Published in final edited form as:

Crit Care Med. 2020 November ; 48(11): e1020–e1028. doi:10.1097/CCM.0000000000004556.

The development and validation of a machine learning model to predict bacteremia and fungemia in hospitalized patients using electronic health record data

Sivasubramaniam V Bhavani, MD¹, Zachary Lonjers, MD², Kyle A Carey, MPH¹, Majid Afshar, MD, MSCR³, Emily R Gilbert, MD³, Nirav S Shah, MD, MPH², Elbert S Huang, MD¹, Matthew M Churpek, MD, PhD⁴

¹University of Chicago Medical Center, Chicago, IL

²Northshore Hospital, Chicago, IL

³Loyola University Medical Center, Chicago, IL

⁴University of Wisconsin, Madison, WI

Abstract

Objective: Bacteremia and fungemia can cause life-threatening illness with high mortality rates, which increase with delays in antimicrobial therapy. The objective of this study is to develop machine learning models to predict blood culture results at the time of the blood culture order using routine data in the electronic health record (EHR).

Design: Retrospective analysis of a large, multicenter inpatient data.

Setting: Two academic tertiary medical centers between the years 2007 and 2018.

Subjects: All hospitalized patients who received a blood culture during hospitalization.

Intervention(s): The dataset was partitioned temporally into development and validation cohorts: the logistic regression and gradient boosting machine (GBM) models were trained on the earliest 80% of hospital admissions and validated on the most recent 20%.

Measurements and Main Results: There were 252,569 blood culture days – defined as non-overlapping 24-hour periods in which one or more blood cultures were ordered. In the validation cohort, there were 50,514 blood culture days, with 3,762 cases of bacteremia (7.5%) and 370 cases of fungemia (0.7%). The GBM model for bacteremia had significantly higher AUC (0.78 [95% CI 0.77-0.78]) than the logistic regression model (0.73 [0.72-0.74]) ($p < 0.001$). The model identified a high-risk group with over 30 times the incidence of bacteremia in the low-risk group (27.4 vs 0.9%, $p < 0.001$). Using the low-risk cut-off, the model identifies bacteremia with 98.7% sensitivity. The GBM model for fungemia had high discrimination (AUC 0.88 [95% CI 0.86-0.90]). The high-risk fungemia group had 252 fungemic cultures compared to one fungemic culture in the low-risk

group (5.0% vs 0.02%, $p<0.001$). Further, the high-risk group had a mortality rate 60 times higher than the low-risk group (28.2 vs 0.4%, $p<0.001$).

Conclusions: Our novel models identified patients at low and high-risk for bacteremia and fungemia using routinely collected EHR data. Further research is needed to evaluate the cost-effectiveness and impact of model implementation in clinical practice.

Keywords

Infection; Sepsis; Machine learning; Artificial intelligence; Fungemia; Bacteremia

Introduction

Bacteremia and fungemia can cause life-threatening illness, with mortality rates between 18% and 35%. (1-3) Currently, blood cultures are the conventional approach to diagnose bloodstream infections, but cultures have significant deficits as a diagnostic test. Blood culture results often require more than 24 hours of incubation. (4) Additionally, the yield of blood cultures is variable while contamination rates are high, with recent studies suggesting up to 17% of positive cultures are contaminants. (5,6) Developing more rapid and accurate approaches to identifying patients at high risk of bacteremia and fungemia could reduce health care costs related to unnecessary testing and treatment, and could increase targeted, early prescribing of antimicrobial therapy that may lower mortality. (7-10)

Apart from blood cultures, biomarker tests have shown inconsistent results in predicting bacteremia. For example, studies have reported area under the receiver operating characteristic curve (AUC) values for procalcitonin ranging between 0.62 and 0.78 depending on the patient cohort. Discrimination for bacteremia is low in patients meeting systemic inflammatory response syndrome (SIRS) criteria or all-comers in the emergency department. (11-14) Importantly, just as with blood cultures, these biomarker results may not be available at the time the clinician first suspects an infection. Molecular methods using polymerase chain reaction (PCR) to detect bacteremia and fungemia have higher accuracy than blood cultures or biomarker tests. (15, 16) However, these methods are more expensive and require complex infrastructure, such as the ability to run PCR around-the-clock, which is not available at most hospitals. (15)

Routinely available data in the electronic health record (EHR) could be used to predict bacteremia and fungemia resulting in real-time risk-stratification of patients. Prior studies have developed predictive models using clinical data in specific patient cohorts, such as patients in the intensive care unit (ICU) and specific infectious processes such as pneumonia or pyelonephritis. (17-22) These studies are limited in generalizability to the overall population of hospitalized patients. Other models have required physician input of clinical signs and symptoms, which may limit automated integration into the EHR. (23, 24) Most studies have used traditional, linear models to build predictive models, with even the best performing models having only moderate discriminative ability. (24, 25) Modern machine learning models may perform better by leveraging the complex interactions in the clinical data. However, there is a paucity of research using machine learning approaches on large multicenter cohorts to predict bacteremia and fungemia.

The objective of this study was to develop and validate machine learning models to predict bacteremia and fungemia using routine EHR data available at the time of the blood culture order. We investigated whether the model could identify low-risk patients who may forego blood cultures, and high-risk patients who may require earlier and more broad-spectrum antimicrobial therapy.

Methods

Patient setting

All adult patients admitted to the University of Chicago Hospital between 2008 and 2018 and Loyola University Medical Center between 2007 and 2017 were eligible for study inclusion in this retrospective analysis of prospectively collected data. Demographic data, International Classification of Diseases (ICD)-9/10 billing codes, clinician orders, blood culture results, vital signs, nursing assessments, and laboratory values were extracted from each institution's clinical data warehouse from their respective EHRs (Epic, Verona, WI), de-identified, and made available for analyses. Based on general impracticability and minimal harm, waivers of consent were granted by the University of Chicago Institutional Review Board (IRB#150195) and Loyola University Chicago IRB (LU#209950).

Blood cultures

The primary outcome in the bacteremia prediction model was whether or not a blood culture was positive for bacteremia. The primary outcome for the fungemia prediction model was whether or not a blood culture was positive for fungemia. Blood culture day was the unit of analysis to account for the fact that patients often receive multiple blood cultures, sometimes minutes to hours apart. A blood culture day was defined as a distinct non-overlapping 24-hour period of hospitalization in which one or more blood cultures were ordered. A blood culture day started at the time of the first blood culture order in that 24-hour period and encompassed all other blood culture orders placed during the subsequent 24 hours. A blood culture day was considered "positive" for bacteremia or fungemia if any blood culture ordered during that 24-hour period resulted in non-contaminant bacterial or fungal growth. Contaminants were identified and excluded based on prior studies: coagulase-negative staphylococci (CONS), *Corynebacterium* species, *Bacillus* species other than *Bacillus anthracis*, *Propionibacterium acnes*, *Micrococcus* species, and *Clostridium perfringens*. (26) The most common contaminant is CONS, but this organism may represent a true infection in approximately 12% of cases. Therefore, cultures growing CONS were considered positive if the organism was present on two separate blood culture days within 72 hours of each other. The contaminant criteria were reviewed by an infectious disease physician, and two trained physicians ran quality checks on blood culture results for incorrect classifications.

Predictor variables

We used EHR data available at the time of the first blood culture order in a given blood culture day (time zero) in the development of the predictive model; these data included demographics, vital signs, laboratory results, nursing assessments, prior culture results, and physician orders. Results of prior cultures were included in the model if the culture was obtained more than 24 hours prior to the current culture. The result of the prior culture was

specified as either positive or negative. Additionally, the total number of prior culture orders were also included as a feature in the model. A comprehensive list of all EHR variables included as predictors along with their corresponding proportion of missing values between bacteremic and non-bacteremic patients are available in eTable 1. Data were considered missing if a vital sign was not available in the EHR in the 24 hours prior to the blood culture order or if a laboratory marker was not available in the EHR 48 hours prior to the blood culture order. If there was only one value for a vital sign or laboratory marker, that value was considered the maximum and minimum value. Standard deviation was considered missing if two or more values were not available for the vital sign. EHR variables were excluded if present in <5% of patients. Missing values were imputed using median/mode values in the logistic regression model and were left missing for the gradient boosting machine (GBM) model, which automatically handles missingness. The missingness itself may carry inherent information and could improve the predictive accuracy of the model. We did not utilize any statistical variable selection process prior to or during model development.

Model development and validation

The earliest 80% of hospital admissions by date were assigned to the training cohort, and the most recent 20% of hospital admissions were assigned to the validation cohort. The time-dependent cohort split was designed to simulate prospective implementation of the predictive model at both hospitals. We developed logistic regression and GBM models on the training cohort using all variables and tested these models in the validation cohort to predict bacteremia. GBM is an ensemble algorithm in which decision trees are sequentially fit to the training data with the difficult-to-predict cases weighted more highly in order to increase model accuracy. Five-fold cross validation was performed in the training set to tune hyperparameters (shrinkage, depth of trees and minimum weight) to maximize AUC. Cross-validation was not performed for logistic regression. Separate logistic regression and GBM models to predict fungemia were developed and validated using the same approach. All models were developed following the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) guidelines.(27)

Evaluating performance of models predicting bacteremia

We compared the performance of the logistic regression and GBM bacteremia prediction models on the validation data for all blood culture days (primary analysis) using AUC.(28) We evaluated the calibration of the most accurate model by plotting deciles of model predicted risk against the incidence of bacteremia in those deciles. To further assess accuracy, we calculated the sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV) of the most accurate model at varying cut-offs. We compared the final models to the SIRS criteria, Infection Probability Score (IPS), and Modified Early Warning Score (MEWS), tools designed to identify and risk stratify patients with infection. (29-32) The most important variables in the GBM model were determined using the scaled metric of mean decrease in impurity.

Identification of low-risk and high-risk bacteremia cohorts

Applying the most accurate model to the validation cohort, we defined a low-risk bacteremia cohort as patients in the lowest decile of predicted risk. We hypothesized that the overall rate

of contamination would be higher than the risk of bacteremia in this cohort. Similarly, we defined a high-risk bacteremia cohort as the highest decile predicted risk of bacteremia. We tested our hypothesis that patients in this high-risk cohort would have a high incidence of resistant organisms that may not be covered with empiric antibiotic therapy (i.e., ESKAPE pathogens: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species).(33) We also tested our hypothesis that high-risk patients may not have received antibiotics in the six hours following the blood culture order. Patients who were not low- or high-risk were classified as intermediate-risk. To test the prognostic implications of risk categories, the length of stay and mortality rates were compared between risk groups.

Evaluating model performance in subgroup analyses

We evaluated the discrimination of the bacteremia prediction model in subgroups based on time from admission to blood culture, location of the blood culture order (emergency department [ED], wards or intensive care unit), and completeness of data (i.e., patients with missing vital signs). We also evaluated model performance separately for the initial blood culture of each admission and then on only subsequent blood cultures.

Evaluating performance of models predicting fungemia

We evaluated the performance and calibration of the fungemia prediction models in the same way as the bacteremia models. Similar to the bacteremia model, we defined a low-risk and high-risk fungemic group based on percentile of predicted risk. The length of stay and mortality rates were compared between risk groups.

Model development for external validation

In a sensitivity analysis, we tested the external generalizability of our modeling strategy. The blood culture data from the University of Chicago were used as the training data to develop a new GBM model to predict bacteremia. Five-fold cross validation was performed in the training set to tune hyperparameters to maximize AUC. This model was then validated in the Loyola University data.

Results

Patient characteristics

Between the two medical centers, there were a total of 252,569 blood culture days (eFigure 1). These blood culture days were from 135,094 unique hospital admissions in 76,688 unique patients. Of the blood culture days, 19,527 were positive for bacteremia (7.7%). There were 12,397 gram-positive (4.9%), 7,570 gram-negative organisms (3.0%), and 9,329 ESKAPE organisms (3.7%). There were 1,776 cultures positive for fungemia (0.7%). Overall, the rate of contamination was 3.5% with 8,851 contaminants. There were significant differences in the patient demographic, physiological, and processes of care characteristics between the bacteremic and non-bacteremic culture days (Table 1). Importantly, bacteremic patients had a greater frequency of comorbidities and more deranged vital signs and laboratory values. Patient admissions with bacteremia had a mortality rate of 14.5% compared to 5.3% for patients without bacteremia ($p<0.001$).

Similarly, fungemic patients had more abnormalities in vital signs and laboratory values (eTable 2), and had a mortality rate of 33% vs. 6.0% for those without fungemia ($p<0.001$).

Model performance for predicting bacteremia

In the validation cohort, there were a total of 50,514 blood culture days, with 3,762 bacteremic cultures (7.5%). The GBM model had significantly greater discrimination (AUC 0.78 [95% CI 0.77-0.78]) than the logistic regression model (AUC 0.73 [95% CI 0.72-0.74]) in predicting bacteremia ($p<0.001$). In a post-hoc analysis, removing collinear variables and using restricted cubic splines for continuous variables did not improve discrimination in the logistic regression model (AUC 0.73 for both models). A plot of deciles of GBM model predicted risk compared to the actual incidence of bacteremia is shown in Figure 1, with incidence of bacteremia increasing across deciles of predicted probabilities from 0.9% to 27.4%. The GBM model's sensitivity, specificity, NPV and PPV based on varying cut-offs for bacteremia are presented in Table 2.

The SIRS criteria, MEWS and IPS had poor discrimination and calibration for predicting bacteremia (eFigures 2-4). Comparisons between the GBM and logistic regression models along with these infection scores are presented in Figure 2.

Important predictors of bacteremia

As shown in Figure 3, time from admission to blood culture order, maximum temperature, and age were important predictors of bacteremia. The majority of the most important predictors were maximum, minimum, and standard deviation of vital signs and measures of interaction between vitals such as heart rate, respiratory rate and blood pressure. The relative importance of all variables in the bacteremia model are detailed in eTable 3.

Low- and high-risk blood cultures in the validation cohort

The rate of bacteremia was 30 times higher in high-risk compared to low-risk cultures (27.4% vs 0.9%, $p<0.001$). One potential workflow would be alerting clinicians if a patient has a high-risk culture. This would result in a clinical workload of 3.6 – which means there would be one case of bacteremia for every 3.6 alerts the clinician would receive. The rate of potentially resistant organisms was over 50 times higher in high-risk compared to low-risk cultures (15.0% vs 0.3%, $p<0.001$). The mortality rate was over 20 times higher in patients with high-risk cultures compared to patients with low-risk cultures (12.7 vs 0.6%, $p<0.001$; eTable 4).

Model performance in predicting bacteremia in subgroups

The GBM model had the highest accuracy in predicting results of blood culture orders placed between 72 and 168 hours of admission with an AUC of 0.82 (95% CI 0.80-0.85), and the lowest performance in blood cultures placed less than one hour into admission (AUC 0.70 [95% CI 0.69-0.72]) (eFigure 5), which likely relates to the amount of non-missing data available for the model. The model performance was the highest in blood cultures obtained on the wards ($n = 20,201$; AUC 0.80, 95% CI 0.79-0.81), followed by the intensive care unit ($n = 12,175$; AUC 0.79, 95% CI 0.78-0.81), and the ED ($n = 17,689$; AUC 0.74, 95% CI 0.72-0.75) (eFigure 6). The model had decreased performance in the subgroup of patients

with missing vital signs ($n = 3,085$; AUC 0.71, 95% CI 0.68-0.75), likely due to the importance of these vital signs in the model. The GBM model's accuracy was 0.75 (95% CI 0.74-0.76) in predicting the results of the initial blood culture, and 0.81 (95% CI 0.80-0.82) in predicting the results of subsequent blood cultures. Low-risk and high-risk cohorts are compared in the initial and subsequent cultures in eTables 5 and 6.

Model performance on external validation

The site-specific GBM model was developed on the 137,589 cultures at University of Chicago and tested on the 114,980 cultures at Loyola. The performance was lower than the temporal validation analysis with an AUC of 0.73 (0.73-0.74). The variable importance was unchanged, with predominantly vital signs as the most important variables.

Model performance in predicting fungemia

Of the 50,514 blood culture days in the validation cohort, there were 370 fungemic cultures (0.7%). Both the GBM model and logistic regression model had high accuracy in predicting fungemia and were not significantly different (GBM model - AUC 0.88 [95% CI 0.86-0.90]; logistic regression - AUC 0.85 [95% CI 0.82-0.87]; $p=0.08$). The deciles of predicted risk of fungemia by the GBM model is compared to the actual incidence of fungemia in eFigure 7. The GBM model's sensitivity, specificity, NPV and PPV for fungemia are presented in eTable 7. The most important predictors of fungemia included prior fungemia and time from admission to the blood culture (Figure 3 and eTable 8).

Low-risk and high-risk cultures were defined for fungemia as the lowest and highest decile of predicted risk. Of the 50,514 validation cultures, 5,047 were defined as low-risk, and only one of these cultures were positive for fungemia (0.02%). In the 5,053 high-risk cultures, there were 252 fungemic cultures (5.0%). The mortality rate was over 60 times higher in high-risk patients compared to low-risk patients (28.2 vs 0.4%, $p<0.001$) (eTable 9).

Discussion

We present a large, multicenter study on the development and validation of novel machine learning models to predict bacteremia and fungemia at the time of blood culture orders. Our model does not require any clinician input or non-routine biomarker data and thus can be integrated into EHR systems. Most importantly, our model identifies clinically relevant low-risk and high-risk patients. The prevalence of bacteremia was 7.7% in our two centers, which is in the lower end of reported rates in the literature; the model's PPV may be higher in other medical centers with higher prevalence.(34-36) The high-risk patients identified by our model had over 30 times the risk of bacteremia and over 50 times the risk of potentially resistant organisms as low-risk patients. Given the high mortality from resistant organisms, it is essential to identify these high-risk patients early and potentially focus antibiotic therapy based on local resistance patterns.(33, 37) Patients identified by the model as high-risk for fungemia had 20 times greater risk of mortality than low-risk patients. Early identification of potentially fungemic patients could result in earlier therapy and decreased mortality.(8-10)

A significant proportion of hospitalized patients receive culture orders during their stay despite the fact that many of these cultures are low yield.(5-7, 38) Our model identifies low

risk groups with positive culture rates of less than 1%; these patients have a higher risk of growing a contaminant than a true pathogen. Integration of our models in the EHR as a decision support tool has the potential to decrease culture ordering and its resultant costs and harms. Furthermore, we found that almost 50% of the highest-risk patients did not receive antibiotics within one hour of their blood culture order; even six hours after the blood culture, 18% of these high-risk patients had not received antibiotics. Given that prior work suggests that every hour of delay in providing appropriate antibiotics is associated with increased mortality, our developed algorithms could potentially prompt lifesaving therapy in the highest risk patients.(8, 10)

While most blood culture prediction models were developed for specific patient cohorts, our models are generalizable to all hospitalized patients with community or hospital-acquired infections of any type.(17-22, 24) For example, Shapiro and colleagues developed a rule-based decision system to identify patients at low-risk for bacteremia, but was limited to patients presenting to the ED.(24) More recently, Roimi et al. published a machine learning model predicting bacteremia in ICU patients using EHR data, with high accuracy on internal validation.(20) However, on external validation, their model performance deteriorated with an AUC of 0.59-0.60. Their tool differs from ours in that their model only works in the ICU and requires 48 hours of clinical data. Importantly, our models performed well across patient subgroups, including patients in the ED, wards and ICU. Reasonable performance was demonstrated even in patients with limited data (e.g., patients who presented less than an hour prior to the blood culture order or patients with missing vital signs and labs). Although our model performance decreased on cross-site analysis, this novel tool still offers significant advantages over the inconsistent wait-times and results from blood cultures. The generalizability and applicability of the model across sites and time allows its use in varied clinical settings.

Our machine learning model was significantly more accurate than established clinical scores, such as the SIRS criteria, IPS, and MEWS.(29, 30, 32) This high accuracy was in part due to the large multicenter training data, large number of variables, and the flexible nature of the GBM model. Importantly, inclusion of measures of interaction between vital signs like shock index and the product of respiratory and heart rates played an important role in the model's performance. Recent studies have shown that dynamic measures that capture complex interactions of hemodynamic parameters may perform better than traditional vital signs in predicting complex outcomes.(38-42)

Our study has several limitations. First, our model was developed in academic centers within a single city, which may limit the generalizability of our findings. Second, we did not have information on when the blood culture results were available in the EHR. We included the results of prior blood cultures in the model if they were ordered more than 24 hours prior, but it is possible that these data may not have been available at the time of the current blood culture order for some patients. Third, although we found high-risk patients did not receive prompt antibiotics, there was not enough granularity in the data to evaluate the potential clinical reasons behind withholding antibiotics. Finally, there is an inherent risk with the model resulting in clinicians foregoing cultures in low-risk patients or over-treating high-risk patients with antibiotics.

In conclusion, we developed novel machine learning algorithms that predict bacteremia and fungemia with high accuracy and provide clinically relevant risk stratification at the time of the blood culture order. Our predictive algorithm has the potential to serve as a clinical decision support tool. Specifically, the model could help reduce unnecessary blood cultures in low-risk patients and prompt earlier antimicrobial therapy in high-risk patients. Transformation into an effective tool will require further engagement with relevant stakeholders to identify appropriate clinician alert thresholds and the workflows that would arise from these alerts. Further research is necessary to evaluate the patient and clinician benefits of integrating these models into clinical practice.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Copyright form disclosure: Drs. Bhavani, Afshar, Shah, Huang, and Churpek received support for article research from the National Institutes of Health (NIH). Dr. Shah's institution received funding from the NIH. Dr. Huang's institution received funding from the NIH/NIDDK. Dr. Churpek's institution received funding from NIH NIGMS (R01 GM123193), NIH NIDDK (R21DK113420-01A1), and EarlySense (Tel Aviv, Israel), and he disclosed that he has a patent pending (ARCD. P0535US.P2) for risk stratification algorithms for hospitalized patients. The remaining authors have disclosed that they do not have any potential conflicts of interest.

References

1. Nielsen SL, Lassen AT, Kolmos HJ, et al.: The daily risk of bacteremia during hospitalization and associated 30-day mortality evaluated in relation to the traditional classification of bacteremia. *American journal of infection control* 2016, 44(2):167–172. [PubMed: 26577629]
2. Goto M, Al-Hasan M: Overall burden of bloodstream infection and nosocomial bloodstream infection in North America and Europe. *Clinical Microbiology and Infection* 2013, 19(6):501–509. [PubMed: 23473333]
3. Søgaaard M, Nørgaard M, Dethlefsen C, et al.: Temporal Changes in the Incidence and 30-Day Mortality associated with Bacteremia in Hospitalized Patients from 1992 through 2006: A Population-based Cohort Study. *Clinical Infectious Diseases* 2011, 52(1):61–69. [PubMed: 21148521]
4. Pardo J, Klinker KP, Borgert SJ, et al.: Time to positivity of blood cultures supports antibiotic de-escalation at 48 hours. *The Annals of pharmacotherapy* 2014, 48(1):33–40. [PubMed: 24259644]
5. Dargère S, Cormier H, Verdon R. Contaminants in blood cultures: importance, implications, interpretation and prevention. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. 2018;24(9):964–969.
6. Coburn B, Morris AM, Tomlinson G, et al.: Does This Adult Patient With Suspected Bacteremia Require Blood Cultures? *JAMA* 2012, 308(5).
7. Gander RM, Byrd L, DeCrescenzo M, et al.: Impact of Blood Cultures Drawn by Phlebotomy on Contamination Rates and Health Care Costs in a Hospital Emergency Department. *Journal of Clinical Microbiology* 2009, 47(4):1021–1024. [PubMed: 19171686]
8. Liu VX, Fielding-Singh V, Greene JD, et al.: The Timing of Early Antibiotics and Hospital Mortality in Sepsis. *Am J Respir Crit Care Med* 2017, 196(7):856–863. [PubMed: 28345952]
9. Garey KW, Rege M, Pai MP, et al.: Time to initiation of fluconazole therapy impacts mortality in patients with candidemia: a multi-institutional study. *Clin Infect Dis* 2006, 43(1):25–31. [PubMed: 16758414]
10. Ibrahim EH, Sherman G, Ward S, et al.: The influence of inadequate antimicrobial treatment of bloodstream infections on patient outcomes in the ICU setting. *Chest* 2000, 118(1):146–155. [PubMed: 10893372]

11. Caffarini EM, DeMott J, Patel G, et al.: Determining the Clinical Utility of an Absolute Procalcitonin Value for Predicting a Positive Culture Result. *Antimicrobial Agents and Chemotherapy* 2017, 61(5):e02007–02016. [PubMed: 28193661]
12. Hoenigl M, Raggam RB, Wagner J, et al.: Procalcitonin fails to predict bacteremia in SIRS patients: a cohort study. *International journal of clinical practice* 2014, 68(10):1278–1281. [PubMed: 24898888]
13. Lin CT, Lu JJ, Chen YC, et al.: Diagnostic value of serum procalcitonin, lactate, and high-sensitivity C-reactive protein for predicting bacteremia in adult patients in the emergency department. *PeerJ* 2017, 5:e4094. [PubMed: 29201568]
14. Yang M, Choi SJ, Lee J, et al.: Serum procalcitonin as an independent diagnostic markers of bacteremia in febrile patients with hematologic malignancies. *PLoS One* 2019, 14(12):e0225765. [PubMed: 31821331]
15. Liesenfeld O, Lehman L, Hunfeld KP, et al.: Molecular diagnosis of sepsis: New aspects and recent developments. *European journal of microbiology & immunology* 2014, 4(1):1–25. [PubMed: 24678402]
16. Bloos F, Hinder F, Becker K, et al.: A multicenter trial to compare blood culture with polymerase chain reaction in severe human sepsis. *Intensive Care Med* 2010, 36(2):241–247. [PubMed: 19924398]
17. Kim B, Choi J, Kim K, et al.: Bacteremia Prediction Model for Community-acquired Pneumonia: External Validation in a Multicenter Retrospective Cohort. *Academic Emergency Medicine* 2017, 24(10):1226–1234. [PubMed: 28714287]
18. Falguera M, Trujillano J, Caro S, et al.: A Prediction Rule for Estimating the Risk of Bacteremia in Patients with Community-Acquired Pneumonia. *Clinical Infectious Diseases* 2009, 49(3):409–416. [PubMed: 19555286]
19. Kim KS, Kim K, Jo YH, et al.: A simple model to predict bacteremia in women with acute pyelonephritis. *The Journal of infection* 2011, 63(2):124–130. [PubMed: 21722666]
20. Roimi M, Neuberger A, Shrot A, et al.: Early diagnosis of bloodstream infections in the intensive care unit using machine-learning algorithms. *Intensive Care Med* 2020.
21. Beeler C, Dbeibo L, Kelley K, et al.: Assessing patient risk of central line-associated bacteremia via machine learning. *American journal of infection control* 2018, 46(9):986–991. [PubMed: 29661634]
22. 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. [PubMed: 27358356]
23. Paul M, Andreassen S, Nielsen AD, et al.: Prediction of Bacteremia Using TREAT, a Computerized Decision-Support System. *Clinical Infectious Diseases* 2006, 42(9):1274–1282. [PubMed: 16586387]
24. Shapiro NI, Wolfe RE, Wright SB, et al.: Who needs a blood culture? A prospectively derived and validated prediction rule. *The Journal of emergency medicine* 2008, 35(3):255–264. [PubMed: 18486413]
25. Eliakim-Raz N, Bates DW, Leibovici L: Predicting bacteraemia in validated models--a systematic review. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 2015, 21(4):295–301.
26. Hall KK, Lyman JA: Updated Review of Blood Culture Contamination. *Clinical Microbiology Reviews* 2006, 19(4):788–802. [PubMed: 17041144]
27. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1–W73. doi:10.7326/M14-0698 [PubMed: 25560730]
28. DeLong ER, DeLong DM, Clarke-Pearson DL: Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988, 44(3):837–845. [PubMed: 3203132]
29. Churpek MM, Snyder A, Sokol S, et al.: Investigating the Impact of Different Suspicion of Infection Criteria on the Accuracy of Quick Sepsis-Related Organ Failure Assessment, Systemic

- Inflammatory Response Syndrome, and Early Warning Scores. *Crit Care Med* 2017, 45(11):1805–1812. [PubMed: 28737573]
30. Subbe CP, Kruger M, Rutherford P, et al.: Validation of a modified Early Warning Score in medical admissions. *QJM : monthly journal of the Association of Physicians* 2001, 94(10):521–526. [PubMed: 11588210]
31. Simpson SQ: SIRS in the Time of Sepsis-3. *Chest* 2018, 153(1):34–38. [PubMed: 29037526]
32. Peres Bota D, Melot C, Lopes Ferreira F, et al.: Infection Probability Score (IPS): A method to help assess the probability of infection in critically ill patients. *Crit Care Med* 2003, 31(11):2579–2584. [PubMed: 14605527]
33. Boucher HW, Talbot GH, Bradley JS, et al.: Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin Infect Dis* 2009, 48(1):1–12. [PubMed: 19035777]
34. Chen AI, Bilker WB, Hamilton KW, et al. Blood culture utilization at an academic hospital: Addressing a gap in benchmarking. *Infect Control Hosp Epidemiol.* 2018;39(11):1353–1359. [PubMed: 30261936]
35. Venturelli C, Righi E, Borsari L, et al. Impact of Pre-Analytical Time on the Recovery of Pathogens from Blood Cultures: Results from a Large Retrospective Survey. *PLoS One.* 2017;12(1):e0169466. [PubMed: 28046040]
36. Nannan Panday RS, Wang S, van de Ven PM, et al. Evaluation of blood culture epidemiology and efficiency in a large European teaching hospital. *PLoS One.* 2019;14(3):e0214052. [PubMed: 30897186]
37. McDanel J, Schweizer M, Crabb V, et al.: Incidence of Extended-Spectrum beta-Lactamase (ESBL)-Producing *Escherichia coli* and *Klebsiella* Infections in the United States: A Systematic Literature Review. *Infect Control Hosp Epidemiol* 2017, 38(10):1209–1215. [PubMed: 28758612]
38. Bhavani SV, Carey KA, Gilbert ER, et al.: Identifying Novel Sepsis Subphenotypes Using Temperature Trajectories. *Am J Respir Crit Care Med* 2019.
39. Fang WF, Huang CH, Chen YM, et al.: Application of dynamic pulse pressure and vasopressor tools for predicting outcomes in patients with sepsis in intensive care units. *Journal of critical care* 2019, 52:156–162. [PubMed: 31078024]
40. Brown SM, Tate MQ, Jones JP, et al.: Coefficient of Variation of Coarsely Sampled Heart Rate is Associated With Early Vasopressor Independence in Severe Sepsis and Septic Shock. *Journal of intensive care medicine* 2015, 30(7):420–425. [PubMed: 24578465]
41. Middleton DJ, Smith TO, Bedford R, et al.: Shock Index Predicts Outcome in Patients with Suspected Sepsis or Community-Acquired Pneumonia: A Systematic Review. *Journal of clinical medicine* 2019, 8(8).
42. Jiang L, Caputo ND, Chang BP: Respiratory adjusted shock index for identifying occult shock and level of Care in Sepsis Patients. *The American journal of emergency medicine* 2019, 37(3):506–509. [PubMed: 30674427]

Figure 1: Deciles of model predicted risk compared to the incidence of bacteremia.

The gradient boosting machine (GBM) model predicted the risk of bacteremia using electronic health record (EHR) data available at the time of the blood culture order. The calibration of the model is shown here with deciles of predicted risk compared to the incidence of bacteremia in the validation cohort (n=50,514). Only 0.9% of patients in the lowest decile of predicted risk have bacteremia, while 27.4% of patients in the highest decile have bacteremia.

Figure 2: Comparison of gradient boosting machine, logistic regression and published risk scores for predicting bacteremia.

The gradient boosting machine (GBM) model had the highest discrimination with an AUC of 0.78 (95% CI 0.77-0.78) followed by the logistic regression (LR) model (AUC 0.73 [95% CI 0.72-0.74]). The infection prediction scores had much lower discrimination, with the Infection Probability Score (IPS) having an AUC of 0.55, Modified Early Warning Score (MEWS) having an AUC of 0.54, and Systemic Inflammatory Response Syndrome (SIRS) criteria having an AUC of 0.52.

Figure 3: Variable importance plot illustrating the most important predictors of bacteremia and fungemia.

In the gradient boosting machine (GBM) models, variable importance was measured by mean decrease in impurity adjusted to a scale of 0 to 100. Time from admission to the blood culture order was the most important predictor of bacteremia, while prior fungemia was the most important predictor of fungemia. The majority of the most important variables in both models were traditional vital signs such as temperature as well as interaction measures between vital signs, such as the product of respiratory and heart rates. *Definition of abbreviations:* max – maximum; min – minimum; sd – standard deviation; RR – respiratory rate; HR – heart rate; dBP – diastolic blood pressure; sBP – systolic blood pressure; PP – pulse pressure; PPI – pulse pressure index; SI – shock index; BUN – blood urea nitrogen.

Table 1.

Clinical characteristics compared between nonbacteremic and bacteremic blood culture days.

Characteristics	Nonbacteremia (233,042)	Bacteremia (19,527)	P-value
Demographics			
Age, yrs, median (IQR)	59 (47-70)	60 (49-70)	<0.001
Male sex (%)	53	56	<0.001
Race (%)			0.4
Black	40	41	
White	49	49	
Other	10	10	
Comorbidities (%)			
CHF	19	23	<0.001
COPD	16	15	0.001
Diabetes mellitus	24	28	<0.001
HTN	44	50	<0.001
Liver disease	12	13	<0.001
Metastatic cancer	8.2	9.0	<0.001
Renal failure	21	26	<0.001
Vital signs, mean (sd)			
Heart rate, max	105 (22)	109 (23)	<0.001
sBP, min	111 (23)	108 (24)	<0.001
dBP, min	60 (16)	57 (16)	<0.001
Respiratory rate, max	23.6 (7)	24.0 (7)	<0.001
Temperature, max	37.4 (1.0)	37.6 (1.1)	<0.001
Temperature, min	36.5 (0.8)	36.6 (1.0)	<0.001
Laboratory values, mean (sd)			
WBC, max	12.9 (11)	13.1 (11)	0.006
WBC, min	9.3 (7.2)	9.3 (7.3)	0.3
Blood urea nitrogen	30 (25)	36 (30)	<0.001
Creatinine	1.9 (2.1)	2.3 (2.6)	<0.001
Lactic acid	2.5 (2.4)	2.8 (2.8)	<0.001
Processes of care			
Time from admission to culture, hrs, median (IQR)	44 (3.0-174)	40 (1.7-197)	<0.001
Location of blood culture			<0.001
Emergency department	27	32	
Wards	45	41	
Intensive care unit	28	27	
Foley (%)	29	30	0.003
Central line (%)	6.9	12	<0.001

Definition of abbreviations: yrs – years; CHF – congestive heart failure; COPD – chronic obstructive pulmonary disease; HTN – hypertension; max – maximum; min – minimum; sBP – systolic blood pressure; dBP – diastolic blood pressure; WBC – white blood cell count; LOS – total length of inpatient stay for that admission; hrs – hours; d – days; IQR – interquartile range; sd – standard deviation.

Table 2.

Sensitivity, specificity, positive predictive value and negative predictive value of the model in predicting bacteremia across varying cut-offs.

Risk cut-off	N	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
0.011	45,422	98.7	10.8	8.2	99.1
0.015	40,354	96.2	21.4	9.0	98.6
0.019	35,370	93.4	31.9	9.9	98.4
0.023	30,258	89.4	42.5	11.1	98.0
0.029	25,231	84.1	52.8	12.5	97.6
0.038	20,192	77.3	63.0	14.4	97.2
0.050	15,159	68.0	73.0	16.9	96.6
0.071	10,099	55.4	82.9	20.6	95.9
0.117	5,062	36.9	92.1	27.4	94.8

Definition of abbreviations: PPV – positive predictive value; NPV – negative predictive value.

* Risk cut-offs were set based on increasing deciles of percentile predicted risk.