Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Machine Learning–Based Identification of Drug-Resistant Pathogen Infection in Patients With Sepsis: A Multicenter Cohort Study

Qiushi Feng

nihaofengqiushi@163.com

The Second Affiliated Hospital of Shenyang Medical College

Hui Miao

Shenyang City University

Wentao Li

Xiayi County Hospital of Traditional Chinese Medicine

Yan Wu

The Second Affiliated Hospital of Shenyang Medical College

Research Article

Keywords: sepsis, multidrug-resistant organisms, MRSA, VRE, carbapenem-resistant Pseudomonas aeruginosa, machine learning, explainable AI, SHAP, decision curve analysis

Posted Date: July 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10188056/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Machine Learning–Based Identification of Drug-Resistant Pathogen Infection in Patients With Sepsis: A Multicenter Cohort Study

Qiushi Feng[1]*†, Hui Miao[2]†, Wentao Li[3], Yan Wu[1]

[1] The Second Affiliated Hospital of Shenyang Medical College, Shenyang, China

[2] Shenyang City University, Shenyang, China

[3] Xiayi County Hospital of Traditional Chinese Medicine, Xiayi, Henan, China

* Corresponding author: Qiushi Feng, Email: nihaofengqiushi@163.com; ORCID iD: https://orcid.org/0009-0001-3639-4326

† These authors contributed equally to this work (co-first author).

ORCID iD of the submitting/corresponding author: Qiushi Feng, 0009-0001-3639-4326

ORCID iD of the submitting/corresponding author: Hui Miao, 0009-0009-3510-8621

Address: The Second Affiliated Hospital of Shenyang Medical College, Shenyang, China

Abstract

Background: Multidrug-resistant organisms (MDROs) such as methicillin-resistant *Staphylococcus aureus* (MRSA), vancomycin-resistant *Enterococcus* (VRE), and carbapenem-resistant *Pseudomonas aeruginosa* (CRPA) are major drivers of adverse outcomes in intensive care unit (ICU) patients with sepsis. At ICU admission, clinicians often must choose empiric antibiotics before culture and susceptibility results are available; inaccurate early risk assessment can delay active therapy or lead to unnecessary broad-spectrum antibiotic exposure.

Methods: We conducted a retrospective multicenter study using two publicly available ICU databases (eICU v2.0 and MIMIC-IV v3.1). Adult ICU admissions (age ≥18 years) with ICU length of stay ≥24 hours were included, and admissions with known MDRO infection or positive MDRO culture before ICU admission were excluded. The outcome window was restricted to clinically ordered cultures obtained from ICU admission through ICU day 7. We extracted 168 candidate predictors from the first 24 hours of ICU data; post-index variables, including ICU or hospital length of stay, were excluded from model development. The merged cohort (N=45,141) was randomly split into a development set (70%, n=31,599) and an independent validation set (30%, n=13,542). Pathogen-specific models were evaluated by AUROC, decision curve analysis (0–50% threshold probabilities), and SHapley Additive exPlanations (SHAP).

Results: Within the prespecified early ICU outcome window, the prevalence of MRSA, VRE, and CRPA infections was 1.5% (n=683), 0.7% (n=296), and 0.2% (n=90), respectively. In the validation set, AUROCs were 0.82 for MRSA, 0.85 for VRE, and 0.82 for CRPA. Decision curve analysis demonstrated positive net benefit across clinically relevant thresholds for all three models. SHAP analyses highlighted pathogen-specific model signals: eosinophil count, PaO2, and platelet count for MRSA; fibrinogen, lymphocyte count, and PaCO2 for VRE; and potassium, LDL cholesterol, and lymphocyte count for CRPA.

Conclusions: Explainable machine-learning models based on routinely available early ICU variables may support early MDRO risk stratification in sepsis and provide transparent,

patient-level explanations. External and prospective validation will be required before clinical implementation.

Keywords: sepsis; multidrug-resistant organisms; MRSA; VRE; carbapenem-resistant Pseudomonas aeruginosa; machine learning; explainable AI; SHAP; decision curve analysis

Introduction

Sepsis is a time-critical syndrome in which clinicians must frequently begin empiric antimicrobial therapy before microbiology results are available. This decision is especially difficult when multidrug-resistant organisms (MDROs) are possible, because undertreatment may delay active therapy whereas overtreatment may increase unnecessary broad-spectrum exposure and antimicrobial resistance pressure [1,10–12]. Methicillin-resistant *Staphylococcus aureus* (MRSA), vancomycin-resistant *Enterococcus* (VRE), and carbapenem-resistant *Pseudomonas aeruginosa* (CRPA) are clinically important resistant phenotypes in ICU patients, but early bedside recognition is challenging because clinical signs are non-specific and culture-based confirmation is delayed.

Electronic health record (EHR) data from modern ICUs provide an opportunity to support earlier risk stratification using variables that are available during the first day of care. Machine-learning (ML) methods have been applied to sepsis detection, deterioration prediction, and antimicrobial resistance prediction [2–9,22–24]. However, many published models are limited by unclear outcome timing, potential leakage from post-index variables, incomplete microbiology definitions, limited interpretability, and insufficient assessment of clinical utility.

Therefore, we developed and validated pathogen-specific ML models to identify early ICU patients with sepsis who were at risk for culture-positive MRSA, VRE, or CRPA infection. We used two large public ICU databases, restricted predictors to the first 24 hours after ICU admission, excluded length-of-stay variables from the candidate predictor set, defined outcomes within an early ICU culture window, and assessed model performance using discrimination, decision curve analysis, and SHAP-based explanations.

Methods

Study design and data sources

We conducted a retrospective multicenter cohort study using a harmonized critical care dataset formed by merging sepsis cohorts from the eICU Collaborative Research Database v2.0 and MIMIC-IV v3.1 [19,20]. In MIMIC-IV, sepsis was identified using a Sepsis-3 framework based on suspected infection and organ dysfunction (SOFA increase ≥2) when derived sepsis indicators were available. In eICU, sepsis was identified using structured diagnosis fields, APACHE admission diagnoses, and ICD diagnosis codes indicating sepsis or septic shock, supplemented by infection-related culture and antimicrobial records when available. These database-specific definitions were used to construct comparable ICU sepsis cohorts while acknowledging differences in source data structure.

Eligibility criteria were age ≥18 years and ICU length of stay ≥24 hours. ICU admission was defined as the index time. Admissions with known MDRO infection or positive MDRO culture before ICU admission were excluded. The primary outcome window was ICU day 0 through ICU day 7; cultures obtained after day 7 were not used to define the primary outcome. The final analytic cohort comprised 45,141 ICU admissions. Predictors were extracted only from the first 24 hours of ICU stay to reflect an early, clinically actionable prediction time point.

Ethics and reporting considerations

Both eICU and MIMIC-IV are publicly available, de-identified critical care databases that have been described in peer-reviewed data descriptor publications [19,20]. This retrospective secondary analysis was conducted in accordance with the principles of the Declaration of Helsinki. The collection of patient information and creation of the MIMIC-IV research resource were reviewed and approved by the Institutional Review Board of Beth Israel Deaconess Medical Center, which granted a waiver of informed consent for the de-identified data-sharing initiative. The eICU Collaborative Research Database was de-identified in accordance with the Health Insurance Portability and Accountability Act (HIPAA) Safe Harbor provision, with re-identification risk certified by Privacert (Cambridge, MA; HIPAA Certification no. 1031219-2). The present study used only credentialed-access, de-identified public data under the PhysioNet data-use agreements;

therefore, additional informed consent was not required. We followed established reporting principles for prediction modeling studies, emphasizing transparent separation of development and validation data, clear definition of outcomes and predictors, and complete reporting of discrimination and clinical-utility measures.

Outcomes (drug-resistant pathogens)

The primary outcomes were binary indicators of culture-positive resistant-pathogen infection: MRSA, VRE, and CRPA. Eligible microbiology records included clinically ordered blood, respiratory, urine, wound/tissue, and other body-fluid cultures obtained from ICU admission through ICU day 7. Resistance was determined from organism identification and phenotypic susceptibility results recorded in the microbiology tables; rapid diagnostic platforms were not consistently distinguishable across databases and therefore were not used as a separate source of resistance classification. To reduce misclassification of colonization as infection, surveillance or screening cultures were excluded when identifiable, and non-sterile-site isolates were considered outcomes only when they were clinically ordered and temporally consistent with treatment for suspected infection. Residual misclassification between colonization and infection remains possible and is acknowledged as a limitation.

Candidate predictors and preprocessing

Candidate predictors were selected a priori from domains routinely available in ICU electronic health records and previously linked to infection severity, antimicrobial resistance risk, physiologic instability, or sepsis outcomes: demographics, comorbidities, early organ-support therapies, vital signs, laboratory tests, and severity scores [10–12,15,22]. All candidate predictors had to be available within the first 24 hours after ICU admission. Post-index variables, including ICU length of stay, hospital length of stay, and variables requiring future information, were excluded before feature selection and model training. For variables measured repeatedly, the first value and clinically meaningful summaries (minimum, maximum, mean, or worst value, depending on the variable type)

were derived within the 24-hour window. Implausible outliers were screened using physiologic plausibility checks and winsorized at prespecified extreme percentiles when necessary.

Development and validation strategy

To support machine-learning development and internal validation, we performed a random split of the merged cohort into a development (training) set and an independent validation set at a 70:30 ratio. The 70% subset was used for feature selection, preprocessing, model training, and hyperparameter tuning; the 30% subset was held out for performance evaluation. All preprocessing parameters, including imputation and scaling, were estimated within the training set and then applied unchanged to the validation set to avoid information leakage.

Baseline comparisons and statistical testing. Baseline characteristics were summarized for the overall cohort and stratified by infection status for each resistant pathogen (MRSA, VRE, CRPA). Continuous variables are presented as median with interquartile range (IQR) due to non-normal distributions typical of ICU biomarkers, and categorical variables as counts with percentages. Between-group comparisons used Wilcoxon rank-sum tests for continuous variables and chi-square tests or Fisher's exact tests for categorical variables, as appropriate, with two-sided P values reported in the baseline tables (Table 1A–1C). These unadjusted comparisons were intended to contextualize clinical differences rather than to infer causality. We further summarized standardized univariable odds ratios to support signal screening (Supplementary Table S1).

Feature selection (Boruta) and interpretability (SHAP)

For each resistant pathogen outcome, we applied Boruta feature selection on the training set after excluding post-index length-of-stay variables. Boruta ranks features based on repeated comparisons to permuted shadow features, retaining variables that consistently outperform noise. To enhance interpretability of the final models, we used SHapley

Additive exPlanations (SHAP) to quantify feature contributions to individual predictions and to summarize global importance using mean absolute SHAP values. Boruta feature importance is visualized in Figures 1–3; the ranked Boruta feature table was moved to the Supplementary Material to reduce redundancy with the figures, while SHAP summaries are presented in Figures 4–6 and Table 4.

Machine-learning models and performance evaluation

We trained multiple classifiers for each resistant pathogen outcome, including penalized logistic regression (LASSO), AdaBoost, LightGBM, XGBoost, and CatBoost, using a consistent training pipeline. Hyperparameters were tuned within the training set using resampling and objective functions aligned with discrimination performance. Model discrimination was assessed on the independent validation set using receiver operating characteristic (ROC) curves and area under the ROC curve (AUC). Additional performance metrics (sensitivity, specificity, and positive predictive value) were summarized to reflect clinical screening considerations, particularly given outcome imbalance. Clinical utility was evaluated with decision curve analysis (DCA) across threshold probabilities from 0% to 50%, comparing each model's net benefit with treat-all and treat-none strategies.

Handling missing data and class imbalance. Missingness patterns differed across laboratories and between source databases. We quantified the proportion of missing values for each candidate predictor and added a variable-level missingness and reporting summary to the Supplementary Material (Supplementary Table S7). For modeling, imputation was performed within the training set to avoid information leakage: median imputation for continuous variables and mode imputation for categorical variables, with imputation parameters fixed and applied to the validation set. Continuous predictors were scaled as required. Given the low prevalence of resistant infections, we evaluated imbalance-aware options such as class-weighting and threshold optimization within the training set and reported performance on the held-out validation set to reflect real-world prevalence.

A total of 168 candidate predictors were extracted across demographics, vital signs, laboratory tests, comorbidity burden (Elixhauser), severity scores (e.g., SOFA), and early organ-support variables. Length-of-stay variables and other post-index variables were excluded from the final modeling pipeline.

In cross-validated LASSO tuning on the development set, the optimal penalty parameter (lambda.min) was 0.000935 for MRSA, 5.69×10^{-6} for VRE, and 0.000305 for CRPA (Supplementary Table S5).

Model training details. For LASSO, we fit penalized logistic regression models using cross-validated selection of the penalty parameter (lambda), considering both lambda.min and lambda.1se for parsimony (Supplementary Table S5). For tree-based ensembles (AdaBoost, LightGBM, XGBoost, CatBoost), hyperparameters were tuned using resampling within the training set. Fixed random seeds were used for data splitting and resampling, and identical preprocessing steps were applied across all classifiers and outcomes. All final models were refit after excluding ICU/hospital length-of-stay variables.

Evaluation metrics and reporting. Discrimination was summarized by AUC with 95% confidence intervals estimated on the validation set, and we additionally reported sensitivity, specificity, and positive predictive value (PPV) at prespecified decision thresholds (Supplementary Table S3). Clinical utility was assessed with DCA. Net benefit was defined as the proportion of true-positive classifications minus the proportion of false-positive classifications weighted by the odds of the selected threshold probability; thus, positive net benefit indicates that model-guided action provides more benefit than harm compared with treat-all or treat-none strategies. Threshold probabilities from 0% to 50% were evaluated because these ranges are most relevant for targeted diagnostics, infection-control precautions, or empiric coverage decisions in low-prevalence MDRO outcomes.

Results

Cohort description and dataset split

The merged sepsis cohort included 45,141 ICU admissions/patients across the two databases. After random allocation, 31,599 admissions (70.0%) were assigned to the

training set and 13,542 admissions (30.0%) to the validation set. Baseline characteristics of the overall cohort and by split are shown in Table 1 and Table 2, respectively. Overall, the training and validation sets were comparable across demographic characteristics, comorbidity profiles, early severity indicators, and commonly measured laboratory and physiologic variables, supporting their use for model development and internal validation.

Prevalence of resistant pathogens

Resistant pathogen infections were uncommon in this sepsis cohort when restricted to the prespecified early ICU outcome window. MRSA infection was identified in 683 admissions (1.5%), VRE in 296 admissions (0.7%), and CRPA in 90 admissions (0.2%). The low prevalence should be interpreted in light of the requirement for clinically ordered cultures and documented susceptibility results; not every ICU admission necessarily had all culture types collected.

Baseline differences by pathogen

In unadjusted comparisons, MRSA-positive, VRE-positive, and CRPA-positive patients differed from patients without the corresponding resistant organism across several demographic, laboratory, and physiologic variables (Table 1A–1C). These comparisons were descriptive and were not interpreted as causal evidence. They were used to contextualize the cohort and to motivate prediction modeling based only on early ICU variables.

Top features identified by Boruta

Boruta feature selection highlighted both shared and pathogen-specific predictors after exclusion of length-of-stay variables (Figures 1–3; Supplementary Table S1). For MRSA, top-ranked features included white blood cell count, age, platelet count, glucose, hemoglobin, and oxygenation measures (e.g., PaO2). For VRE, white blood cell count, platelet count, fibrinogen, PCO2, and lymphocyte-related variables ranked highly. For CRPA, PCO2, fibrinogen, white blood cell count, lymphocyte measures, LDL, platelet

count, and oxygenation gradients were among the leading predictors. We report these rankings descriptively and avoid causal interpretation.

Cross-outcome overlap and pathogen specificity. Across all three outcomes, several variable domains repeatedly appeared among top-ranked features, including inflammatory markers, hematologic measures, and respiratory physiology. However, the relative ordering differed by pathogen. Because these are model-derived associations, they should be interpreted as predictive signals rather than mechanistic explanations.

Model performance across classifiers

On the validation set, all model families demonstrated clinically meaningful discrimination, with AUCs ranging from 0.75 to 0.85 depending on outcome (Figures 7–9; Supplementary Table S4). For MRSA, the highest AUC was achieved by LASSO (AUC=0.83), closely followed by LightGBM, XGBoost, and CatBoost (each AUC=0.82–0.81), whereas AdaBoost showed slightly lower performance (AUC=0.80). For VRE, AdaBoost and CatBoost achieved the highest AUCs (both AUC=0.85), outperforming LASSO and LightGBM (AUC=0.82) and XGBoost (AUC=0.80). For CRPA, LightGBM provided the best discrimination (AUC=0.82), followed by XGBoost (AUC=0.81) and LASSO (AUC=0.80), while AdaBoost (AUC=0.77) and CatBoost (AUC=0.75) were less discriminative.

Performance metrics beyond AUC are summarized in Supplementary Table S3. For MRSA, the LASSO model achieved sensitivity of 0.75 and specificity of 0.75 with AUC 0.83 (95% CI 0.77–0.87) and PPV 0.70. For VRE, CatBoost demonstrated strong overall performance with AUC 0.85 (95% CI 0.81–0.89), specificity 0.84, sensitivity 0.74, and PPV 0.77; AdaBoost showed comparable AUC 0.85 (95% CI 0.82–0.89) with a higher sensitivity (0.76) but lower specificity (0.77). For CRPA, LightGBM had the best discrimination (AUC 0.82, 95% CI 0.77–0.86) and favorable operating characteristics (specificity 0.78; sensitivity 0.71; PPV 0.71).

Taken together, these results suggest that different model families may be preferred depending on the pathogen and the intended clinical trade-off. Penalized regression

(LASSO) performed competitively for MRSA, whereas gradient-boosted and categorical-boosting methods were advantageous for VRE and CRPA. Notably, performance differences were modest for MRSA and CRPA, indicating that model choice could also be guided by implementation considerations such as computational cost, interpretability, and calibration.

Model interpretability (SHAP)

SHAP analyses provided model-level explanations of the final length-of-stay-excluded models (Figures 4–6; Table 3). For MRSA, eosinophils, PaO2, platelet count, white blood cell count, hemoglobin, glucose, age, and SOFA emerged as leading contributors. For VRE, fibrinogen, lymphocytes, PCO2, white blood cell count, and platelet count were most influential. For CRPA, potassium, LDL, lymphocytes, ALT, oxygenation gradient (A–a gradient), basophils, age, phosphate, cholesterol, and pH contributed prominently.

Directionality patterns in SHAP plots are best interpreted as model behavior rather than causal effects. For example, abnormal inflammatory or oxygenation-related values often shifted predicted risk upward, whereas more favorable physiologic values generally shifted predictions toward lower risk. The SHAP distributions also illustrated substantial inter-individual heterogeneity, supporting the use of patient-level explanations rather than a single universal rule.

Risk factor signals and clinical utility

Decision curve analysis suggested that the final models could yield positive net benefit across clinically relevant threshold probabilities (0–50%), outperforming treat-all and treat-none strategies (Figures 10–12; Supplementary Table S6). For MRSA, LASSO provided the greatest net benefit over approximately 10–35% thresholds, with curves converging at higher thresholds. For VRE, CatBoost showed consistently higher net benefit across most thresholds. For CRPA, LightGBM tended to dominate at moderate-to-high thresholds (~20–50%). Together, ROC and DCA results support the feasibility of using

early ICU variables to guide targeted diagnostic or infection-control strategies, pending external validation.

Clinical interpretation of DCA results. In the DCA framework, net benefit expresses the clinical value of a model after accounting for the trade-off between true-positive and false-positive classifications at a chosen threshold probability. A model is clinically useful when its net benefit exceeds both the treat-none strategy (no targeted action) and the treat-all strategy (universal targeted action) within threshold ranges where clinicians would reasonably intervene. In this study, positive net benefit over wide threshold ranges suggests potential utility for targeted culture acquisition, rapid molecular testing, early infection-control precautions, or empiric antimicrobial escalation in selected high-risk patients.

Discussion

The main finding of this study is that pathogen-specific, explainable ML models using only early ICU variables achieved moderate-to-good discrimination for MRSA, VRE, and CRPA risk in patients with sepsis. This addresses a practical clinical problem: at ICU admission, clinicians must make empiric antimicrobial decisions before culture and susceptibility results are available. A model that identifies patients at higher risk for resistant pathogens could support earlier targeted diagnostics, infection-control precautions, and stewardship-informed empiric therapy, while avoiding universal broad-spectrum escalation.

Our findings align with prior work showing that EHR-based ML can support sepsis and antimicrobial-resistance risk prediction [2–5,22–24], but extend this literature by focusing on three explicit resistant phenotypes and by evaluating net benefit using DCA. The model signals identified by Boruta and SHAP were descriptive rather than causal.

After removing length-of-stay variables, the remaining predictors largely reflected early inflammation, hematology, organ dysfunction, respiratory physiology, and metabolic status. These variables are routinely available in ICU data and are therefore more plausible for early implementation than variables that require future information.

The low prevalence of MRSA, VRE, and CRPA in this cohort is an important consideration. The outcomes required clinically ordered cultures and documented susceptibility results, and not every patient underwent all culture types. Therefore, the models should be interpreted as predicting documented early culture-positive resistant infection rather than the true biologic prevalence of MDRO infection. In addition, although we attempted to exclude surveillance cultures and align non-sterile-site isolates with clinical treatment, residual misclassification between colonization and infection is possible.

Several limitations should be emphasized. First, the outcome window was restricted to ICU day 0 through day 7 to match the early clinical question, but resistant infections occurring later in the ICU stay were not counted as primary outcomes. Second, prior antibiotic exposure is a known risk factor for resistance, but harmonized pre-index antibiotic exposure was not consistently available across the two databases and was therefore not included in the main models. Third, missingness varied by predictor and database, requiring imputation and careful interpretation of laboratory-derived predictors. Fourth, although the study used two large public ICU databases and a held-out validation set, external temporal and prospective validation will be required before clinical

deployment.

Implementation should be cautious and workflow-centered. These models are not intended to replace clinician judgment or microbiology testing. Rather, they may provide an early, transparent risk-stratification layer that can help prioritize rapid diagnostics, infection-control interventions, or antimicrobial stewardship review. Future work should validate the models across hospitals and time periods, incorporate prior antibiotic exposure when reliably available, and test whether model-guided workflows improve time to active therapy without increasing unnecessary broad-spectrum antibiotic use.

In summary, early ICU EHR data can be used to construct interpretable, pathogen-specific prediction models for resistant infections in sepsis. The clinical value of such models will depend on external validation, calibration, workflow integration, and prospective assessment of whether model-guided actions improve both patient-level outcomes and antimicrobial stewardship.

Conclusion

Explainable machine-learning models built from early ICU EHR variables can identify sepsis patients at increased risk for documented early MRSA, VRE, and CRPA infection and provide transparent, patient-level explanations of model behavior. These models may support targeted diagnostics and infection-control decisions, but external and prospective validation is required before clinical implementation.

Declarations

Ethics approval and consent to participate

This study analyzed publicly available, de-identified data from eICU (v2.0) and MIMIC-IV (v3.1) and was conducted in accordance with the principles of the Declaration of Helsinki. For MIMIC-IV, the collection of patient information and creation of the research resource were reviewed and approved by the Institutional Review Board of Beth Israel Deaconess Medical Center, which granted a waiver of informed consent for the de-identified data-sharing initiative. For eICU, the database was de-identified in accordance with the HIPAA Safe Harbor provision, and the re-identification risk was certified by Privacert (Cambridge, MA; HIPAA Certification no. 1031219-2). The present secondary analysis was performed by credentialed users under the applicable PhysioNet data-use agreements; because only de-identified public data were used, additional individual informed consent was not required.

Data Availability Statement

The eICU Collaborative Research Database and MIMIC-IV are available through PhysioNet to credentialed users under applicable data use agreements. The research team has completed the required human subjects/data-use training (CITI Program: Human Research—Data or Specimens Only Research, Basic Course; Record ID: 74861575) and accessed the data in compliance with PhysioNet requirements.

Competing interests

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Didelot X, Pouwels KB. Machine-learning-assisted selection of antibiotic prescription. Nat Med. 2019;25(7):1033-1034. doi:10.1038/s41591-019-0517-0. https://doi.org/10.1038/s41591-019-0517-0

2. Delahanty RJ, Alvarez J, Flynn LM, Sherwin RL, Jones SS. Development and Evaluation of a Machine Learning Model for the Early Identification of Patients at Risk for Sepsis. Ann Emerg Med. 2019;73(4):334-344. doi:10.1016/j.annemergmed.2018.11.036. https://doi.org/10.1016/j.annemergmed.2018.11.036

3. Kong G, Lin K, Hu Y. Using machine learning methods to predict in-hospital mortality of sepsis patients in the ICU. BMC Med Inform Decis Mak. 2020;20:251. doi:10.1186/s12911-020-01271-2. https://doi.org/10.1186/s12911-020-01271-2

4. Islam MM, Nasrin T, Walther BA, et al. Prediction of sepsis patients using machine learning approach: a meta-analysis. Comput Methods Programs Biomed. 2019;170:1-9. doi:10.1016/j.cmpb.2018.12.027. https://doi.org/10.1016/j.cmpb.2018.12.027

5. Shimabukuro DW, Barton CW, Feldman MD, Mataraso SJ, Das R. Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial. BMJ Open Respir Res. 2017;4:e000234. doi:10.1136/bmjresp-2017-000234. https://doi.org/10.1136/bmjresp-2017-000234

6. Tang K, Tang D, Wang Q, et al. MALDI-TOF MS platform combined with machine learning to establish a model for rapid identification of methicillin-resistant Staphylococcus aureus. J Microbiol Methods. 2021;180:106109. doi:10.1016/j.mimet.2020.106109. https://doi.org/10.1016/j.mimet.2020.106109

7. Goodswen SJ, Barratt JLN, Kennedy PJ, Kaufer A, Calarco L, Ellis JT. Machine learning and applications in microbiology. FEMS Microbiol Rev. 2021;45(5):fuab015. doi:10.1093/femsre/fuab015. https://doi.org/10.1093/femsre/fuab015

8. Giacobbe DR, Mora S, Giacomini M, Bassetti M. Machine Learning and Multidrug-Resistant Gram-Negative Bacteria: An Interesting Combination for Current and Future Research. Antibiotics (Basel). 2020;9(2):54. doi:10.3390/antibiotics9020054. https://doi.org/10.3390/antibiotics9020054

9. Serafim MSM, Kronenberger T, Oliveira PR, Poso A, Honorio KM, Maltarollo VG. The application of machine learning techniques to innovative antibacterial discovery and development. Expert Opin Drug Discov. 2020;15(10):1165-1179. doi:10.1080/17460441.2020.1776696. https://doi.org/10.1080/17460441.2020.1776696

10. Laxminarayan R, Duse A, Wattal C, et al. Antibiotic resistance—the need for global solutions. Lancet Infect Dis. 2013;13(12):1057-1098. doi:10.1016/S1473-3099(13)70318-9. https://doi.org/10.1016/S1473-3099(13)70318-9

11. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016;315(8):801-810. doi:10.1001/jama.2016.0287. https://doi.org/10.1001/jama.2016.0287

12. Evans L, Rhodes A, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. Intensive Care Med. 2021;47(11):1181-1247. doi:10.1007/s00134-021-06506-y. https://doi.org/10.1007/s00134-021-06506-y

13. Mekontso-Dessap A, Kirsch M, Brun-Buisson C, et al. Poststernotomy mediastinitis due to Staphylococcus aureus: comparison of methicillin-resistant and methicillin-susceptible cases. Clin Infect Dis. 2001;32(6):877-883.

14. Kim SH, Park WB, Lee KD, et al. Outcome of Staphylococcus aureus bacteremia in patients with eradicable foci versus noneradicable foci. Clin Infect Dis. 2003;37(6):794-799. doi:10.1086/377540. https://doi.org/10.1086/377540

15. Jee Y, Carlson J, Rafai E, et al. Antimicrobial resistance: a threat to global health. Lancet Infect Dis. 2018;18(9):939-940. doi:10.1016/S1473-3099(18)30471-7. https://doi.org/10.1016/S1473-3099(18)30471-7

16. Dantas G, Sommer MOA. How to Fight Back Against Antibiotic Resistance. Am Sci. 2014;102(1):42-51. Available from: https://www.americanscientist.org/article/how-to-fight-back-against-antibiotic-resistance

17. Centers for Disease Control and Prevention. Newly reported gene, mcr-1, threatens last-resort antibiotics. Atlanta (GA): CDC; 2016. Available from: https://www.cdc.gov/media/releases/2016/s0531-mcr1.html

18. Luepke KH, Suda KJ, Boucher H, et al. Past, Present, and Future of Antibacterial Economics: Increasing Bacterial Resistance, Limited Antibiotic Pipeline, and Societal Implications. Pharmacotherapy. 2017;37(1):71-84. doi:10.1002/phar.1868. https://doi.org/10.1002/phar.1868

19. Pollard TJ, Johnson AEW, Raffa JD, et al. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. Sci Data. 2018;5:180178. doi:10.1038/sdata.2018.178. https://doi.org/10.1038/sdata.2018.178

20. Johnson AEW, Bulgarelli L, Shen L, et al. MIMIC-IV, a freely accessible electronic health record dataset. Sci Data. 2023;10(1):1. doi:10.1038/s41597-022-01899-x. https://doi.org/10.1038/s41597-022-01899-x

21. Hyland SL, Faltys M, Hüser M, et al. Early prediction of circulatory failure in the intensive care unit using machine learning. Nat Med. 2020;26(3):364-373. doi:10.1038/s41591-020-0789-4. https://doi.org/10.1038/s41591-020-0789-4

22. Lewin-Epstein O, Baruch S, Hadany L, et al. Predicting antibiotic resistance in hospitalized patients by applying machine learning to electronic medical records. Clin Infect Dis. 2021;72(11):e848-e855. doi:10.1093/cid/ciaa1576. https://doi.org/10.1093/cid/ciaa1576

23. Fleuren LM, Klausch TLT, Zwager CL, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. Intensive Care Med. 2020;46(3):383-400. doi:10.1007/s00134-019-05872-y. https://doi.org/10.1007/s00134-019-05872-y

24. Moor M, Brückner M, Bialonski S, et al. Toward optimal sepsis prediction in intensive care units using machine learning: a systematic review. Front Med (Lausanne). 2021;8:607952. doi:10.3389/fmed.2021.607952. https://doi.org/10.3389/fmed.2021.607952

25. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26(6):565-574. doi:10.1177/0272989X06295361. https://doi.org/10.1177/0272989X06295361

26. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. BMJ. 2016;352:i6. doi:10.1136/bmj.i6. https://doi.org/10.1136/bmj.i6

27. Hastie T, Tibshirani R, Friedman JH. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. New York: Springer; 2009.

28. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 2017;30:4765-4774. Available from: https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

29. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med. 2019;170(1):51-58. doi:10.7326/M18-1376. https://doi.org/10.7326/M18-1376

30. Collins GS, Reitsma JB, Altman DG, Moons KGM. 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. https://doi.org/10.7326/M14-0697

31. World Health Organization. Global action plan on antimicrobial resistance. Geneva: World Health Organization; 2015. Available from: https://iris.who.int/handle/10665/193736

Figure legends

Figure 1. Boruta feature importance for MRSA infection prediction after excluding length-of-stay and other post-index variables. Boxplots show feature importance scores across Boruta iterations, with shadow features used as the reference for acceptance.

Figure 2. Boruta feature importance for VRE infection prediction after excluding length-of-stay and other post-index variables. Feature importance distributions are compared with shadow feature maxima to identify robust predictors.

Figure 3. Boruta feature importance for CRPA infection prediction after excluding length-of-stay and other post-index variables. Importance distributions across iterations highlight top-ranked predictors relative to shadow controls.

Figure 4. SHAP summary plot for MRSA infection prediction in the final model excluding length-of-stay variables. Each point represents an individual admission and shows the SHAP value for a given feature; positive values increase predicted risk and negative values decrease risk.

Figure 5. SHAP summary plot for VRE infection prediction in the final model excluding length-of-stay variables. Global feature importance is summarized by the distribution of SHAP values across admissions.

Figure 6. SHAP summary plot for CRPA infection prediction in the final model excluding length-of-stay variables. Features are ordered by overall contribution, illustrating heterogeneity of individual-level model contributions.

Figure 7. Receiver operating characteristic (ROC) curves for MRSA infection prediction in the validation set using final models based only on early ICU variables. Curves compare

LASSO, AdaBoost, LightGBM, XGBoost, and CatBoost, with AUC values shown in the legend.

Figure 8. Receiver operating characteristic (ROC) curves for VRE infection prediction in the validation set using final models based only on early ICU variables. Model discrimination is summarized by AUC values in the legend.

Figure 9. Receiver operating characteristic (ROC) curves for CRPA infection prediction in the validation set using final models based only on early ICU variables. ROC curves and AUC values illustrate discrimination differences across classifiers.

Figure 10. Decision curve analysis (DCA) for MRSA infection prediction using final models based only on early ICU variables. Net benefit is plotted across threshold probabilities (0–50%) and compared with treat-all and treat-none strategies.

Figure 11. Decision curve analysis (DCA) for VRE infection prediction using final models based only on early ICU variables. Net benefit curves indicate potential clinical utility across threshold probabilities relative to treat-all and treat-none strategies.

Figure 12. Decision curve analysis (DCA) for CRPA infection prediction using final models based only on early ICU variables. Net benefit across threshold probabilities (0–50%) compares classifiers and benchmark strategies.

Table 1A. Baseline characteristics of the merged eICU and MIMIC-IV sepsis cohort stratified by MRSA status (Yes vs No). Rows indicating absence of binary conditions were removed for brevity. Outcome ascertainment used clinically ordered blood, respiratory, urine, wound/tissue, and other body-fluid cultures.

Variable	Overall	MRSA No	MRSA Yes	P value
N	45141	44458	683	
Database: eICU	30380 (67.3%)	30335 (68.2%)	45 (6.6%)	
Database: MIMIC-IV	14761 (32.7%)	14123 (31.8%)	638 (93.4%)	
Sex: Male	24495 (54.3%)	24114 (54.2%)	381 (55.8%)	0.866

Sex: Female	20640 (45.7%)	20338 (45.7%)	302 (44.2%)	
Sex: Other/Unknown	2 (0.0%)	2 (0.0%)	0 (0.0%)	
Sex: Unknown	4 (0.0%)	4 (0.0%)	0 (0.0%)	
Race/ethnicity: White	9672 (21.4%)	9258 (20.8%)	414 (60.6%)	<0.001
Race/ethnicity: Black	1078 (2.4%)	1031 (2.3%)	47 (6.9%)	
Race/ethnicity: Hispanic	456 (1.0%)	441 (1.0%)	15 (2.2%)	
Race/ethnicity: Asian	391 (0.9%)	381 (0.9%)	10 (1.5%)	
Race/ethnicity: Other	1026 (2.3%)	977 (2.2%)	49 (7.2%)	
Race/ethnicity: Unknown	32518 (72.0%)	32370 (72.8%)	148 (21.7%)	
Age, years	67.00 (55.00, 77.00)	67.00 (55.00, 77.00)	72.00 (59.00, 82.00)	<0.001
Sodium, mmol/L	143.00 (140.00, 146.00)	143.00 (140.00, 146.00)	144.00 (140.00, 147.00)	<0.001
Serum creatinine	1.50 (0.98, 2.72)	1.50 (0.98, 2.72)	1.40 (0.90, 2.50)	0.044
White blood cell count	16.80 (12.00, 23.20)	16.80 (12.00, 23.20)	16.30 (11.43, 22.20)	0.120
Potassium, mmol/L	4.70 (4.30, 5.20)	4.70 (4.30, 5.20)	4.80 (4.40, 5.40)	<0.001
Congestive heart failure (Yes)	7466 (16.5%)	7279 (16.4%)	187 (27.4%)	<0.001
Diabetes (Yes)	4406 (9.8%)	4169 (9.4%)	237 (34.7%)	<0.001
Chloride, mmol/L	105.00 (2.70, 112.00)	105.00 (2.70, 112.00)	2.50 (2.20, 2.80)	<0.001
Platelet count	204.00 (152.00, 308.00)	205.00 (152.00, 310.00)	165.00 (149.00, 190.00)	<0.001
Mechanical ventilation (Yes)	7231 (16.0%)	7212 (16.2%)	19 (2.8%)	<0.001
Vasopressor use (Yes)	4814 (10.7%)	4574 (10.3%)	240 (35.1%)	<0.001
INR	1.25 (1.10, 1.53)	1.26 (1.10, 1.55)	1.14 (1.01, 1.29)	<0.001
Total bilirubin	0.80 (0.50, 1.60)	0.80 (0.50, 1.60)	0.70 (0.40, 1.50)	0.004
Lactate, mmol/L	2.50 (1.60, 4.20)	2.50 (1.60, 4.20)	2.40 (1.60, 3.70)	0.245
GCS score	9.00 (3.80, 15.00)	9.00 (3.84, 15.00)	3.24 (2.43, 4.18)	<0.001
Arterial pH	7.45 (7.39, 7.49)	7.44 (7.39, 7.49)	7.46 (7.41, 7.50)	<0.001

Anion gap	16.00 (12.00, 25.00)	15.80 (12.00, 24.80)	137.00 (133.00, 140.00)	<0.001
Blood urea nitrogen	37.00 (23.00, 60.00)	37.00 (23.00, 60.00)	39.00 (26.00, 51.00)	0.937
Chronic kidney disease (Yes)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1.000
Liver disease/cirrhosis (Yes)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1.000
Malignancy (Yes)	104 (0.2%)	104 (0.2%)	0 (0.0%)	1.000
Bicarbonate, mmol/L	22.10 (6.10, 27.90)	22.30 (6.30, 28.00)	4.60 (3.90, 5.60)	<0.001
Magnesium	2.20 (2.00, 2.50)	2.20 (2.00, 2.50)	2.30 (2.20, 2.60)	0.008
PaO2, mmHg	140.00 (92.80, 228.00)	140.00 (92.50, 228.00)	172.50 (136.00, 226.00)	0.040
FiO2	70.00 (40.00, 100.00)	70.00 (40.00, 100.00)	1.00 (0.60, 100.00)	<0.001
APTT	36.00 (30.00, 48.50)	36.00 (30.00, 48.50)	38.00 (29.60, 54.00)	0.650
ICU mortality (Yes)	2552 (5.7%)	2404 (5.4%)	148 (21.7%)	<0.001
Hypertension (Yes)	7173 (15.9%)	6882 (15.5%)	291 (42.6%)	0.129
ALT	4.80 (4.40, 5.30)	4.80 (4.40, 5.30)	4.80 (4.40, 5.50)	0.348
AST	143.00 (140.00, 146.00)	143.00 (140.00, 146.00)	144.00 (140.00, 148.00)	<0.001
Glucose	122.82 (107.66, 149.07)	122.60 (107.62, 148.44)	129.00 (109.14, 159.22)	<0.001
Hemoglobin, g/dL	35.20 (31.70, 39.20)	35.20 (31.70, 39.20)	34.70 (31.40, 38.70)	0.106
Phosphate	4.40 (3.70, 5.50)	4.40 (3.70, 5.50)	4.50 (3.80, 5.70)	0.098
SOFA score	4.57 (3.60, 5.31)	4.58 (3.62, 5.32)	4.14 (3.25, 5.13)	<0.001
Lymphocyte count	13.40 (8.30, 20.40)	13.50 (8.40, 20.40)	11.80 (7.20, 18.60)	<0.001
Albumin, g/dL	3.20 (2.80, 3.70)	3.30 (2.80, 3.70)	3.10 (2.70, 3.50)	<0.001
Body mass index, kg/m^2	29.10 (25.10, 34.70)	29.10 (25.10, 34.70)	28.80 (25.40, 34.90)	0.714
Calcium	138.00 (135.00, 140.00)	138.00 (135.00, 140.00)	138.00 (135.00, 141.00)	0.393
Temperature, °C	37.17 (36.62, 37.64)	37.17 (36.62, 37.64)	37.26 (36.49, 37.65)	0.608
C-reactive protein	138.50 (94.00, 229.00)	139.00 (94.00, 231.00)	126.50 (90.00, 192.25)	0.184

Table 1B. Baseline characteristics of the merged eICU and MIMIC-IV sepsis cohort stratified by VRE status (Yes vs No). Rows indicating absence of binary conditions were removed for brevity. Outcome ascertainment used clinically ordered blood, respiratory, urine, wound/tissue, and other body-fluid cultures.

Variable	Overall	VRE No	VRE Yes	P value
N	45141	44845	296	
Database: eICU	30380 (67.3%)	30372 (67.7%)	8 (2.7%)	
Database: MIMIC-IV	14761 (32.7%)	14473 (32.3%)	288 (97.3%)	
Sex: Male	24495 (54.3%)	24339 (54.3%)	156 (52.7%)	0.953
Sex: Female	20640 (45.7%)	20500 (45.7%)	140 (47.3%)	
Sex: Other/Unknown	2 (0.0%)	2 (0.0%)	0 (0.0%)	
Sex: Unknown	4 (0.0%)	4 (0.0%)	0 (0.0%)	
Race/ethnicity: White	9672 (21.4%)	9499 (21.2%)	173 (58.4%)	<0.001
Race/ethnicity: Black	1078 (2.4%)	1058 (2.4%)	20 (6.8%)	
Race/ethnicity: Hispanic	456 (1.0%)	443 (1.0%)	13 (4.4%)	
Race/ethnicity: Asian	391 (0.9%)	383 (0.9%)	8 (2.7%)	
Race/ethnicity: Other	1026 (2.3%)	999 (2.2%)	27 (9.1%)	

Race/ethnicity: Unknown	32518 (72.0%)	32463 (72.4%)	55 (18.6%)	
Age, years	67.00 (55.00, 77.00)	67.00 (55.00, 77.00)	65.00 (54.00, 75.00)	0.069
Sodium, mmol/L	143.00 (140.00, 146.00)	143.00 (140.00, 146.00)	144.00 (141.00, 148.00)	<0.001
Serum creatinine	1.50 (0.98, 2.72)	1.50 (0.97, 2.71)	1.90 (1.05, 4.15)	<0.001
White blood cell count	16.80 (12.00, 23.20)	16.80 (12.00, 23.20)	19.40 (12.45, 28.25)	<0.001
Potassium, mmol/L	4.70 (4.30, 5.20)	4.70 (4.30, 5.20)	5.20 (4.70, 5.80)	<0.001
Congestive heart failure (Yes)	7466 (16.5%)	7425 (16.6%)	41 (13.9%)	0.217
Diabetes (Yes)	4406 (9.8%)	4319 (9.6%)	87 (29.4%)	<0.001
Chloride, mmol/L	105.00 (2.70, 112.00)	105.00 (2.70, 112.00)	2.60 (2.40, 2.90)	<0.001
Platelet count	204.00 (152.00, 308.00)	204.00 (152.00, 309.00)	185.00 (166.75, 216.25)	<0.001
Mechanical ventilation (Yes)	7231 (16.0%)	7228 (16.1%)	3 (1.0%)	<0.001
Vasopressor use (Yes)	4814 (10.7%)	4660 (10.4%)	154 (52.0%)	<0.001
INR	1.25 (1.10, 1.53)	1.26 (1.10, 1.54)	1.09 (0.97, 1.23)	<0.001
Total bilirubin	0.80 (0.50, 1.60)	0.80 (0.50, 1.60)	1.60 (0.60, 7.22)	<0.001
Lactate, mmol/L	2.50 (1.60, 4.20)	2.50 (1.60, 4.20)	3.10 (1.92, 5.67)	<0.001
GCS score	9.00 (3.80, 15.00)	9.00 (3.81, 15.00)	3.14 (2.41, 4.08)	<0.001
Arterial pH	7.45 (7.39, 7.49)	7.44 (7.39, 7.49)	7.47 (7.42, 7.52)	<0.001
Anion gap	16.00 (12.00, 25.00)	15.90 (12.00, 25.00)	138.00 (134.00, 141.00)	<0.001
Blood urea nitrogen	37.00 (23.00, 60.00)	37.00 (23.00, 60.00)	69.50 (63.00, 96.00)	<0.001
Chronic kidney disease (Yes)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1.000
Liver disease/cirrhosis (Yes)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1.000
Malignancy (Yes)	104 (0.2%)	104 (0.2%)	0 (0.0%)	1.000
Bicarbonate, mmol/L	22.10 (6.10, 27.90)	22.20 (6.20, 27.90)	4.50 (3.90, 5.30)	<0.001
Magnesium	2.20 (2.00, 2.50)	2.20 (2.00, 2.50)	2.50 (2.50, 2.90)	0.004
PaO2, mmHg	140.00 (92.80, 228.00)	140.00 (92.78, 228.00)	214.00 (214.00, 312.00)	0.048

FiO2	70.00 (40.00, 100.00)	70.00 (40.00, 100.00)	70.00 (70.00, 100.00)	0.886
APTT	36.00 (30.00, 48.50)	36.00 (30.00, 48.40)	65.70 (59.17, 67.05)	0.080
ICU mortality (Yes)	2552 (5.7%)	2471 (5.5%)	81 (27.4%)	<0.001
Hypertension (Yes)	7173 (15.9%)	7037 (15.7%)	136 (45.9%)	0.672
ALT	4.80 (4.40, 5.30)	4.80 (4.40, 5.30)	5.10 (4.65, 5.85)	<0.001
AST	143.00 (140.00, 146.00)	143.00 (140.00, 146.00)	144.00 (141.00, 148.00)	<0.001
Glucose	122.82 (107.66, 149.07)	122.82 (107.62, 149.01)	123.17 (109.26, 151.15)	0.643
Hemoglobin, g/dL	35.20 (31.70, 39.20)	35.20 (31.70, 39.30)	33.80 (30.20, 37.45)	<0.001
Phosphate	4.40 (3.70, 5.50)	4.40 (3.70, 5.50)	5.50 (4.40, 7.15)	<0.001
SOFA score	4.57 (3.60, 5.31)	4.57 (3.61, 5.31)	4.16 (3.32, 5.15)	<0.001
Lymphocyte count	13.40 (8.30, 20.40)	13.40 (8.30, 20.30)	14.70 (8.80, 24.35)	0.035
Albumin, g/dL	3.20 (2.80, 3.70)	3.30 (2.80, 3.70)	3.10 (2.60, 3.80)	0.055
Body mass index, kg/m^2	29.10 (25.10, 34.70)	29.10 (25.10, 34.70)	28.00 (23.80, 31.70)	0.003
Calcium	138.00 (135.00, 140.00)	138.00 (135.00, 140.00)	138.00 (134.00, 141.00)	0.789
Temperature, °C	37.17 (36.62, 37.64)	37.17 (36.62, 37.64)	38.94 (37.18, 42.74)	0.021
C-reactive protein	138.50 (94.00, 229.00)	138.00 (93.00, 229.00)	160.00 (114.00, 230.00)	0.047

Table 1C. Baseline characteristics of the merged eICU and MIMIC-IV sepsis cohort stratified by CRPA status (Yes vs No). Rows indicating absence of binary conditions were removed for brevity. Outcome ascertainment used clinically ordered blood, respiratory, urine, wound/tissue, and other body-fluid cultures.

Variable	Overall	CRPA No	CRPA Yes	P value
N	45141	45051	90	
Database: eICU	30380 (67.3%)	30376 (67.4%)	4 (4.4%)	
Database: MIMIC-IV	14761 (32.7%)	14675 (32.6%)	86 (95.6%)	
Sex: Male	24495 (54.3%)	24436 (54.2%)	59 (65.6%)	0.200
Sex: Female	20640 (45.7%)	20609 (45.7%)	31 (34.4%)	
Sex: Other/Unknown	2 (0.0%)	2 (0.0%)	0 (0.0%)	
Sex: Unknown	4 (0.0%)	4 (0.0%)	0 (0.0%)	
Race/ethnicity: White	9672 (21.4%)	9613 (21.3%)	59 (65.6%)	<0.001
Race/ethnicity: Black	1078 (2.4%)	1073 (2.4%)	5 (5.6%)	
Race/ethnicity: Hispanic	456 (1.0%)	454 (1.0%)	2 (2.2%)	
Race/ethnicity: Asian	391 (0.9%)	388 (0.9%)	3 (3.3%)	
Race/ethnicity: Other	1026 (2.3%)	1021 (2.3%)	5 (5.6%)	
Race/ethnicity: Unknown	32518 (72.0%)	32502 (72.1%)	16 (17.8%)	
Age, years	67.00 (55.00, 77.00)	67.00 (55.00, 77.00)	65.00 (52.00, 75.75)	0.291
Sodium, mmol/L	143.00 (140.00, 146.00)	143.00 (140.00, 146.00)	147.00 (140.00, 149.00)	<0.001
Serum creatinine	1.50 (0.98, 2.72)	1.50 (0.98, 2.72)	1.50 (1.00, 2.80)	0.885
White blood cell count	16.80 (12.00, 23.20)	16.80 (12.00, 23.20)	19.80 (11.80, 29.10)	0.064
Potassium, mmol/L	4.70 (4.30, 5.20)	4.70 (4.30, 5.20)	5.20 (4.60, 5.90)	<0.001
Congestive heart failure (Yes)	7466 (16.5%)	7449 (16.5%)	17 (18.9%)	0.675
Diabetes (Yes)	4406 (9.8%)	4374 (9.7%)	32 (35.6%)	<0.001
Chloride, mmol/L	105.00 (2.70, 112.00)	105.00 (2.70, 112.00)	2.50 (2.30, 2.90)	<0.001

Platelet count	204.00 (152.00, 308.00)	204.00 (152.00, 308.00)	182.00 (165.00, 204.00)	0.005
Mechanical ventilation (Yes)	7231 (16.0%)	7229 (16.0%)	2 (2.2%)	<0.001
Vasopressor use (Yes)	4814 (10.7%)	4762 (10.6%)	52 (57.8%)	<0.001
INR	1.25 (1.10, 1.53)	1.25 (1.10, 1.53)	1.10 (0.98, 1.23)	<0.001
Total bilirubin	0.80 (0.50, 1.60)	0.80 (0.50, 1.60)	0.80 (0.40, 2.08)	0.682
Lactate, mmol/L	2.50 (1.60, 4.20)	2.50 (1.60, 4.20)	2.60 (1.60, 3.90)	0.929
GCS score	9.00 (3.80, 15.00)	9.00 (3.80, 15.00)	2.54 (2.19, 3.43)	<0.001
Arterial pH	7.45 (7.39, 7.49)	7.45 (7.39, 7.49)	7.49 (7.45, 7.51)	<0.001
Anion gap	16.00 (12.00, 25.00)	16.00 (12.00, 25.00)	137.00 (129.00, 142.00)	<0.001
Blood urea nitrogen	37.00 (23.00, 60.00)	37.00 (23.00, 60.00)	72.00 (65.00, 79.00)	0.026
Chronic kidney disease (Yes)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1.000
Liver disease/cirrhosis (Yes)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1.000
Malignancy (Yes)	104 (0.2%)	104 (0.2%)	0 (0.0%)	1.000
Bicarbonate, mmol/L	22.10 (6.10, 27.90)	22.10 (6.10, 27.90)	4.60 (4.20, 5.40)	<0.001
Magnesium	2.20 (2.00, 2.50)	2.20 (2.00, 2.50)	2.35 (1.90, 2.80)	0.745
PaO2, mmHg	140.00 (92.80, 228.00)	140.00 (92.80, 228.00)	197.00 (159.00, 235.00)	0.247
FiO2	70.00 (40.00, 100.00)	70.00 (40.00, 100.00)	75.00 (60.00, 90.00)	0.958
APTT	36.00 (30.00, 48.50)	36.00 (30.00, 48.50)	45.00 (45.00, 45.00)	
ICU mortality (Yes)	2552 (5.7%)	2532 (5.6%)	20 (22.2%)	0.185
Hypertension (Yes)	7173 (15.9%)	7147 (15.9%)	26 (28.9%)	<0.001
ALT	4.80 (4.40, 5.30)	4.80 (4.40, 5.30)	5.20 (4.60, 6.05)	<0.001
AST	143.00 (140.00, 146.00)	143.00 (140.00, 146.00)	147.00 (140.00, 149.00)	0.004
Glucose	122.82 (107.66, 149.07)	122.79 (107.62, 149.02)	129.78 (115.04, 153.43)	0.042
Hemoglobin, g/dL	35.20 (31.70, 39.20)	35.20 (31.70, 39.20)	33.85 (30.92, 37.68)	0.061
Phosphate	4.40 (3.70, 5.50)	4.40 (3.70, 5.50)	5.00 (4.10, 6.70)	<0.001

SOFA score	4.57 (3.60, 5.31)	4.57 (3.61, 5.31)	3.43 (2.77, 4.35)	<0.001
Lymphocyte count	13.40 (8.30, 20.40)	13.40 (8.30, 20.30)	14.40 (7.65, 23.90)	0.364
Albumin, g/dL	3.20 (2.80, 3.70)	3.30 (2.80, 3.70)	3.10 (2.73, 3.50)	0.015
Body mass index, kg/m^2	29.10 (25.10, 34.70)	29.10 (25.10, 34.70)	29.15 (22.65, 38.83)	0.888
Calcium	138.00 (135.00, 140.00)	138.00 (135.00, 140.00)	138.00 (131.50, 142.00)	0.940
Temperature, °C	37.17 (36.62, 37.64)	37.17 (36.62, 37.64)	37.69 (37.69, 37.69)	
C-reactive protein	138.50 (94.00, 229.00)	138.00 (93.00, 229.00)	184.50 (147.50, 252.50)	0.002

Table 2. Baseline characteristics of the merged cohort by development and validation sets (random 70/30 split).

Variable	**Overall**	**Development (70%)**	**Validation (30%)**	**P value**
N	45141	31599	13542	
Database: eICU	30380 (67.3%)	21232 (67.2%)	9148 (67.6%)	0.461
Database: MIMIC-IV	14761 (32.7%)	10367 (32.8%)	4394 (32.4%)	
Sex: Male	24495 (54.3%)	17175 (54.4%)	7320 (54.1%)	0.857
Sex: Female	20640 (45.7%)	14420 (45.6%)	6220 (45.9%)	
Sex: Other/Unknown	2 (0.0%)	1 (0.0%)	1 (0.0%)	
Sex: Unknown	4 (0.0%)	3 (0.0%)	1 (0.0%)	
Race/ethnicity: White	9672 (21.4%)	6792 (21.5%)	2880 (21.3%)	0.952
Race/ethnicity: Black	1078 (2.4%)	751 (2.4%)	327 (2.4%)	
Race/ethnicity: Hispanic	456 (1.0%)	312 (1.0%)	144 (1.1%)	
Race/ethnicity: Asian	391 (0.9%)	270 (0.9%)	121 (0.9%)	
Race/ethnicity: Other	1026 (2.3%)	723 (2.3%)	303 (2.2%)	

Race/ethnicity: Unknown	32518 (72.0%)	22751 (72.0%)	9767 (72.1%)	
Age, years	67.00 (55.00, 77.00)	66.00 (55.00, 77.00)	67.00 (56.00, 77.00)	0.151
CRPA infection (Yes)	90 (0.2%)	65 (0.2%)	25 (0.2%)	0.730
MRSA infection (Yes)	683 (1.5%)	472 (1.5%)	211 (1.6%)	0.637
VRE infection (Yes)	296 (0.7%)	207 (0.7%)	89 (0.7%)	1.000
Sodium, mmol/L	143.00 (140.00, 146.00)	143.00 (140.00, 147.00)	143.00 (140.00, 146.00)	0.128
Serum creatinine	1.50 (0.98, 2.72)	1.50 (0.97, 2.76)	1.50 (0.99, 2.70)	0.758
White blood cell count	16.80 (12.00, 23.20)	16.80 (12.00, 23.20)	16.80 (12.00, 23.20)	0.577
Potassium, mmol/L	4.70 (4.30, 5.20)	4.70 (4.30, 5.30)	4.70 (4.30, 5.20)	0.944
Congestive heart failure (Yes)	7466 (16.5%)	5181 (16.4%)	2285 (16.9%)	0.226
Diabetes (Yes)	4406 (9.8%)	3068 (9.7%)	1338 (9.9%)	0.600
Chloride, mmol/L	105.00 (2.70, 112.00)	105.00 (2.70, 112.00)	105.00 (2.70, 112.00)	0.654
Platelet count	204.00 (152.00, 308.00)	203.00 (152.00, 307.00)	205.00 (153.00, 310.00)	0.259
Mechanical ventilation (Yes)	7231 (16.0%)	5039 (15.9%)	2192 (16.2%)	0.523
Vasopressor use (Yes)	4814 (10.7%)	3389 (10.7%)	1425 (10.5%)	0.539
INR	1.25 (1.10, 1.53)	1.25 (1.10, 1.53)	1.26 (1.10, 1.54)	0.946
Total bilirubin	0.80 (0.50, 1.60)	0.80 (0.50, 1.60)	0.80 (0.50, 1.60)	0.036
Lactate, mmol/L	2.50 (1.60, 4.20)	2.50 (1.60, 4.20)	2.50 (1.50, 4.20)	0.214
GCS score	9.00 (3.80, 15.00)	9.00 (3.79, 15.00)	9.00 (3.80, 15.00)	0.572
Arterial pH	7.45 (7.39, 7.49)	7.45 (7.39, 7.49)	7.44 (7.39, 7.49)	0.379
Anion gap	16.00 (12.00, 25.00)	16.00 (12.00, 25.00)	16.00 (12.00, 25.00)	0.973
Blood urea nitrogen	37.00 (23.00, 60.00)	37.00 (23.00, 60.00)	37.00 (23.00, 60.00)	0.912
Chronic kidney disease (Yes)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1.000
Liver disease/cirrhosis (Yes)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1.000
Malignancy (Yes)	104 (0.2%)	76 (0.2%)	28 (0.2%)	0.544
Bicarbonate, mmol/L	22.10 (6.10, 27.90)	22.00 (6.10, 27.90)	22.20 (6.10, 27.80)	0.834
Magnesium	2.20 (2.00, 2.50)	2.20 (2.00, 2.50)	2.20 (2.00, 2.50)	0.245
PaO2, mmHg	140.00 (92.80, 228.00)	140.85 (93.00, 227.00)	140.00 (92.00, 232.00)	0.915
FiO2	70.00 (40.00, 100.00)	70.00 (40.00, 100.00)	70.00 (40.00, 100.00)	0.830
APTT	36.00 (30.00, 48.50)	36.00 (30.00, 48.80)	36.00 (30.00, 48.00)	0.136
ICU mortality (Yes)	2552 (5.7%)	1818 (5.8%)	734 (5.4%)	0.231
Hypertension (Yes)	7173 (15.9%)	5040 (15.9%)	2133 (15.8%)	0.919
ALT	4.80 (4.40, 5.30)	4.80 (4.40, 5.30)	4.80 (4.40, 5.30)	0.968
AST	143.00 (140.00, 146.00)	143.00 (140.00, 146.00)	143.00 (140.00, 146.00)	0.035
Glucose	122.82 (107.66, 149.07)	123.00 (107.67, 149.18)	122.53 (107.61, 148.65)	0.852
Hemoglobin, g/dL	35.20 (31.70, 39.20)	35.20 (31.70, 39.30)	35.20 (31.60, 39.10)	0.930
Phosphate	4.40 (3.70, 5.50)	4.40 (3.70, 5.50)	4.40 (3.70, 5.50)	0.735
SOFA score	4.57 (3.60, 5.31)	4.57 (3.59, 5.30)	4.57 (3.63, 5.33)	0.399
Lymphocyte count	13.40 (8.30, 20.40)	13.30 (8.20, 20.30)	13.55 (8.50, 20.60)	0.137
Albumin, g/dL	3.20 (2.80, 3.70)	3.30 (2.80, 3.70)	3.20 (2.80, 3.70)	0.452
Body mass index, kg/m^2	29.10 (25.10, 34.70)	29.00 (25.00, 34.60)	29.10 (25.20, 34.80)	0.525
Calcium	138.00 (135.00, 140.00)	138.00 (135.00, 140.00)	138.00 (135.00, 140.00)	0.501
Temperature, °C	37.17 (36.62, 37.64)	37.16 (36.61, 37.63)	37.20 (36.63, 37.66)	0.304

C-reactive protein	138.50	(94.00, 229.00)	143.00	(95.25, 241.00)	131.00	(91.00, 204.00)	0.003

Table 3. Top 10 features by mean absolute SHAP value across outcomes (MRSA, VRE, CRPA) in final models excluding length-of-stay variables.

Rank	**MRSA feature**	**MRSA mean|SHAP|**	**VRE feature**	**VRE mean|SHAP|**	**CRPA feature**	**CRPA mean|SHAP|**
1	eosinophils	0.450	fibrinogen	1.421	potassium	0.837
2	po2	0.261	lymphocytes	0.774	ldl	0.598
3	platelet.1	0.216	pco2	0.456	lymphocytes	0.346
4	white_blood_cell_count	0.211	white_blood_cell_count	0.355	alt	0.236
5	hemoglobin	0.209	platelet	0.333	alveolar_arterial_oxygen_gradient	0.234
6	glucose	0.202	—	—	basophils	0.199
7	age	0.199	—	—	age	0.197
8	avg_sofa	0.165	—	—	phosphate	0.183
9	—	—	—	—	cholesterol	0.179
10	—	—	—	—	ph	0.175

Figure 1. Boruta feature importance for MRSA infection prediction. Boxplots show feature importance scores across Boruta iterations, with shadow features used as the reference for acceptance. Higher importance indicates a robust contribution beyond randomized noise.

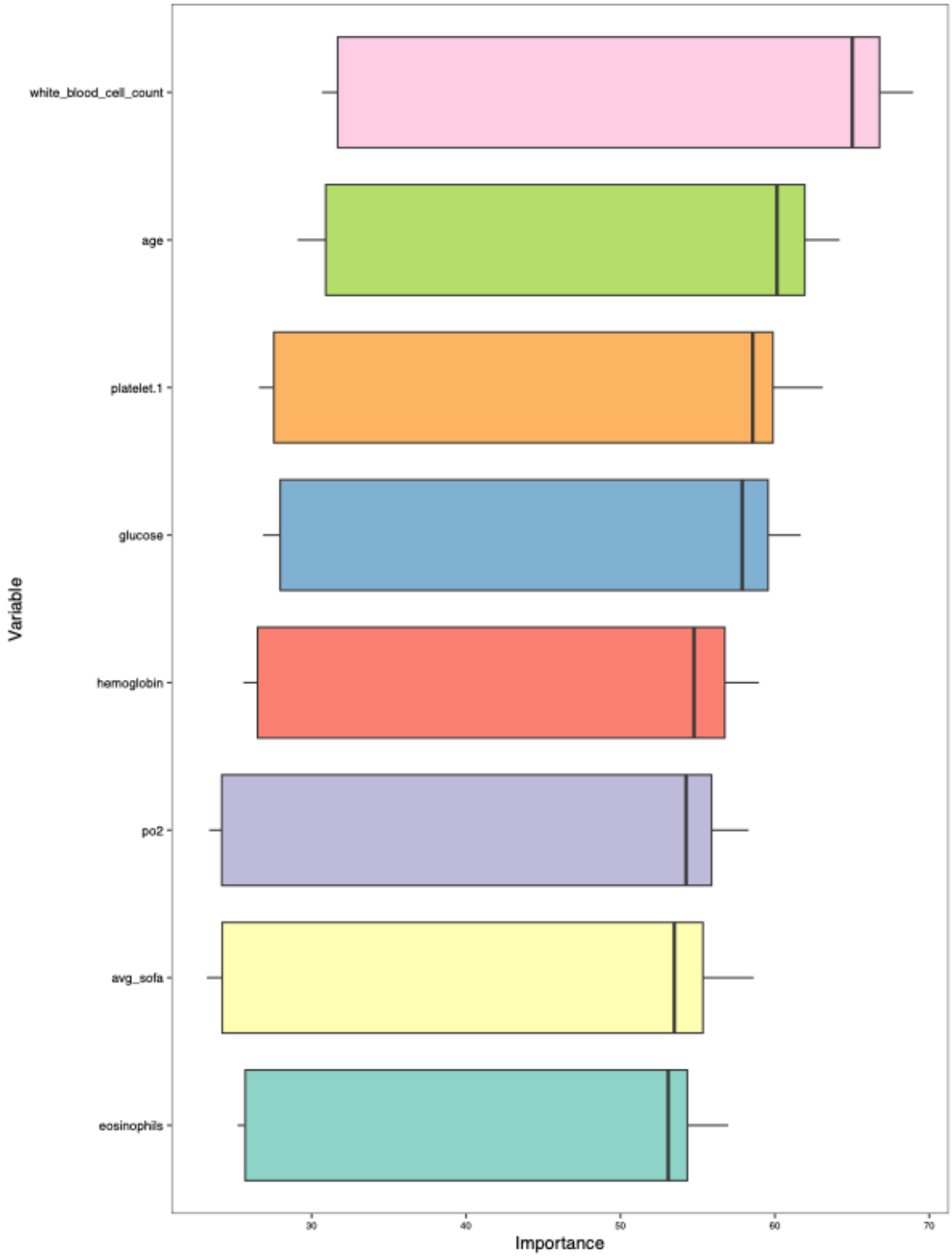


Figure 2. Boruta feature importance for VRE infection prediction. Feature importance distributions are compared with shadow feature maxima to identify robust predictors; accepted features consistently exceed the shadow reference.

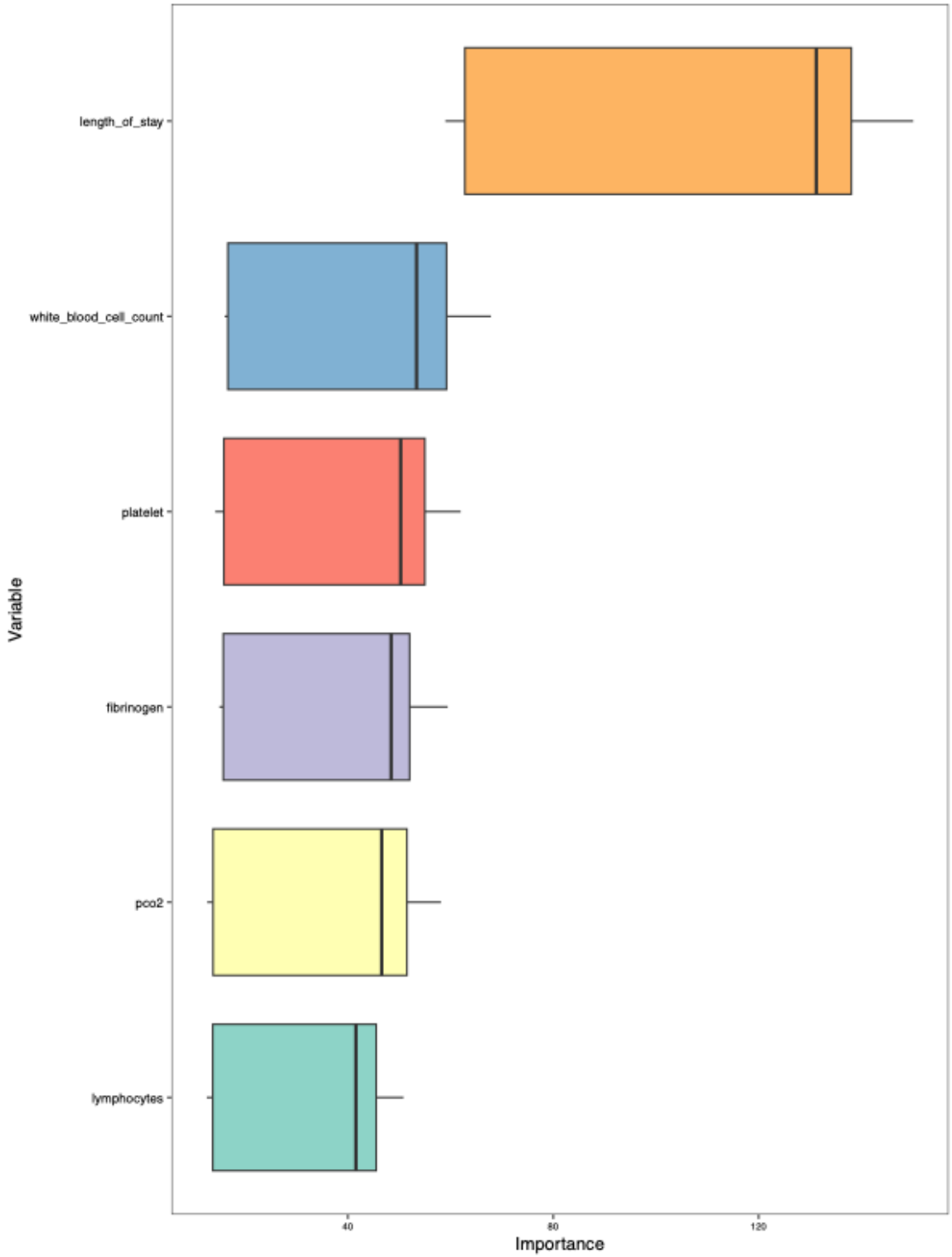


Figure 3. Boruta feature importance for CRPA infection prediction. Importance distributions across iterations highlight top-ranked predictors and distinguish accepted, tentative, and rejected features relative to shadow controls.

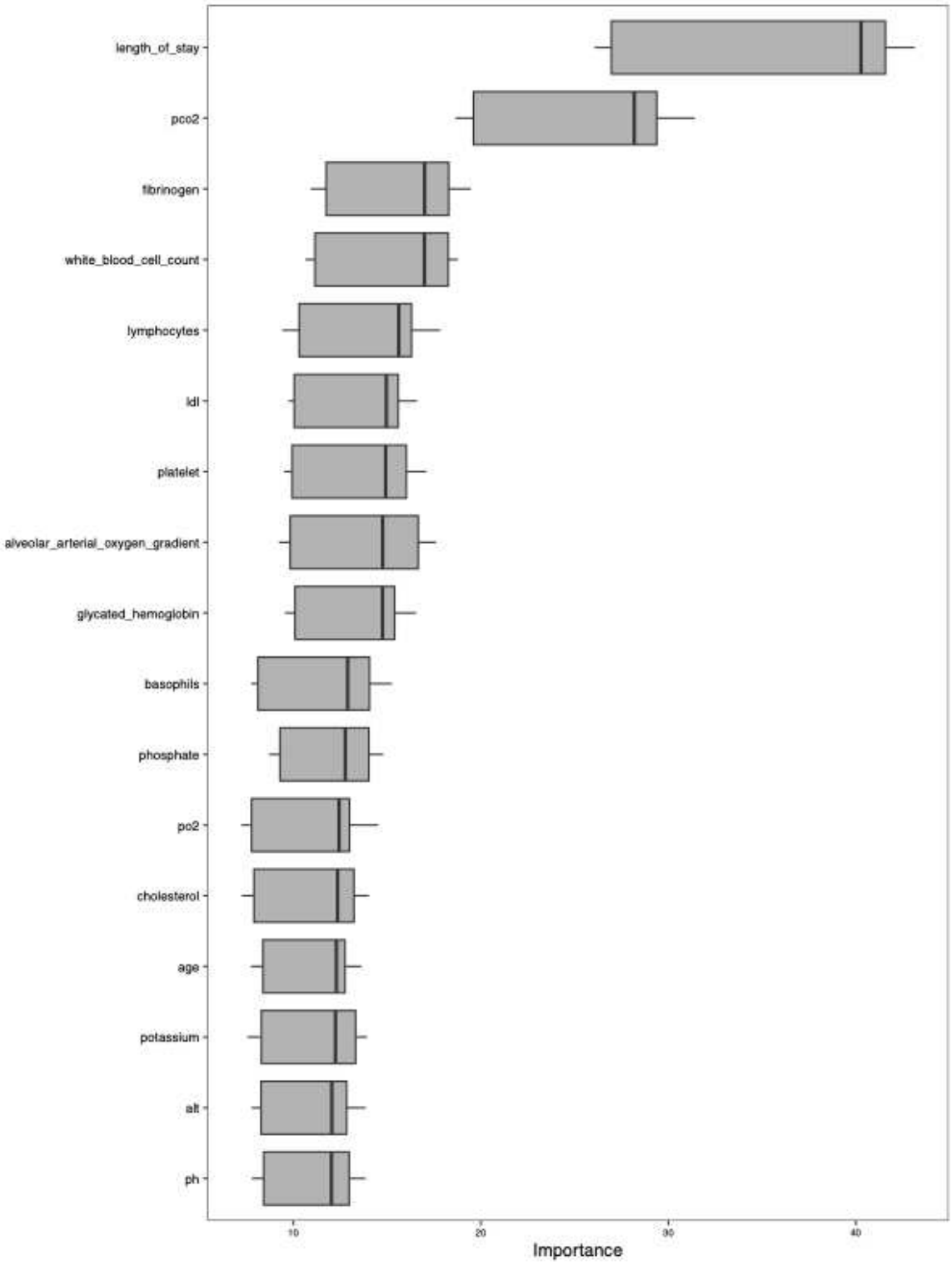


Figure 4. SHAP summary plot for MRSA infection prediction. Each point represents an individual admission and shows the SHAP value (x-axis) for a given feature (y-axis), with

features ordered by mean absolute impact. Positive SHAP values increase predicted risk, and negative values decrease risk.

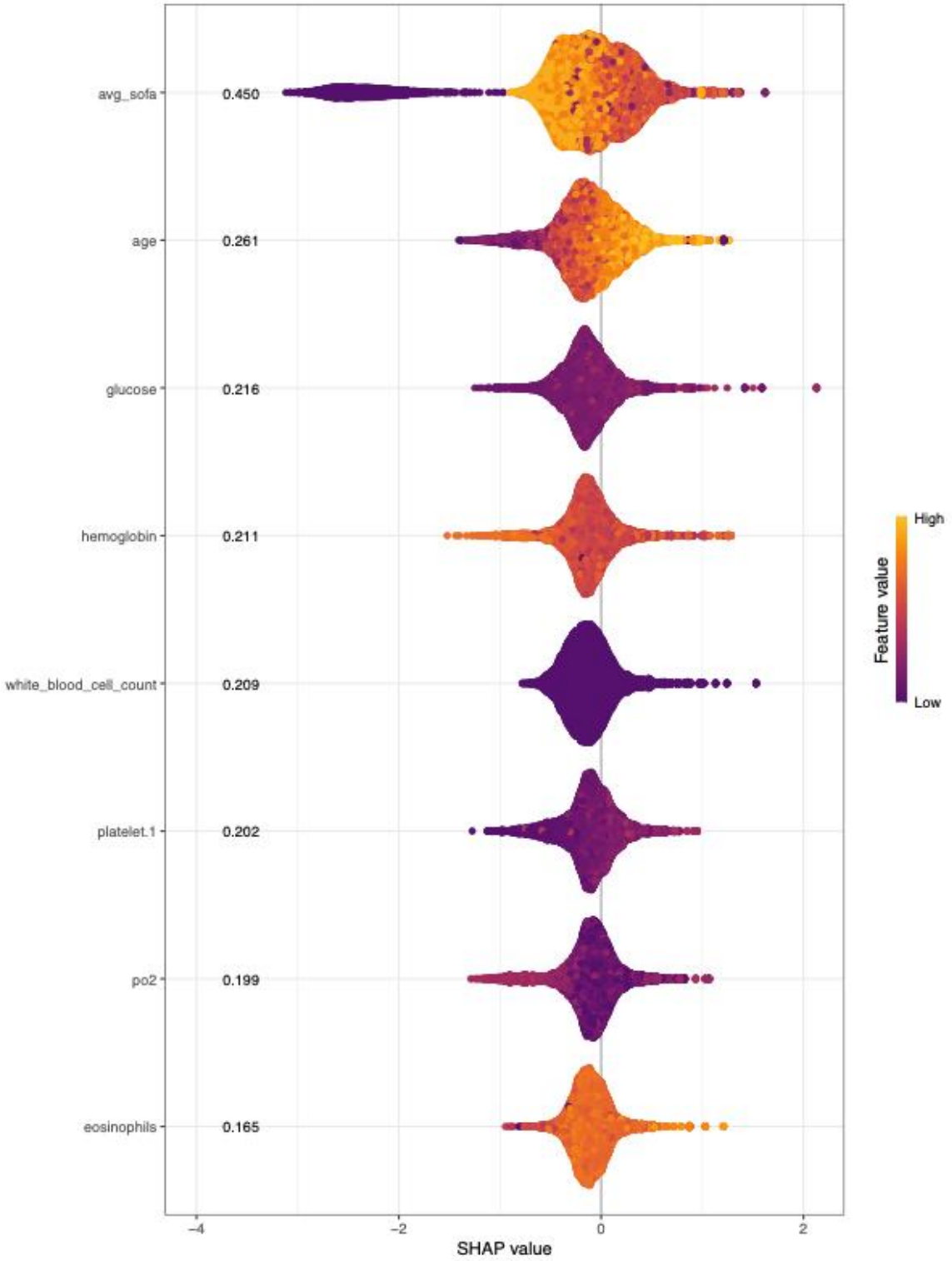


Figure 5. SHAP summary plot for VRE infection prediction. Global feature importance is summarized by the distribution of SHAP values across admissions, enabling interpretation of both magnitude and directionality of effects.

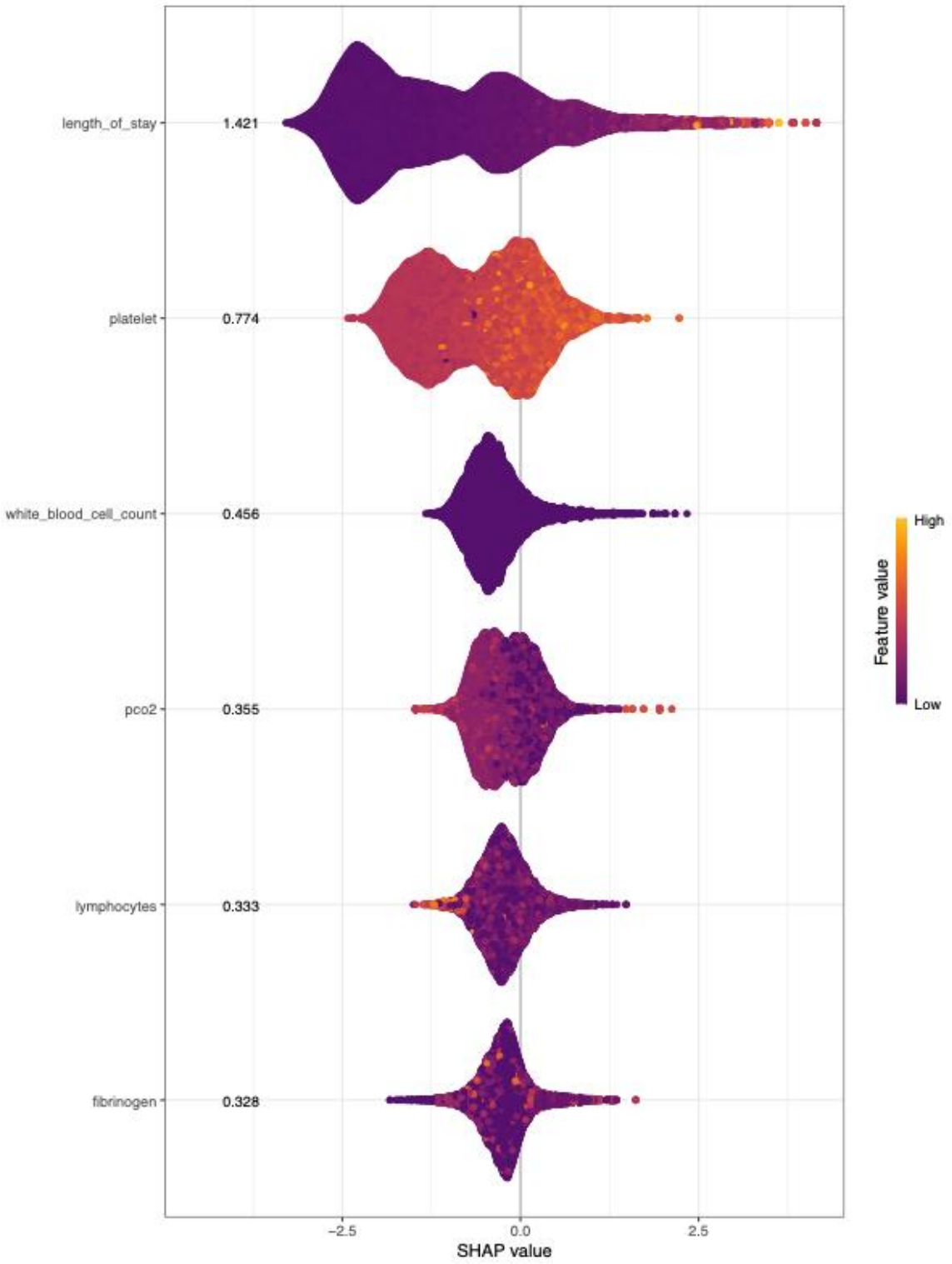


Figure 6. SHAP summary plot for CRPA infection prediction. Features are ordered by overall contribution, illustrating heterogeneity of individual-level contributions and dominant physiologic/laboratory signals. Source: crpa_shap.pdf.

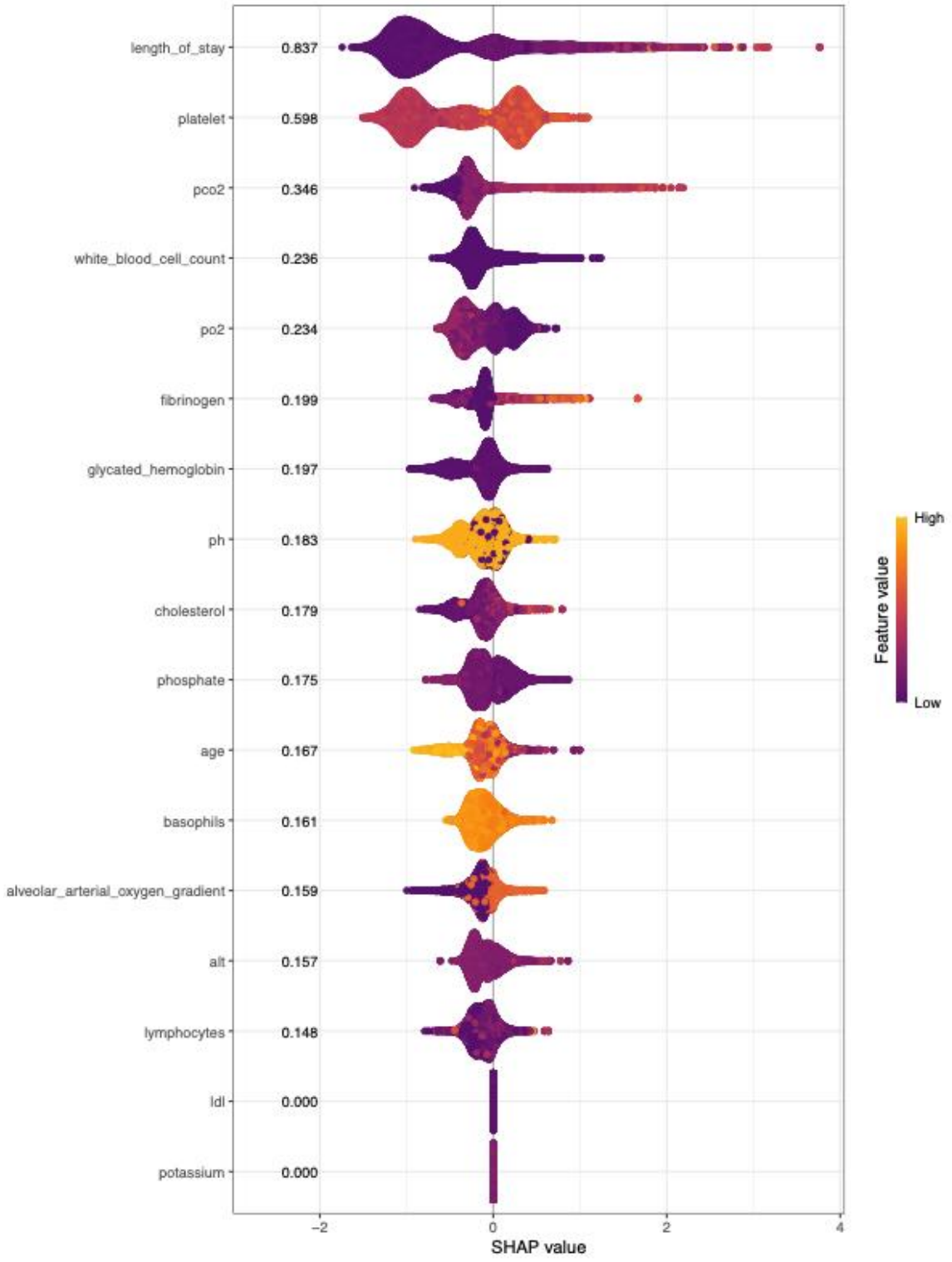


Figure 7. Receiver operating characteristic (ROC) curves for MRSA infection prediction in the validation set. Curves compare LASSO, AdaBoost, LightGBM, XGBoost, and CatBoost, with AUC values shown in the legend.

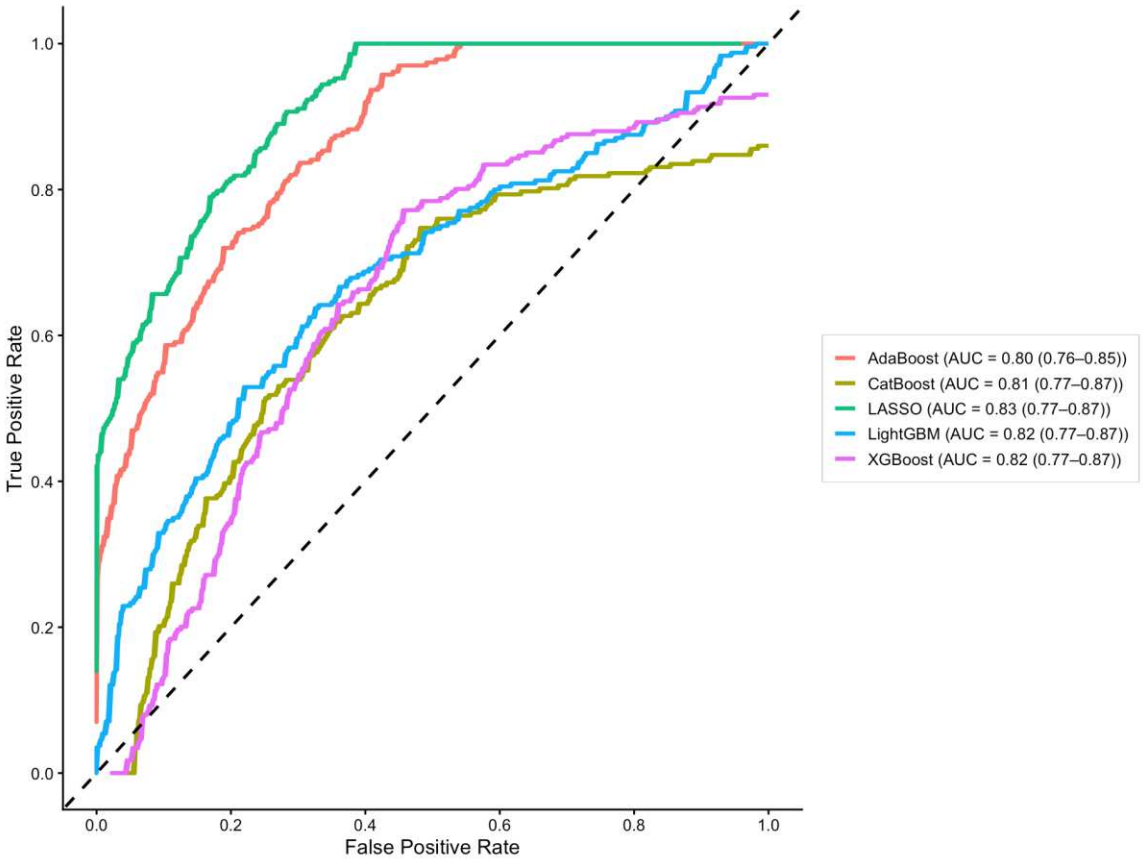


Figure 8. Receiver operating characteristic (ROC) curves for VRE infection prediction in the validation set. Model discrimination is summarized by AUC values in the legend, highlighting comparative performance across classifiers.

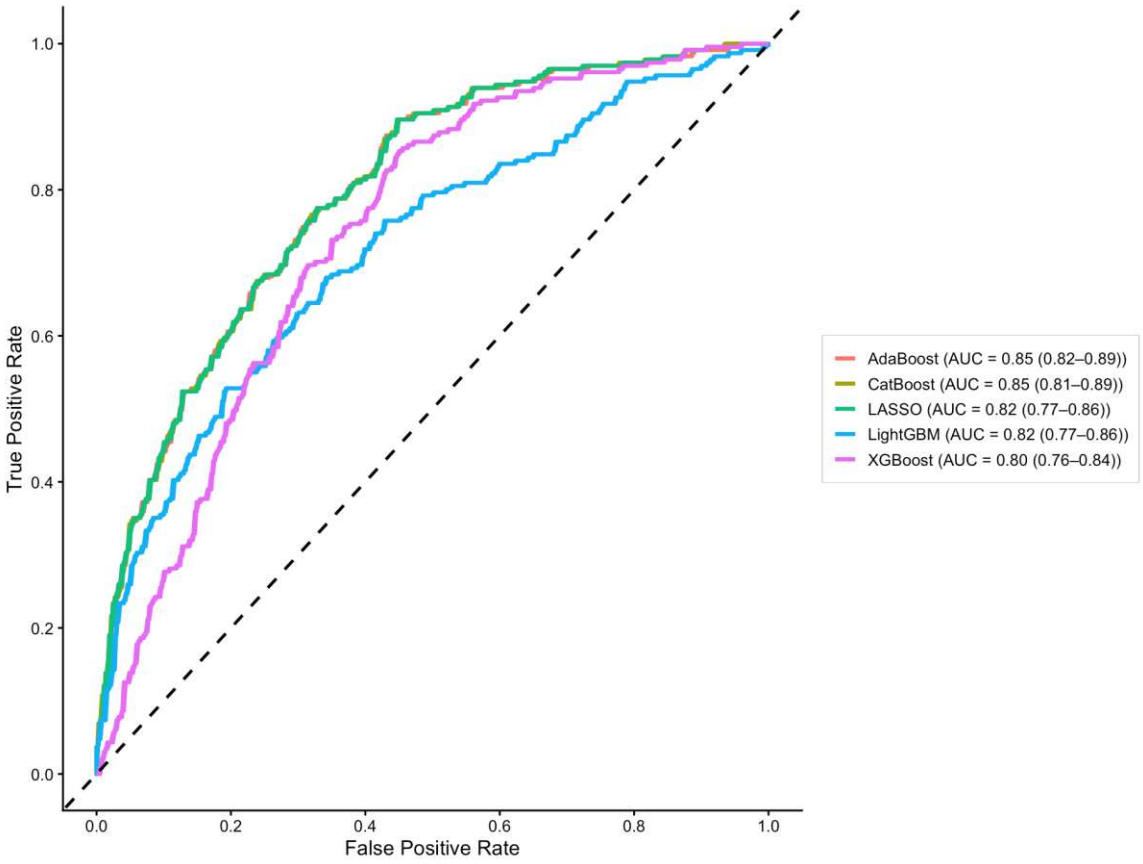


Figure 9. Receiver operating characteristic (ROC) curves for CRPA infection prediction in the validation set. ROC curves and AUC values illustrate discrimination differences across model families.

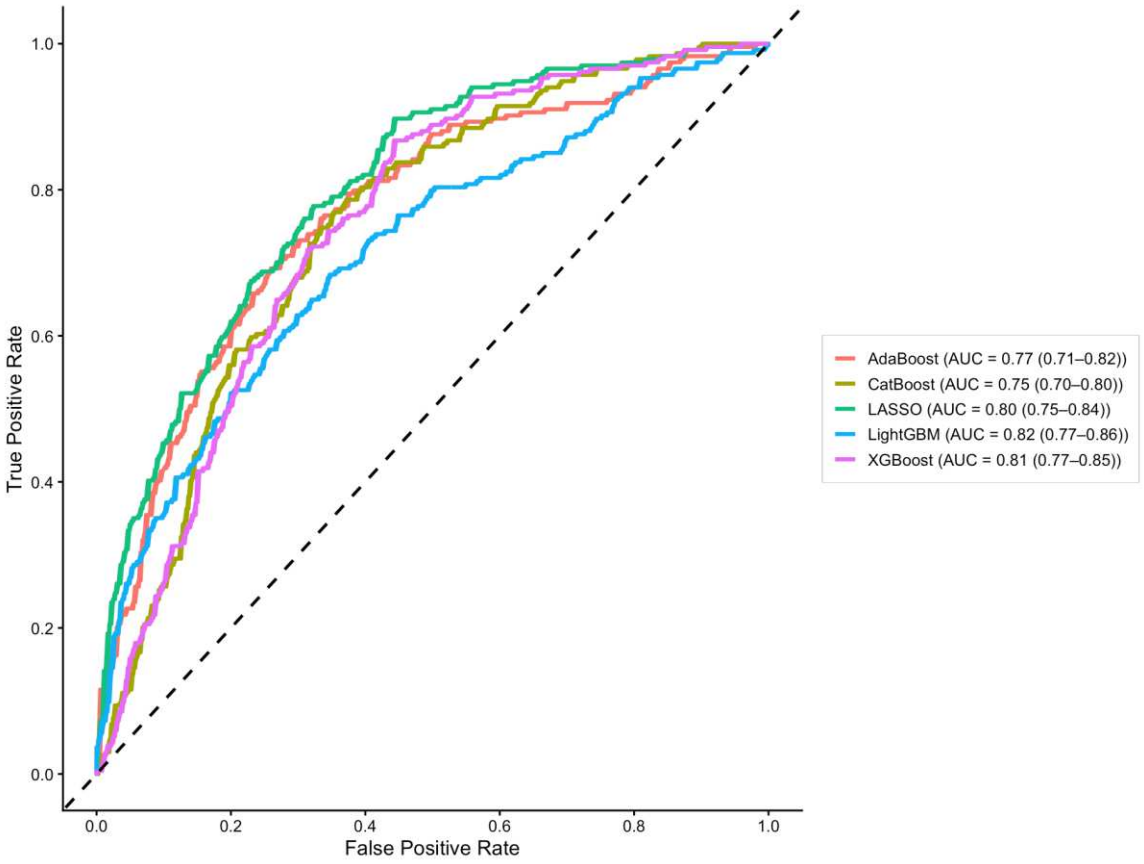


Figure 10. Decision curve analysis (DCA) for MRSA infection prediction. Net benefit is plotted across threshold probabilities (0–50%) for each classifier and compared with treat-all and treat-none strategies.

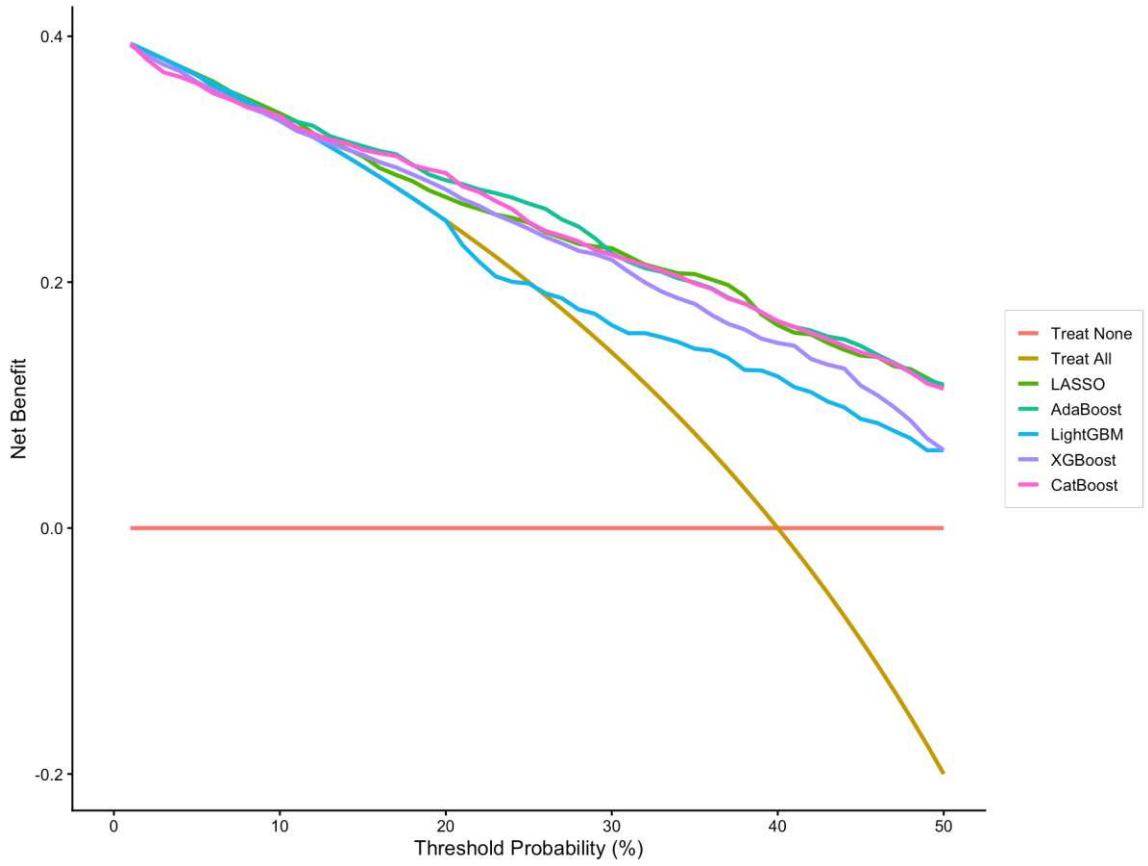


Figure 11. Decision curve analysis (DCA) for VRE infection prediction. Net benefit curves indicate potential clinical utility across threshold probabilities relative to treat-all and treat-none strategies.

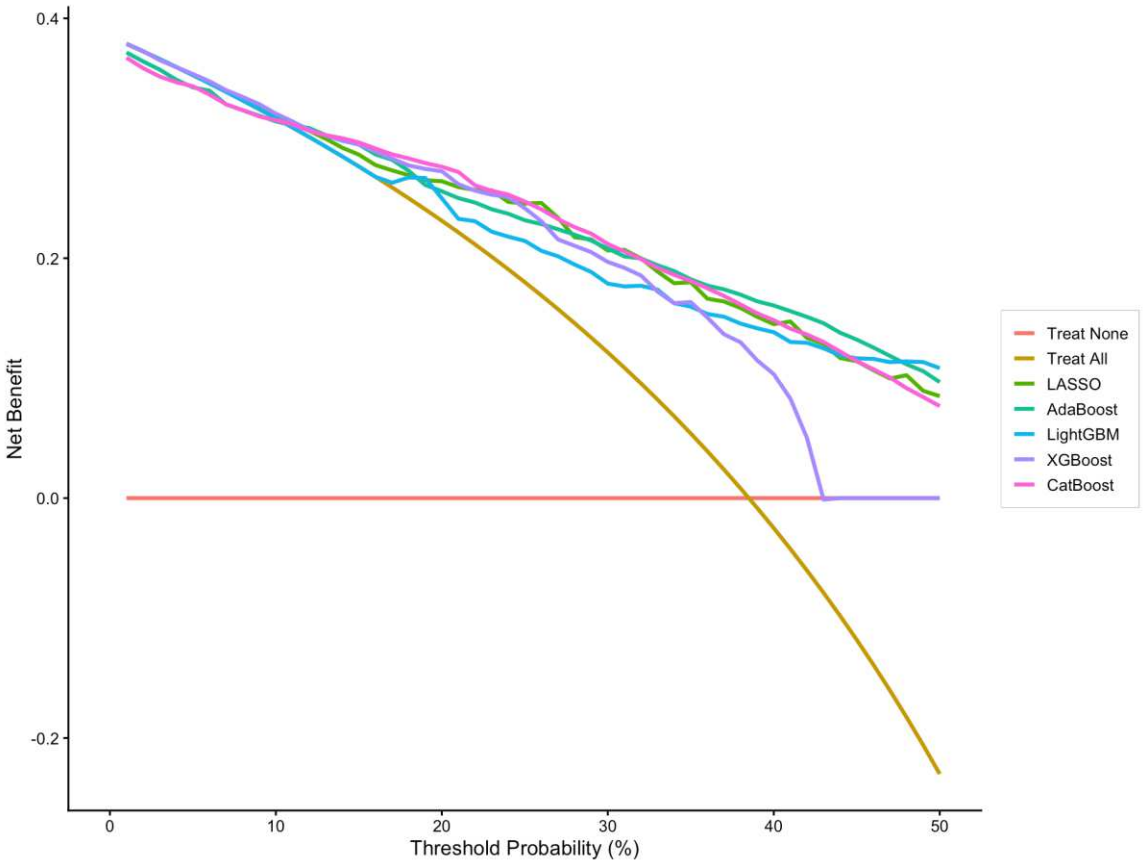
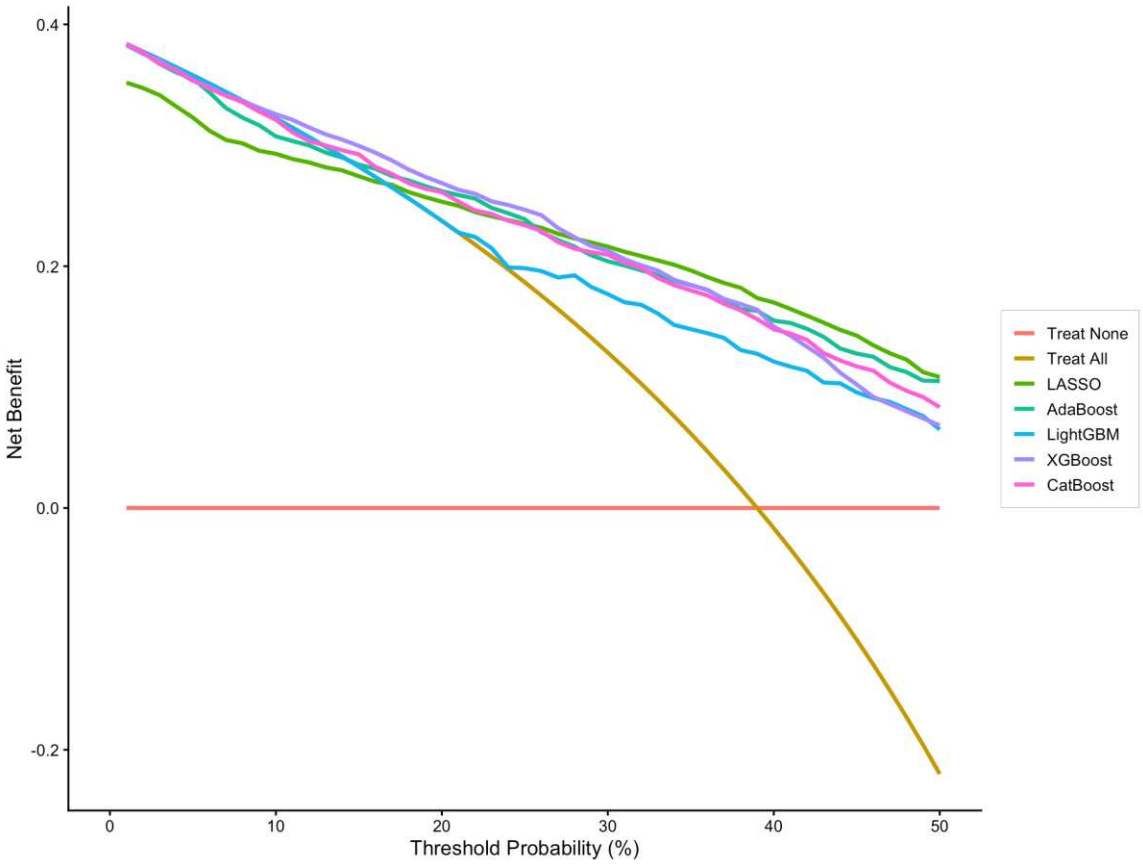


Figure 12. Decision curve analysis (DCA) for CRPA infection prediction. Net benefit across threshold probabilities (0–50%) compares classifiers and benchmark strategies to assess clinical usefulness.


0.4
0.2
0.0
-0.2
Net Benefit
0
10
20
30
40
50
Threshold Probability (%)
Treat None
Treat All
LASSO
AdaBoost
LightGBM
XGBoost
CatBoost

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformationTablesS1S7withCaptions.docx