times 31 = 310$ . To account for potential data attrition, the sample size was increased by 15%, resulting in a minimum requirement of 357 MDRO infection cases.

## Data Extraction

Data were sourced from two primary systems: (1) the Hospital Medical Record System, which provided demographic data, underlying disease diagnoses, admission assessments, clinical nursing scores, and records of invasive procedures, and (2) the Xinglin Hospital Infection Real-Time Monitoring System, which contained antibiotic usage, and infection-related records.

## Statistical Analysis

Statistical analyses were conducted using SPSS 26.0 and Python 3.7. The Shapiro–Wilk test ( $\alpha = 0.05$ ) was used to assess the normality of continuous variables. Normally distributed data were presented as mean  $\pm$  standard deviation ( $\bar{x} \pm s$ ) and compared using the independent samples *t*-test, while non-normally distributed data were reported as median (25th percentile, 75th percentile) [*M* (*Q*<sub>1</sub>, *Q*<sub>3</sub>)] and analyzed using the Mann–Whitney *U*-test. Categorical variables were expressed as frequencies and percentages, with comparisons performed using the Chi-square test ( $\chi^2$  test) or Fisher's exact test, as appropriate. A two-tailed *p*-value  $< 0.05$  was considered statistically significant.

## Handling Missing Data

Variables with more than 30% missing data were excluded from the analysis. For variables with less than 30% missing data, multiple imputation by chained equations (MICE) was performed using SPSS 26.0. Five imputation iterations were conducted, and the estimates from the final iteration were used to replace missing values, ensuring data integrity and robustness.

## Data Partitioning and Model Development

The study sample was randomly partitioned into training and test sets in a 7:3 ratio. Model development and evaluation were conducted using Python, with the pandas library handling data preprocessing and the Sklearn library facilitating the construction of six machine learning models: LR, NN, DT, RF, SVM, and GNB.

## Model Evaluation Metrics

Predictive performance was assessed using multiple evaluation metrics, including accuracy, sensitivity, specificity, error rate, F1 score, Brier score, G-means, and Area Under the Curve (AUC). The models were compared based on AUC, accuracy, and specificity, with the highest-performing algorithm selected as the optimized model.

## Calibration and Clinical Utility

Calibration curves were constructed to compare predicted and observed MDRO infection rates, while decision curve analysis (DCA) was used to assess the clinical utility of the model across different probability thresholds.

## Interpretability Analysis

SHAP was applied for feature attribution analysis, quantifying the contribution of individual variables to model predictions. The model demonstrating optimal predictive performance in the validation cohort was designated as the final model.

# Results

## General Clinical Data

A total of 888 cases were initially collected for this study. After applying the inclusion and exclusion criteria, 63 cases were excluded, yielding a final cohort of 825 patients, of whom 375 were diagnosed with MDRO infections. The study participant selection process is illustrated in [Figure S1](#), while [Table 1](#) presents the baseline characteristics of patients with MDRO and non-MDRO infections. The distribution of missing data is visualized in [Figure S2](#).

**Table 1** Demographic and Clinical Characteristics at Baseline

Variable	Total (n = 825)	Non MDRO Infection (n = 450)	MDRO Infection (n = 375)	P
Age, year, M (Q <sub>1</sub> , Q <sub>3</sub> )	65.00 (53.00–74.00)	62.00 (53.00–72.00)	68.00 (54.00–77.00)	<0.001 <sup>a</sup>
Gender				0.003 <sup>b</sup>
Female, n (%)	490 (59.39)	246 (54.67)	244 (65.07)	
Male, n (%)	335 (40.61)	204 (45.33)	131 (34.93)	
BMI, M (Q <sub>1</sub> , Q <sub>3</sub> )	24.03 (21.79–26.36)	24.08 (22.03–26.75)	23.80 (21.48–26.12)	0.150 <sup>a</sup>
Rural patients, n (%)	331 (40.12)	179 (39.11)	155 (41.33)	0.564 <sup>b</sup>
LOS, days, M (Q <sub>1</sub> , Q <sub>3</sub> )	36.06 (24.46–54.06)	35.66 (24.76–53.90)	36.97 (23.85–54.61)	0.771 <sup>a</sup>
SDH, M (Q <sub>1</sub> , Q <sub>3</sub> )	1.00 (0.00–1.00)	1.00 (0.00–1.00)	1.00 (0.00–1.00)	0.012 <sup>a</sup>
Number of hospitalizations, M (Q <sub>1</sub> , Q <sub>3</sub> )	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	0.858 <sup>a</sup>
Times for ventilator therapy, M (Q <sub>1</sub> , Q <sub>3</sub> )	1.00 (1.00–2.00)	1.00 (1.00–2.00)	2.00 (1.00–2.00)	<0.001 <sup>a</sup>
Days for ventilator therapy, M (Q <sub>1</sub> , Q <sub>3</sub> )	15.00 (5.00–27.00)	12.00 (4.00–24.00)	17.00 (6.00–29.00)	<0.001 <sup>a</sup>
Days of central venous catheterization, M (Q <sub>1</sub> , Q <sub>3</sub> )	14.00 (0.00–28.00)	14.00 (0.00–26.00)	15.00 (0.00–31.50)	0.201 <sup>a</sup>
Times of central venous catheterization, M (Q <sub>1</sub> , Q <sub>3</sub> )	2.00 (1.00–2.00)	1.00 (1.00–2.00)	2.00 (1.00–2.50)	0.003 <sup>a</sup>
Days of indwelling catheter, M (Q <sub>1</sub> , Q <sub>3</sub> )	29.00 (18.00–46.00)	26.00 (17.00–41.00)	34.00 (21.00–52.50)	<0.0001 <sup>a</sup>
Times of indwelling catheter, M (Q <sub>1</sub> , Q <sub>3</sub> )	2.00 (1.00–3.00)	2.00 (1.00–3.00)	2.00 (1.00–3.00)	0.022 <sup>a</sup>
Length of fever, M (Q <sub>1</sub> , Q <sub>3</sub> )	3.00 (0.00–9.00)	2.00 (0.00–9.00)	5.00 (0.00–10.00)	0.014 <sup>a</sup>
Operation history, n (%)	516 (62.55)	264 (58.67)	252 (67.20)	0.014 <sup>b</sup>
Scores of fall down risk, M (Q <sub>1</sub> , Q <sub>3</sub> )	12.00 (9.00–14.00)	12.00 (9.00–14.00)	12.00 (9.00–15.00)	0.167 <sup>a</sup>
Scores of pressure sores, M (Q <sub>1</sub> , Q <sub>3</sub> )	18.00 (14.00–23.00)	18.00 (15.00–23.00)	18.00 (14.00–23.00)	0.303 <sup>a</sup>
Scores of ADL, M (Q <sub>1</sub> , Q <sub>3</sub> )	25.00 (10.00–49.00)	25.00 (14.25–45.00)	20.00 (5.00–55.00)	0.268 <sup>a</sup>
Nutrition risk scores, M (Q <sub>1</sub> , Q <sub>3</sub> )	2.00 (1.00–2.00)	2.00 (1.00–2.00)	2.00 (1.00–2.00)	<0.001 <sup>a</sup>
Days of using antibiotics, M (Q <sub>1</sub> , Q <sub>3</sub> )	29.00 (20.00–44.00)	27.00 (19.00–40.00)	33.00 (23.00–49.00)	<0.001 <sup>a</sup>
Penicillin, n (%)	234 (28.36)	120 (26.67)	114 (30.40)	0.268 <sup>b</sup>
Cephalosporin, n (%)	803 (97.33)	439 (97.56)	364 (97.07)	0.828 <sup>b</sup>
Aminoglycosides, n (%)	254 (30.79)	118 (26.22)	136 (36.27)	0.002 <sup>b</sup>
Carbapenem, n (%)	373 (45.21)	197 (43.78)	176 (46.93)	0.403 <sup>b</sup>
Tetracyclines, n (%)	254 (30.79)	119 (26.44)	135 (36.00)	0.004 <sup>b</sup>
Macrolides, n (%)	21 (2.55)	4 (0.89)	17 (4.53)	0.002 <sup>b</sup>
Glycopeptides, n (%)	400 (48.48)	215 (47.78)	185 (49.33)	0.708 <sup>b</sup>
Quinolones, n (%)	377 (45.70)	188 (41.78)	189 (50.40)	0.016 <sup>b</sup>
Combined use of antibiotics, n (%)	817 (99.03)	449 (99.78)	368 (98.13)	0.026 <sup>c</sup>
Diabetes, n (%)	161 (19.52)	72 (16.00)	89 (23.73)	0.007 <sup>b</sup>
Hypertension, n (%)	321 (38.91)	165 (36.67)	156 (41.60)	0.169 <sup>b</sup>
Hypoproteinemia, n (%)	434 (52.61)	235 (52.22)	199 (53.07)	0.864 <sup>b</sup>
Respiratory failure, n (%)	219 (26.55)	107 (23.78)	112 (29.87)	0.058 <sup>b</sup>
Renal failure, n (%)	52 (6.3)	21 (4.67)	31 (8.27)	0.048 <sup>b</sup>
Liver insufficiency, n (%)	190 (23.03)	98 (21.78)	92 (24.53)	0.394 <sup>b</sup>
Hypokalemia, n (%)	211 (25.58)	118 (26.22)	93 (24.80)	0.699 <sup>b</sup>
Anemia, n (%)	282 (34.18)	141 (31.33)	141 (37.60)	0.069 <sup>b</sup>
AKI, n (%)	40 (4.85)	26 (5.78)	14 (3.73)	0.231 <sup>b</sup>
MODs, n (%)	50 (6.06)	17 (3.78)	33 (8.80)	0.004 <sup>b</sup>
Pneumonia, n (%)	431 (52.24)	223 (49.56)	208 (55.47)	0.105 <sup>b</sup>
Cerebral infarction, n (%)	245 (29.70)	108 (24.00)	137 (36.53)	<0.001 <sup>b</sup>
Cerebral hemorrhage, n (%)	134 (16.24)	57 (12.67)	77 (20.53)	0.003 <sup>b</sup>
Sepsis, n (%)	87 (10.55)	37 (8.22)	50 (13.33)	0.023 <sup>b</sup>
Electrolyte disturbance, n (%)	117 (14.18)	53 (11.78)	64 (17.07)	0.039 <sup>b</sup>
Acidosis, n (%)	110 (13.33)	46 (10.22)	64 (17.07)	0.005 <sup>b</sup>

Notes: <sup>a</sup>Mann–Whitney U-Test, <sup>b</sup>Chi-square test, <sup>c</sup>Fisher's Exact test.

Abbreviations: BMI, Body Mass Index; LOS, Length of Stay; SDH, Surgery during hospitalization; Scores of ADL, Scores of Activity of Daily Living; AKI, Acute Renal Injury; MODS, Multiple Organ Dysfunction Syndrome.



## Feature Selection

Within the first 24 hours of ICU admission, 45 clinical variables were recorded, with no variable exceeding a 30% missing data rate. Feature selection was conducted using LASSO regression, as illustrated in [Figure S3](#), identifying 31 key predictor variables detailed in [Table S2](#).

## Model Performance Comparison

Six machine learning models were developed to predict MDRO infections. [Table 2](#) summarizes their performance: accuracy: RF (0.767) DT (0.672), GPB (0.561), LR (0.637), NN (0.748), SVM (0.740); sensitivity: RF (0.747), DT (0.650), GPB (0.247), LR (0.615), NN (0.745), SVM (0.714); specificity: RF (0.788), DT (0.695), GPB (0.882), LR (0.662), NN (0.753), SVM (0.767). The ROC curve analysis ([Figure 2](#)) demonstrated that the RF model outperformed the others, achieving an AUC of 0.83 and an accuracy of 0.767 in the validation cohort. The calibration curve of the RF model ([Figure 3](#)) showed strong concordance between predicted and observed risks, confirming its reliability. Based on its superior performance, the RF model was selected for deployment. Detailed performance metrics for all models are provided in [Table 2](#), and the complete parameter configuration code is available in [Supplementary Materials File S1](#).

## Model Explanation

### Global Explanation

SHAP analysis was conducted to interpret the RF model at both global and local levels. [Figure 4A](#) displays the top 20 clinical features ranked by their average SHAP values, indicating their relative contributions to the model. Higher SHAP values denote greater predictive significance. Medical treatment factors, including prolonged ventilator use, extended urinary catheterization, and prolonged antimicrobial therapy, exhibited strong associations with increased MDRO infection risk. Additionally, demographic variables such as advanced age and abnormal body mass index (BMI) were linked to a higher likelihood of MDRO infection.

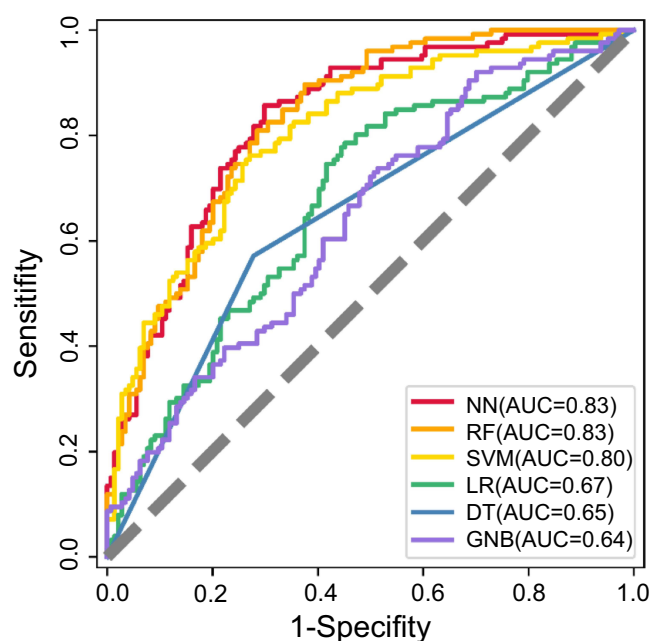
[Figure 4B](#) illustrates the directional impact of these features on the RF model. Each row represents a feature, while each point corresponds to an individual sample. The horizontal axis denotes the SHAP value, reflecting a feature's influence on model output. Red indicates a positive correlation, with darker shades representing higher feature values, whereas blue represents a negative correlation, with darker shades indicating lower feature values. These interpretations are derived from objective data, free from subjective bias.

### SHAP Interpretation of Significant Features

Further SHAP dependency plot analyses elucidated the relationship between clinical characteristics and MDRO infection risk. Variables such as the duration of urinary catheterization, age, length of antibiotic treatment, and ventilator use demonstrated a positive correlation with infection risk. [Figure 5](#) visualizes the SHAP interpretation of these key predictors, where the vertical axis represents the SHAP value and the horizontal axis denotes the feature's range of variation. Higher SHAP values indicate an increased risk of MDRO infection.

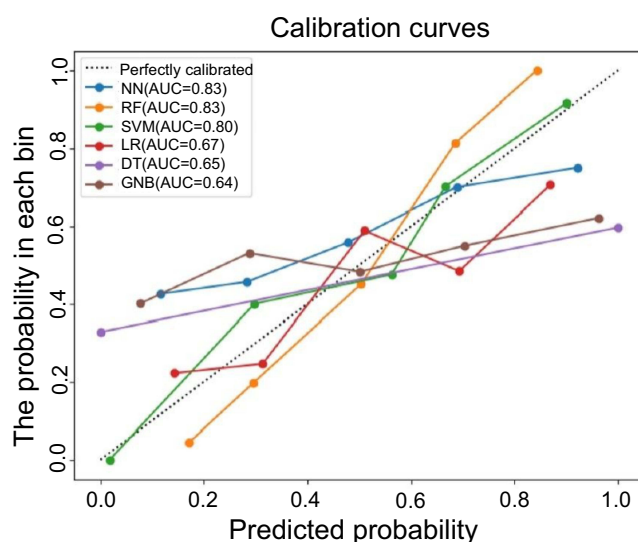
**Table 2** Performances of the Six Machine Learning Models for Predicting MDRO Infection

Model	Accuracy	Brier Scores	Error Rate	F1 Score	G-Means	Positive Predictive Value	Sensitivity	Specificity
Decision Tree	0.672	0.328	0.328	0.666	0.671	0.686	0.650	0.695
Gaussian Naive Bayes	0.561	0.439	0.439	0.340	0.438	0.689	0.247	0.882
Logistic Regression	0.637	0.363	0.363	0.630	0.637	0.650	0.615	0.662
Neural Network	0.748	0.252	0.252	0.748	0.747	0.756	0.745	0.753
Random Forest	0.767	0.233	0.233	0.763	0.767	0.783	0.747	0.788
Support Vector Machine	0.740	0.260	0.260	0.734	0.739	0.758	0.714	0.767



**Figure 2** ROC curves for the six ML models to predict MDRO infection.

**Abbreviations:** NN, Neural Network; RF, Random Forest; SVM, Support Vector Machine; LR, Logistic Regression; DT, Decision Tree; GNB, Gaussian Naive Bayes; AUC, area under the curve.



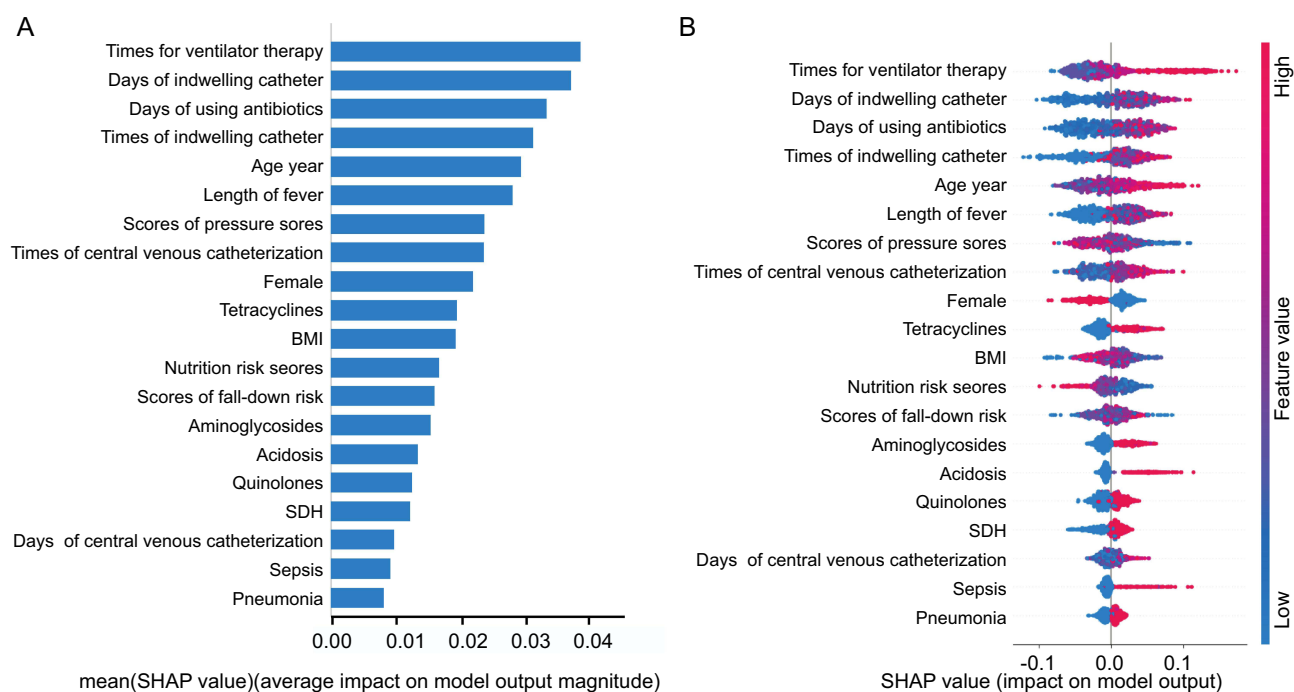
**Figure 3** Calibration plot of the MDRO infection prediction model.

**Abbreviations:** NN, Neural Network; RF, Random Forest; SVM, Support Vector Machine; LR, Logistic Regression; DT, Decision Tree; GNB, Gaussian Naive Bayes.

### Application of the Model

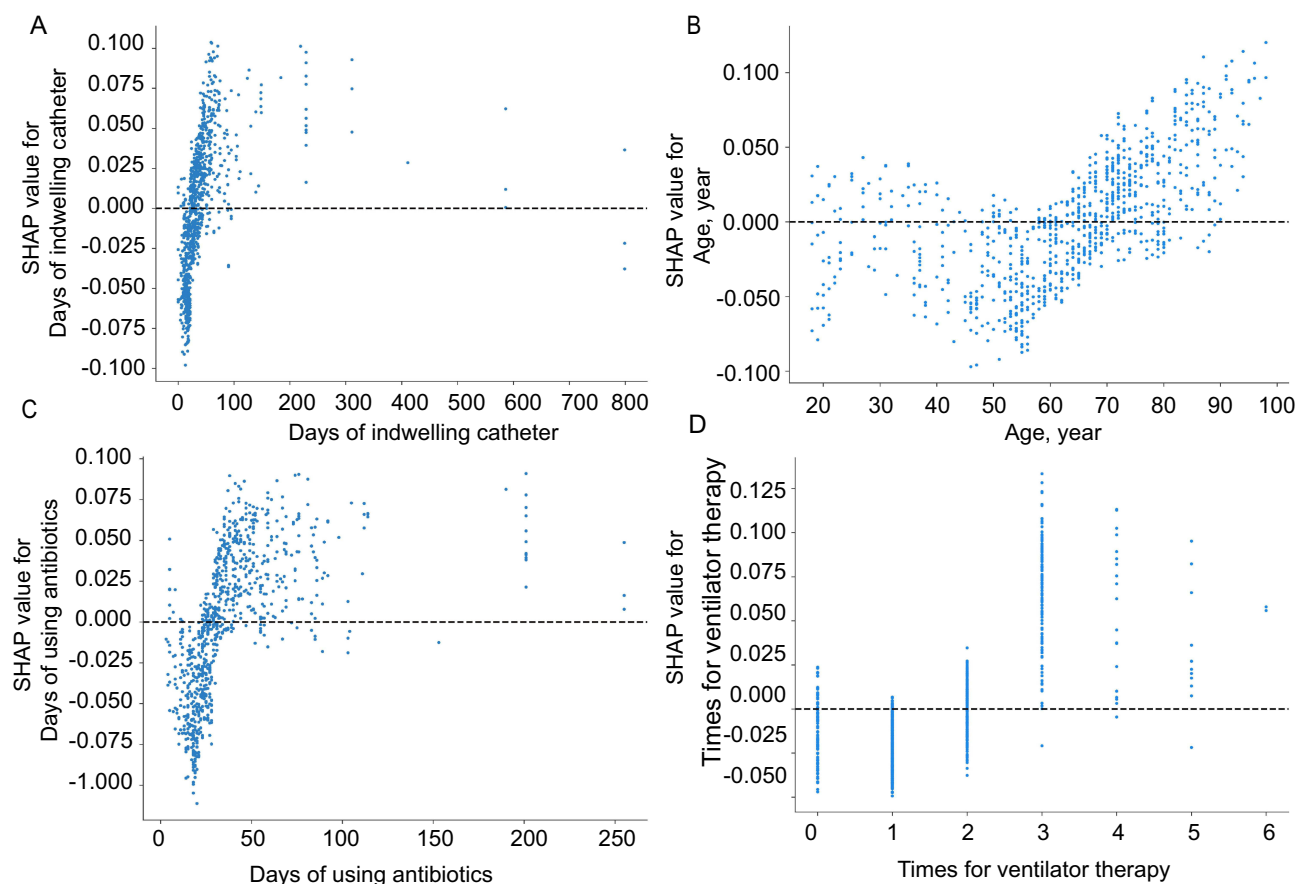
The SHAP force plot provides individualized risk assessments by visualizing how specific features influence model predictions. For instance, in patient 120, prolonged fever duration, catheterization, and antibiotic usage were negatively correlated with MDRO infection risk, whereas lower nutritional scores, prior tetracycline use, and reduced stress injury scores positively contributed to infection risk. The model assigned a 0.30 probability of MDRO infection for this patient, which fell below the threshold, leading to a prediction of no infection (Figure 6).



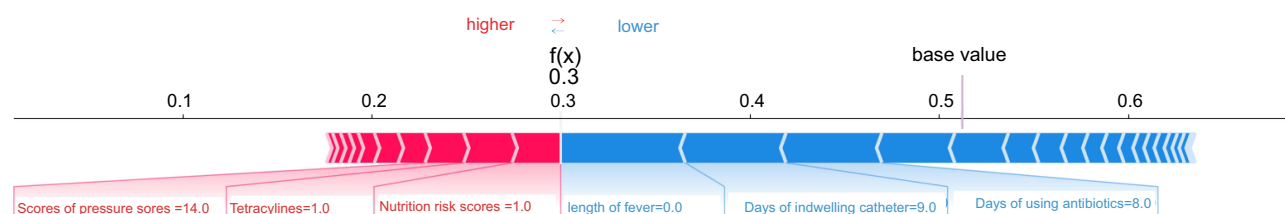


**Figure 4** SHAP summary plot of the top 20 clinical features contributing to the random forest model. **(A)** SHAP feature importance measured as the mean absolute Shapley values. **(B)** The attributes of the features in the model.

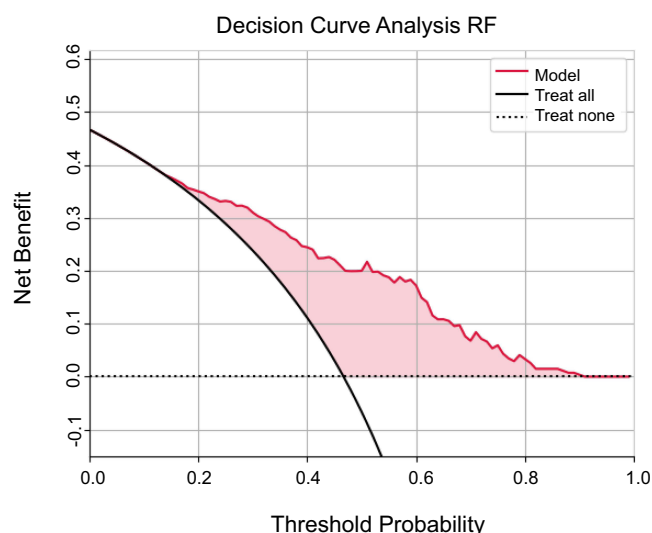
**Abbreviations:** SHAP, SHapley Additive exPlanation; BMI, Body Mass Index; SDH, Surgery During Hospitalization.



**Figure 5** SHAP Dependency Plots of Clinical Characteristics Contributing to Random Forest Models. **(A)** Day of indwelling catheter; **(B)** Age, year; **(C)** Day of using antibiotics; **(D)** Times for ventilator therapy.



**Figure 6** SHAP force plot for the interpretation of individual predictions in the validation cohort. Red color indicates a contributing effect on the outcome, blue color indicates a negative effect on the outcome, and a larger area indicates a higher degree of influence.



**Figure 7** Decision curve analysis for predicting MDRO infections for patients in intensive care unit. Solid line: prophylactic intervention for MDRO infection assumed to be performed in all patients. Dashed line: Assuming no prophylactic intervention for MDRO infection was given to any patient. Red line: assumes prophylactic intervention for patients at high risk of MDRO infection identified according to the Random Forest predictive model.

## Clinical Applications and Dissemination

Clinical DCA demonstrated that interventions targeting patients with a threshold probability between 0.44 and 0.60 provided greater net benefits compared to universal prophylaxis or no intervention (Figure 7 and Figure S4). These findings underscore the model's potential clinical utility and highlight its value in guiding evidence-based decision-making to improve patient outcomes.

## Discussion

### Summary

This study developed and evaluated six machine learning models for predicting MDRO infections in ICU patients using data from a tertiary hospital in Taiyuan, Shanxi Province. Among them, the RF model exhibited the highest predictive accuracy, providing a reliable tool for the early identification of high-risk patients. By prioritizing clinically accessible and relevant variables, the study improved model generalizability while effectively addressing challenges related to missing data.<sup>18</sup> Additionally, SHAP analysis was employed to enhance model interpretability, mitigating the “black box” limitations inherent in machine learning. However, clinical applicability must account for the threshold probability of adverse outcomes.<sup>35</sup> To assess the real-world utility of the RF model, DCA was performed, demonstrating its potential for guiding targeted interventions. Notably, this study focused exclusively on patients with documented infections, as MDRO colonization does not necessarily correlate with increased mortality.<sup>36</sup>

## Establishment and Evaluation of Prediction Models

Predicting MDRO infections is inherently complex due to their spatial clustering and clinical resemblance to non-MDRO infections. Prior research has underscored the efficacy of machine learning in this domain. For instance, Liang et al compared multiple models and found that Random Forest outperformed XGBoost, Decision Tree, and multivariable logistic regression (accuracy:  $84\% > 82\% > 81\% > 72\%$ ; AUROC:  $0.91 > 0.89 = 0.89 > 0.78$ ).<sup>37</sup> Despite the demonstrated advantages of machine learning-based approaches over traditional statistical methods, direct cross-study performance comparisons remain challenging due to differences in study settings, patient populations, and variable selection.<sup>38</sup>

## Interpretability of the Model

Model interpretability is a critical factor in healthcare applications, where predictive tools influence clinical decision-making and patient management.<sup>39</sup> The inherent opacity of complex machine learning algorithms often impedes their adoption, as healthcare professionals require transparent, explainable predictions to inform treatment decisions.<sup>39</sup> To address this limitation, the SHAP framework was integrated to provide granular insights into the RF model's decision-making process. By quantifying each feature's contribution to individual predictions, SHAP enhances model transparency and facilitates the identification of key clinical factors driving MDRO infection risk.<sup>30</sup>

The SHAP global explanation analysis identified invasive procedures and prolonged antimicrobial exposure as primary contributors to MDRO infections. Mechanical ventilation and indwelling urinary catheters emerged as significant risk factors, as these interventions are commonly administered to critically ill and immunocompromised patients requiring extended hospitalization and prolonged antibiotic therapy. Repeated or prolonged invasive procedures further facilitate MDRO colonization and transmission, a finding corroborated by Uluç et al.<sup>11</sup>

Beyond these procedural factors, antibiotic duration, advanced age, febrile days, and stress injury scores were pivotal in assessing MDRO infection risk, consistent with prior clinical studies.<sup>9,40</sup> SHAP dependency plots further delineated the linear associations between clinical variables and infection probability, demonstrating positive correlations with catheterization duration, age, antimicrobial exposure, and ventilatory support frequency. Atkinson et al reported a 1.4-fold increase in infection probability per decade of age, aligning with these findings.<sup>41</sup> Likewise, Mangioni et al<sup>42</sup> and Alsehem et al<sup>43</sup> established a robust link between extensive antimicrobial use and MDRO infections, corroborating the SHAP dependency analysis. Jiang et al<sup>32</sup> identified endotracheal intubation, arterial blood pressure monitoring, fever, antimicrobial administration, and pneumonia as independent risk factors. However, these variables do not necessarily imply direct avenues for preventive intervention.<sup>9,40</sup>

## Interpretability and Clinical Application of the RF Model

The interpretative analysis of the RF model using the SHAP Force Plot marks a advancement in individualized risk stratification for MDRO infections. This approach enables dynamic visualization of risk factors, deconstructing their directional impact (positive/negative) and magnitude, as quantified by absolute SHAP values.<sup>30</sup>

In the evaluation of case No.120, clinical parameters marked in blue (including fever duration, indwelling catheter days, and antibiotic usage days) demonstrated negative correlations with MDRO infection risk. Conversely, red-coded risk factors (nutrition risk scores, tetracycline-class antibiotics administration, and pressure ulcer severity scores) showed positive associations with infection probability. The model calculated a 0.03 predicted probability of infection for this case, which fell below the predetermined risk threshold. This individualized risk linkage analysis highlights both the universal and case-specific dimensions of clinical decision-making. To enhance clinical applicability, integrating the SHAP interpretation module into hospital information systems (HIS) and developing a real-time risk heatmap dashboard for continuous monitoring and intervention is recommended.<sup>44</sup>

## Antibiotic Use and MDRO Infection Risk

This study established a strong correlation between prolonged antibiotic exposure and an elevated risk of MDRO infections, corroborating the findings of Costelloe et al,<sup>45</sup> who reported a similar association between antibiotic use

and antimicrobial resistance. Prior research underscores that optimizing antibiotic stewardship can significantly reduce MDRO incidence.<sup>23</sup> Therefore, implementing structured antibiotic stewardship programs is imperative to curtail MDRO emergence and transmission in healthcare settings.<sup>46</sup> However, achieving complete infection prevention remains challenging due to resource constraints. Healthcare professionals must carefully balance intervention costs and benefits, leveraging DCA at appropriate thresholds to guide clinical decision-making.<sup>47</sup>

## Clinical Implementation and Future Directions

The proposed machine learning model offers substantial potential for clinical integration via hospital EHR systems. Real-time risk stratification could facilitate dynamic patient categorization and trigger automated alerts for infection control teams.<sup>18,25,37</sup> Integration with computerized physician order entry (CPOE) platforms would enable context-aware decision support, such as recommending spectrum narrowing in high-risk cases, while predictive analytics could optimize preemptive isolation strategies.<sup>24,42,48</sup> At an institutional level, aggregated risk assessments could inform targeted infection prevention policies through unit-specific risk heatmaps and antimicrobial benchmarking.<sup>41</sup> Effective deployment necessitates prospective validation across diverse healthcare settings, user-centered interface designs to mitigate alert fatigue, and robust ethical governance to ensure algorithmic accountability—ultimately bridging the gap between predictive modeling and actionable clinical interventions.<sup>31,39</sup>

## Limitations and Future Research

Several limitations should be acknowledged. First, spatial layout and healthcare worker activities were not included due to data constraints, potentially overlooking spatial clustering and procedural factors influencing infection risk. Second, the absence of pathogen-specific analyses limits the granularity of predictions, as bacterial species and resistance profiles were not incorporated. Third, the single-center design may constrain the generalizability of findings.

Future research should address these limitations by incorporating strain-specific classifications for MDRO infections, integrating spatial mapping of healthcare environments, and accounting for procedural workflows. Expanding the dataset to encompass multiple institutions and larger cohorts will enhance model robustness and external validity. Additionally, periodic model retraining every 2–3 years with updated datasets is essential to maintain predictive accuracy. Our research team remains committed to these iterative refinements to ensure the model's continued clinical relevance and efficacy.

## Conclusion

This study establishes the feasibility and clinical utility of machine learning in predicting MDRO infections. The Random Forest model demonstrated strong predictive performance, identifying key risk factors, with prolonged antibiotic use and invasive medical procedures emerging as the most influential modifiable determinants. SHAP analysis enhanced model interpretability, offering actionable insights to support clinical decision-making. These findings lay the groundwork for developing automated early warning systems to strengthen infection prevention and antimicrobial stewardship programs in healthcare settings. Future research should prioritize prospective validation and broader clinical integration.

## Data Sharing Statement

The datasets generated and analyzed in this study are not publicly available due to privacy and ethical constraints. However, raw data supporting the study's findings can be obtained from the corresponding author upon reasonable request. Requests must include a detailed research proposal outlining the intended data usage and comply with all relevant ethical and legal standards. Data sharing will be subject to institutional review board approval and adherence to data protection regulations.

## Compliance with Ethics Guidelines

This study was conducted in accordance with the principles of the Declaration of Helsinki and approved by the Ethics Committee of the First Hospital of Shanxi Medical University (Approval No.: 2022-K159). The study did not involve patient-identifiable information and was exempt from obtaining informed consent.

## Consent for Publication

All authors have consented to publication.

## Acknowledgments

The authors acknowledge the technical support provided by the Information Department of the First Hospital of Shanxi Medical University. This manuscript has been uploaded as a preprint on ResearchSquare: <https://www.researchsquare.com/article/rs-3409615/v1>.

## Author Contributions

Conceptualization: Wenting Zhao, Pei Sun; Methodology: Wei Li; Formal analysis and investigation: Pei Sun; Writing: Wenting Zhao, Pei Sun; Funding acquisition: Wenting Zhao; Supervision: Linping Shang. All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This research was supported by the Shanxi Province Graduate Education Innovation Project (Project No.: 2021Y371) and the Shanxi Provincial Health Commission Research Foundation Project (Grant No.: 2024212).

## Disclosure

The authors report no conflicts of interest in this work.

## References

- Magiorakos AP, Srinivasan A, Carey RB, et al. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin Microbiol Infect.* **2012**;18(3):268–281. doi:10.1111/j.1469-0691.2011.03570.x
- Murray CJL, Ikuta KS, Sharara F, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet.* **2022**;399(10325):629–655. doi:10.1016/S0140-6736(21)02724-0
- World Health Organization. Ten threats to global health in 2019. Available from: <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019>. Accessed April 21, 2025.
- Li ZJ, Wang KW, Liu B, et al. The distribution and source of MRDOs infection: a retrospective study in 8 ICUs, 2013–2019. *Infect Drug Resist.* **2021**;14:4983–4991. doi:10.2147/IDR.S332196
- Oliveira ABS, Sacilotto GH, Neves MFB, et al. Prevalence, outcomes, and predictors of multidrug-resistant nosocomial lower respiratory tract infections among patients in an ICU. *J Bras Pneumol.* **2023**;49(1):e20220235. doi:10.36416/1806-3756/e20220235
- Tsuzuki S, Yu J, Matsunaga N, Ohmagari N. Length of stay, hospitalisation costs and in-hospital mortality of methicillin-susceptible and methicillin-resistant *Staphylococcus aureus* bacteremia in Japan. *Public Health.* **2021**;198:292–296. doi:10.1016/j.puhe.2021.07.046
- Nelson RE, Hyun D, Jezek A, Samore MH. Mortality, length of stay, and healthcare costs associated with multidrug-resistant bacterial infections among elderly hospitalized patients in the United States. *Clin Infect Dis.* **2022**;74(6):1070–1080. doi:10.1093/cid/ciab696
- Abubakar U, Zulkarnain AI, Rodríguez-Baño J, et al. Treatments and predictors of mortality for carbapenem-resistant gram-negative bacilli infections in Malaysia: a retrospective cohort study. *Trop Med Infect Dis.* **2022**;7(12). doi:10.3390/tropicalmed7120415
- El Mekes A, Zahlane K, Ait Said L, Tadlaoui Ouafi A, Barakate M. The clinical and epidemiological risk factors of infections due to multi-drug resistant bacteria in an adult intensive care unit of University hospital center in Marrakesh-Morocco. *J Infect Public Health.* **2020**;13(4):637–643. doi:10.1016/j.jiph.2019.08.012
- Liu D, Li G, Hong Z, et al. Prevalence of multidrug-resistant organisms in healthy adults in Shenzhen, China. *Health Secur.* **2023**;21(2):122–129. doi:10.1089/hs.2022.0111
- Uluç K, Kutbay özçelik H, Akkütük Öngel E, et al. The prevalence of multidrug-resistant and extensively drug-resistant infections in respiratory intensive care unit, causative microorganisms and mortality. *Infect Drug Resist.* **2024**;17:4913–4919. doi:10.2147/IDR.S480829
- Chen G, Xu K, Sun F, et al. Risk factors of multidrug-resistant bacteria in lower respiratory tract infections: a systematic review and meta-analysis. *Can J Infect Dis Med Microbiol.* **2020**;2020:7268519. doi:10.1155/2020/7268519
- Tseng WP, Chen YC, Yang BJ, et al. Predicting multidrug-resistant gram-negative bacterial colonization and associated infection on hospital admission. *Infect Control Hosp Epidemiol.* **2017**;38(10):1216–1225. doi:10.1017/ice.2017.178
- Robben PM, Ayalew MD, Chung KK, Ressler RA. Multi-drug-resistant organisms in burn infections. *Surg Infect.* **2021**;22(1):103–112. doi:10.1089/sur.2020.129
- Al-Sunaidar KA, Aziz NA, Hassan Y, Jamshed S, Sekar M. Association of multidrug resistance bacteria and clinical outcomes of adult patients with sepsis in the intensive care unit. *Trop Med Infect Dis.* **2022**;7(11). doi:10.3390/tropicalmed7110365



16. Liu X, Liu X, Jin C, et al. Prediction models for diagnosis and prognosis of the colonization or infection of multidrug-resistant organisms in adults: a systematic review, critical appraisal, and meta-analysis. *Clin Microbiol Infect.* **2024**;30(11):1364–1373. doi:10.1016/j.cmi.2024.07.005
17. González Del Castillo J, Julián-Jiménez A, Gamazo-Del Rio JJ, et al. A multidrug-resistant microorganism infection risk prediction model: development and validation in an emergency medicine population. *Eur J Clin Microbiol Infect Dis.* **2020**;39(2):309–323. doi:10.1007/s10096-019-03727-4
18. Seo SM, Jeong IS, Song JY, Lee S. Development of a nomogram for carbapenem-resistant Enterobacteriaceae acquisition risk prediction among patients in the intensive care unit of a secondary referral hospital. *Asian Nurs Res.* **2021**;15(3):174–180. doi:10.1016/j.anr.2021.02.005
19. Song JY, Jeong IS. Development of a risk prediction model of carbapenem-resistant Enterobacteriaceae colonization among patients in intensive care units. *Am J Infect Control.* **2018**;46(11):1240–1244. doi:10.1016/j.ajic.2018.05.001
20. Bi Q, Goodman KE, Kaminsky J, Lessler J. What is machine learning? A primer for the epidemiologist. *Am J Epidemiol.* **2019**;188(12):2222–2239. doi:10.1093/aje/kwz189
21. Dantas LF, Peres IT, Antunes BBP, et al. Prediction of multidrug-resistant bacteria (MDR) hospital-acquired infection (HAI) and colonisation: a systematic review. *Infect Dis Health.* **2025**;30(1):50–60. doi:10.1016/j.idh.2024.07.003
22. Heo J, Yoon JG, Park H, et al. Machine learning-based model for prediction of outcomes in acute stroke. *Stroke.* **2019**;50(5):1263–1265. doi:10.1161/STROKEAHA.118.024293
23. Chen Y, Chen X, Liang Z, et al. Epidemiology and prediction of multidrug-resistant bacteria based on hospital level. *J Glob Antimicrob Resist.* **2022**;29:155–162. doi:10.1016/j.jgar.2022.03.003
24. Goodman KE, Lessler J, Cosgrove SE, et al. A clinical decision tree to predict whether a bacteremic patient is infected with an extended-spectrum  $\beta$ -lactamase-producing organism. *Clin Infect Dis.* **2016**;63(7):896–903. doi:10.1093/cid/ciw425
25. Hartvigsen T, Sen C, Rundensteiner EA. Detecting MRSA infections by fusing structured and unstructured electronic health record data. In: *Biomedical Engineering Systems and Technologies*. Cham: Springer International Publishing; **2019**:399–419.
26. Li Y, Cao Y, Wang M, et al. Development and validation of machine learning models to predict MDRO colonization or infection on ICU admission by using electronic health record data. *Antimicrob Resist Infect Control.* **2024**;13(1):74. doi:10.1186/s13756-024-01428-y
27. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell.* **2019**;1(5):206–215. doi:10.1038/s42256-019-0048-x
28. Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. *Nature Mach Intell.* **2020**;2(1):56–67. doi:10.1038/s42256-019-0138-9
29. Fahmy AS, Csecs I, Arafati A, et al. An explainable machine learning approach reveals prognostic significance of right ventricular dysfunction in nonischemic cardiomyopathy. *JACC Cardiovasc Imaging.* **2022**;15(5):766–779. doi:10.1016/j.jcmg.2021.11.029
30. Hu C, Li L, Huang W, et al. Interpretable machine learning for early prediction of prognosis in sepsis: a discovery and validation study. *Infect Dis Ther.* **2022**;11(3):1117–1132. doi:10.1007/s40121-022-00628-6
31. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med.* **2015**;162(1):55–63. doi:10.7326/M14-0697
32. Li J, Li Y, Song N, Chen Y. Risk factors for carbapenem-resistant *Klebsiella pneumoniae* infection: a meta-analysis. *J Glob Antimicrob Resist.* **2020**;21:306–313. doi:10.1016/j.jgar.2019.09.006
33. Raman G, Avendano EE, Chan J, Merchant S, Puzniak L. Risk factors for hospitalized patients with resistant or multidrug-resistant *Pseudomonas aeruginosa* infections: a systematic review and meta-analysis. *Antimicrob Resist Infect Control.* **2018**;7(1):79. doi:10.1186/s13756-018-0370-9
34. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ.* **2020**;368:m441. doi:10.1136/bmj.m441
35. Yue S, Li S, Huang X, et al. Construction and validation of a risk prediction model for acute kidney injury in patients suffering from septic shock. *Dis Markers.* **2022**;2022:9367873. doi:10.1155/2022/9367873
36. Junior APN, Del Missier GM, Praça APA, Silva I, Caruso P. In-hospital mortality and one-year survival of critically ill patients with cancer colonized or not with carbapenem-resistant gram-negative bacteria or vancomycin-resistant enterococci: an observational study. *Antimicrob Resist Infect Control.* **2023**;12(1):8. doi:10.1186/s13756-023-01214-2
37. Liang Q, Zhao Q, Xu X, Zhou Y, Huang M. Early prediction of carbapenem-resistant Gram-negative bacterial carriage in intensive care units using machine learning. *J Glob Antimicrob Resist.* **2022**;29:225–231. doi:10.1016/j.jgar.2022.03.019
38. Bozorgmehr A, Thielmann A, Weltermann B. Chronic stress in practice assistants: an analytic approach comparing four machine learning classifiers with a standard logistic regression model. *PLoS One.* **2021**;16(5):e0250842. doi:10.1371/journal.pone.0250842
39. Petch J, Di S, Nelson W. Opening the black box: the promise and limitations of explainable machine learning in cardiology. *Can J Cardiol.* **2022**;38(2):204–213. doi:10.1016/j.cjca.2021.09.004
40. Herrera S, Torralbo B, Herranz S, et al. Carriage of multidrug-resistant Gram-negative bacilli: duration and risk factors. *Eur J Clin Microbiol Infect Dis.* **2023**;42(5):631–638. doi:10.1007/s10096-023-04581-1
41. Atkinson A, Ellenberger B, Piezzi V, et al. Extending outbreak investigation with machine learning and graph theory: benefits of new tools with application to a nosocomial outbreak of a multidrug-resistant organism. *Infect Control Hosp Epidemiol.* **2023**;44(2):246–252. doi:10.1017/ice.2022.66
42. Mangioni D, Chatenoud L, Colombo J, et al. Multidrug-resistant bacterial colonization and infections in large retrospective cohort of mechanically ventilated COVID-19 patients(1). *Emerg Infect Dis.* **2023**;29(8):1598–1607. doi:10.3201/eid2908.230115
43. Alseheri AF, Alharbi EA, Alammash BB, et al. Assessment of risk factors associated with multidrug-resistant organism infections among patients admitted in a tertiary hospital - a retrospective study. *Saudi Pharm J.* **2023**;31(6):1084–1093. doi:10.1016/j.jsps.2023.03.019
44. Tan TH, Hsu CC, Chen CJ, et al. Predicting outcomes in older ED patients with influenza in real time using a big data-driven and machine learning approach to the hospital information system. *BMC Geriatr.* **2021**;21(1):280. doi:10.1186/s12877-021-02229-3
45. Costelloe C, Metcalfe C, Lovering A, Mant D, Hay AD. Effect of antibiotic prescribing in primary care on antimicrobial resistance in individual patients: systematic review and meta-analysis. *BMJ.* **2010**;340(may18 2):c2096. doi:10.1136/bmj.c2096
46. Pouwels KB, Butler CC, Robotham JV. Comment on ‘The distribution of antibiotic use and its association with antibiotic resistance’. *Elife.* **2019**;8. doi:10.7554/eLife.46561



47. Jayaraman SP, Jiang Y, Resch S, Askari R, Klompas M. Cost-effectiveness of a model infection control program for preventing multi-drug-resistant organism infections in critically ill surgical patients. *Surg Infect.* 2016;17(5):589–595. doi:10.1089/sur.2015.222
48. Hur EY, Jin YJ, Jin TX, Lee SM. Development and evaluation of the automated risk assessment system for multidrug-resistant organisms (autoRAS-MDRO). *J Hosp Infect.* 2018;98(2):202–211. doi:10.1016/j.jhin.2017.08.004

## Infection and Drug Resistance

### Publish your work in this journal

Infection and Drug Resistance is an international, peer-reviewed open-access journal that focuses on the optimal treatment of infection (bacterial, fungal and viral) and the development and institution of preventive strategies to minimize the development and spread of resistance. The journal is specifically concerned with the epidemiology of antibiotic resistance and the mechanisms of resistance development and diffusion in both hospitals and the community. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/infection-and-drug-resistance-journal>

**Dovepress**  
Taylor & Francis Group