Enhancing the Fever Workup Utilizing a Multi-Technique Modeling Approach to Diagnose Infections More Accurately

Adam M.A. Fadlalla,¹ Joseph F. Golob Jr.,² and Jeffrey A. Claridge²

Abstract

Background: Differentiation between infectious and non-infectious etiologies of the systemic inflammatory response syndrome (SIRS) in trauma patients remains elusive. We hypothesized that mathematical modeling in combination with computerized clinical decision support would assist with this differentiation. The purpose of this study was to determine the capability of various mathematical modeling techniques to predict infectious complications in critically ill trauma patients and compare the performance of these models with a standard fever workup practice (identifying infections on the basis of fever or leukocytosis).

Methods: An 18-mo retrospective database was created using information collected daily from critically ill trauma patients admitted to an academic surgical and trauma intensive care unit. Two hundred forty-three non-infected patient-days were chosen randomly to combine with the 243 infected-days, which created a modeling sample of 486 patient-days. Utilizing ten variables known to be associated with infectious complications, decision trees, neural networks, and logistic regression analysis models were created to predict the presence of urinary tract infections (UTIs), bacteremia, and respiratory tract infections (RTIs). The data sample was split into a 70% training set and a 30% testing set. Models were compared by calculating sensitivity, specificity, positive predictive value, negative predictive value, overall accuracy, and discrimination.

Results: Decision trees had the best modeling performance, with a sensitivity of 83%, an accuracy of 82%, and a discrimination of 0.91 for identifying infections. Both neural networks and decision trees outperformed logistic regression analysis. A second analysis was performed utilizing the same 243 infected days and only those non-infected patient-days associated with negative microbiologic cultures ($n = 236$). Decision trees again had the best modeling performance for infection identification, with a sensitivity of 79%, an accuracy of 83%, and a discrimination of 0.87.

Conclusion: The use of mathematical modeling techniques beyond logistic regression can improve the robustness and accuracy of predicting infections in critically ill trauma patients. Decision tree analysis appears to have the best potential to use in assisting physicians in differentiating infectious from non-infectious SIRS.

INFECTIONS ARE COMMON COMPLICATIONS in the surgical and trauma intensive care unit (STICU) [1]. Approximately two million nosocomial infections occur annually, and the cost of diagnosing and treating these infections exceeds \$4.5 billion each year [2, 3]. Despite being common, infections in the STICU can be difficult to diagnose. Common clinical signs such as fever and leukocytosis become less reliable in a surgical and trauma patient population because of their association with inflammation rather than infection [4–9].

We previously demonstrated that obtaining cultures of blood, urine, and respiratory secretions in critically ill trauma

patients was associated with fever ($\geq 38.5^{\circ}\text{C}$) or leukocytosis ($\geq 12.0 \times 10^3/\text{mm}^3$). However, neither fever nor leukocytosis was associated with a positive culture [10–12]. Similar results have been described by others for a variety of populations, including general surgery and burn patients [4, 8, 13–15].

In the future, we expect that computers with clinical decision support will assist physicians and improve the diagnostic efficiency of identifying infections in the STICU. To do this, one needs an accurate prospective medical informatics system and validated prediction rules. We have developed and validated the Surgical Intensive Care–Infection Registry

¹Department of Computer and Information Science, Cleveland State University, Cleveland, Ohio.

²Department of Surgery, MetroHealth Medical Center, Cleveland.

Presented at the Twenty-eighth Annual Meeting of the Surgical Infection Society, Hilton Head, South Carolina, May 7–9, 2008.

(SIC-IR)[®] as our informatics tool [16, 17]. This present study was performed to evaluate the potential ability of mathematical models to assist in predicting infections.

The SIC-IR is an electronic medical record system designed specifically for research as well as physician and administrative documentation in our two STICUs. The system collects more than 100 clinical variables daily on each STICU patient at the point of care. These variables include all the National Healthcare Safety Network data, as well as all the Joint Commission ICU core measures. The Registry was created specifically with the ability to incorporate advanced clinical decision support generated from robust mathematical models to assist clinicians with the diagnosis of infectious complications. We hypothesize that the addition of multiple clinical and demographic variables to computerized mathematical models will improve the accuracy and efficiency of identifying infections in critically ill trauma patients. Possible modeling techniques that can be utilized within SIC-IR include decision trees, artificial neural networks, and logistic regression analysis.

The purpose of this pilot study was twofold: (1) To define our STICU culture practice and the impact the fever workup (obtaining blood, urine, or respiratory secretion cultures on the basis of fever, leukocytosis, or both) has on this practice; and (2) to calculate the ability of various mathematical modeling techniques to assist beyond our current practice in the identification of infectious complications (urinary tract infections [UTIs], bacteremia, or respiratory tract infections [RTIs]) in critically ill trauma patients. The goal of this study was not to create an infection workup, but rather to determine the feasibility of using mathematical modeling techniques and to identify the technique that has the greatest potential to assist in the development of clinical decision support within the SIC-IR system.

Patients and Methods

Database creation and study sample selection

An 18-mo (January 2004–June 2005) retrospective database was created with consecutive critically ill trauma patients admitted to our institution's 27-bed STICU. To be included in the database, patients had to be at least 18 years old, could not have been admitted to the regular nursing ward or another hospital for more than 24 h prior to STICU admission, and to have stayed at least two days in the STICU. Data collected included patient demographics, injury mechanism (blunt vs. penetrating), and Injury Severity Score (ISS). Clinical data were collected daily during the patient's initial 14 hospital days and included maximum body temperature, maximum leukocyte count, antibiotic treatments, need for mechanical ventilation, presence of a central venous catheter, microbiologic culture results, and confirmed infectious complications. The U.S. Centers for Disease Control and Prevention definitions for UTI, bacteremia, and RTI were utilized. Bronchoalveolar lavage samples with $\geq 10^4$ bacteria/mL and urine cultures with $\geq 10^5$ colony-forming units/mL were considered positive. All positive blood cultures were analyzed individually for contaminating organisms such as a single bottle of coagulase-negative staphylococci or diphtheroid species, which were not considered a positive result.

The population used for this analysis was created by identifying all infected patient-days in the database. An in-

fecting patient-day was defined as the day when an infectious complication was confirmed by a positive microbiologic culture (blood, urine, or respiratory secretions). To identify non-infected patient-days, the remaining patient-days were refined by removing any day from patients with an identified infection. For example, if a patient was identified as having a UTI on STICU day three, all days before and after STICU day three were removed from the analysis. This resulted in a group of non-infected patient-days, which were defined as any day with negative microbiologic cultures, or days when the clinical suspicion of infection was so low that no cultures were obtained. This patient-day removal was done because we were unable to determine when an infected patient should be considered disease-free (i.e., on what day after antimicrobial treatment initiation is an infected patient considered non-infected?).

Evaluation of our STICU culture practice and the influence of the fever workup

These infected and non-infected patient-days were analyzed to determine our STICU practice of obtaining cultures of blood, urine, or respiratory secretions (tracheal aspirate or bronchoalveolar lavage) and to determine the influence the fever workup has on this practice. This was accomplished by calculating the relative risk that fever or leukocytosis would trigger a microbiologic culture.

A bivariable analysis was then performed to identify variables associated with obtaining a culture. The variables evaluated are listed in Table 1 and include those known to be associated with infectious complications or to have a direct impact on culture results [7, 18–25]. Any variable with a *p* value < 0.1 was placed in a backward stepwise logistic regression analysis to identify those that were associated independently with a culture. The odds ratios were then compared to determine the impact fever or leukocytosis has on triggering a culture workup.

TABLE 1. VARIABLES USED IN DETERMINATION OF OUR SURGICAL AND TRAUMA INTENSIVE CARE UNIT MICROBIOLOGIC CULTURE PRACTICE AND IN MATHEMATICAL MODELING OF INFECTION IDENTIFICATION

<i>Variable name</i>	<i>Variable type</i>
Intensive care unit stay (days)	Continuous
Age (years)	Continuous
Max. body temperature (°C)	Continuous
Max. leukocyte count ($10^3/\text{mL}^3$)	Continuous
Injury Severity Score	Continuous
Sex	Dichotomous
Mechanism of injury	Dichotomous
Presence of a central line	Dichotomous
Presence of mechanical ventilation	Dichotomous
Use of antibiotics the day prior to cultures	Dichotomous
Fever or leukocytosis	Dichotomous

These variables were collected daily on each critically ill trauma patient and placed in various mathematical models to predict infections (urinary tract, bacteremia, respiratory tract). All mathematical models utilized temperature and leukocyte count as continuous variables. Fever, leukocytosis, or both replaced maximum body temperature and maximum leukocyte count when evaluating our culture practice.

TABLE 2. DESCRIPTION OF MATHEMATICAL MODELS

	Decision tree	Artificial neural network	Logistic regression analysis
Goal	Provides algorithms for generating rules that partition data leading to prediction of target outcome variable	Computerized intelligent systems that simulate the inductive power of the human brain to abstract essential data needed to predict specific outcomes	Predicts probabilities of a binary outcome as a function of one or more independent variables
Types	<p><i>Classification trees</i>: decision trees that predict categorical variable outcomes</p> <p><i>Regression trees</i>: decision trees that predict continuous variable outcomes</p>	<p><i>Recurrent or non-recurrent</i>: how the layer of data processing elements (neurons) are organized and connected</p>	<p><i>Direct</i>: All independent variables enter the model at the same time</p> <p><i>Sequential</i>: Using a block chi-square method (likelihood ratio test), the full model with independents and covariates is evaluated. Then the model is run again with the block of independents that were dropped to determine the importance of the covariates</p> <p><i>Stepwise</i>: Model determines automatically which independent variables to add/drop</p>
Method of prediction	<ul style="list-style-type: none"> Variety of algorithms that generate rules (rules are stored in leaf nodes) <ul style="list-style-type: none"> CHAID (CHi-square Automatic Interaction Detection) [44] CART (Classification and Regression Trees) [45] QUEST [46] C5.0 [47] Iterative splitting of independent variables into discrete groups that are as different as possible Ideally, a split leads to a pure leaf node that contains only one of the outcomes being predicted 	<ul style="list-style-type: none"> Separates independent variables into interconnected network of simple units called neurons (connections) Neurons lead data through a simple input layer, any number of hidden layers, and an output layer Similar to a human brain, neural networks can “learn” from mistakes by modifying connections between neurons (e.g., delta learning rule) 	<ul style="list-style-type: none"> Evaluates a group of independent variables collectively termed x and their probability (p) of an outcome $\log it(p) = \log_2 \left\{ \frac{p}{1-p} \right\} = \alpha + x'b$ <p>where α is the intercept parameter and b is a vector of slope parameters</p>

TABLE 3. CRITERIA USED TO COMPARE THE MATHEMATICAL MODELS WITH EACH OTHER AND THE FEVER WORKUP^a

Model comparison criterion	Equation/method of calculation
Sensitivity	True positive/(true positive + false negative)
Specificity	True negative/(false positive + true negative)
Positive predictive value	True positive/(true positive + false positive)
Negative predictive value	True negative/(false negative + true negative)
Accuracy	(True positive + true negative)/(true positive + false positive + true negative + false negative)
Discrimination	Measured by the area under the receiver operating characteristic curve

^aObtaining cultures of blood, urine, and respiratory secretions because of fever, leukocytosis, or both.

Description of mathematical models and their statistical evaluation

Decision tree, artificial neural network, and logistic regression models were created to predict infectious complications in critically ill trauma patients during their initial 14 hospital days. Decision trees are tools for partitioning data recursively to assign cases to specific target classes such as infected or non-infected. Artificial neural networks were developed from engineering research and attempt to mimic the action of human neurons within computer software. These networks model inputs and outputs to identify patterns leading the “neurons” to a specific outcome. Logistic regression analysis is a more familiar modeling tool used in the medical literature. This tool determines if a group of clinical variables including covariables has a unique predictive relation to a specific outcome. Table 2 provides a more detailed description of these mathematical models.

For unbiased results, models should be developed on a different data set from the set on which they are tested [26–28]. For developing the models in this analysis, a random sample of non-infected patient-days was selected to match the number of infected days in a 1:1 ratio. This sample was then split 70% for a training set and 30% for a testing set. Decision trees, neural networks, and logistic regression models were trained using the training set to predict the presence of infection. These models were created utilizing the clinical and demo-

graphic variables listed in Table 1. The models were then tested on the testing set and compared with each other and with the fever workup by calculating sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), overall accuracy, and discrimination.

All models were designed using SPSS (SPSS Inc, Chicago, IL). The neural network was created using a multi-perceptron implementation from SPSS and the rules of the decision tree using the SPSS implementation of the C5.0 algorithm. The definitions of sensitivity, specificity, PPV, NPV, overall accuracy, discrimination, and H-L calibration statistic are listed in Table 3. All categorical variables are expressed as percentages and compared using the chi-square or Fisher exact test as appropriate. Continuous variables are expressed as the means $\pm$ the standard error of the mean and compared using the Student *t*-test after evaluation with the Levene test to determine if equal or unequal variances should be assumed. A *p* value of 0.05 was used to define statistical significance. This study was reviewed and approved by the MetroHealth Institutional Review Board.

Results

The retrospective database used for this analysis consisted of 510 patients with 3,839 patient-days of data. There were a total of 243 infected patient-days, which left 3,596 patient-days remaining. All days from patients who had an identified infection were excluded. This left 2,243 non-infected patient-days to combine with the 243 infected-days, which created the 2,486 patient-day study sample used in this analysis (Fig. 1). This study sample had a 10% prevalence of infection (bacteremia, UTI, RTI) and a total of 2% missing data for the ten variables evaluated. Table 4 lists the clinical and demographic data of the study cohort. The 243 infected-days consisted of 60

FIG. 1. Flow diagram of patient selection to evaluate our surgical and trauma intensive care unit culture practice and impact of fever workup. Infected patient-day = day an infectious complication was identified by positive culture (blood, urine, or respiratory secretions); non-infected patient-day: any day with negative cultures or when clinical suspicion of infection was so low that no cultures were obtained; fever workup = cultures of blood, urine, or respiratory secretions because of fever, leukocytosis, or both.

TABLE 4. CLINICAL AND DEMOGRAPHIC DATA OF STUDY COHORT (N = 2,486 PATIENT-DAYS)

Age (years)	48.8 $\pm$ 0.4
Daily max. temperature ($^{\circ}$ C)	37.9 $\pm$ 0.02
Daily max. leukocyte count ($\times 10^3$ /mL ³)	11.2 $\pm$ 0.1
Injury Severity Score	20.2 $\pm$ 0.2
Male (%)	74
Blunt trauma (%)	91
Patient-days with central venous catheter (%)	40
Patient-days with ventilator requirement (%)	55

Continuous variables are expressed as means $\pm$ standard error of the mean per patient-day. Dichotomous variables are expressed as percents per patient-day.

TABLE 5. 2×2 TABLE EVALUATING SURGICAL OR TRAUMA INTENSIVE CARE UNIT CULTURE PRACTICE

		Blood, urine, and/or respiratory secretion culture obtained	
		Yes	No
Fever And/Or Leukocytosis	Yes	378	873
	No	101	1134

Relative risk of fever and/or leukocytosis triggering a culture = 3.4 (95% CI = 3.0-4.5)

This table demonstrates that our STICU practices the fever workup (obtaining cultures of blood, urine, or respiratory secretions in the presence of fever, leukocytosis, or both) and obtains microbiologic cultures on the basis of fever ($\geq 38.5^{\circ}\text{C}$) or leukocytosis ($\geq 12.0 \times 10^3/\text{mm}^3$ of blood). Values are expressed in patient-days.

UTIs, 27 bacteremias, and 178 RTIs. Twenty-two infected patient-days had multiple infections identified.

Evaluation of our STICU culture practice and the influence of the fever workup

Four hundred seventy-nine patient-days (19%) were associated with the obtaining of at least one of the three cultures of interest: Blood, urine, or respiratory secretions. One thousand two hundred fifty-one patient-days (50%) were associated with either fever or leukocytosis. The relative risk of fever or leukocytosis triggering a culture was 3.4 (95% confidence interval [CI] 3.0, 4.5)(Table 5). Seventy-nine percent of the blood, urine, and respiratory secretion cultures were obtained on patient-days with fever or leukocytosis. These data support the idea that our STICU culture practice is influenced strongly by the fever workup.

Bivariable analysis using the variables listed in Table 1 identified ICU day ($p < 0.001$), age ($p = 0.002$), Injury Severity

TABLE 6. LOGISTIC REGRESSION ANALYSIS IDENTIFYING VARIABLES ASSOCIATED WITH OUR SURGERY OR TRAUMA INTENSIVE CARE UNIT CULTURE PRACTICE

Variable associated with obtaining a culture	Odds ratio	95% Confidence interval
Intensive care unit day	1.2	1.1, 1.2
Presence of central venous catheter	1.3	1.0, 1.6
Mechanical ventilation requirement	3.6	2.7, 4.7
Presence of fever or leukocytosis	3.8	2.9, 4.8

Presence of fever, leukocytosis, or both had the highest odds ratio, suggesting the importance of the fever workup (obtaining blood, urine, and respiratory secretion cultures because of fever or leukocytosis) in our STICU practice. The area under the receiver operating characteristic curve was 0.80, which suggests not all variables associated with our culture practice were recorded and analyzed.

FIG. 2. Flow diagram of patient selection to evaluate mathematical models. Sample was created by randomly choosing 243 non-infected patient-days to combine with the 243 infected patient-days. This sample was then split into a 70% model training set and a 30% testing set.

Score ($p < 0.001$), male sex ($p = 0.009$), presence of a central venous catheter ($p < 0.001$), requirement for mechanical ventilation ($p < 0.001$), and presence of fever, leukocytosis, or both ($p < 0.001$) as associated with obtaining a culture. These variables were placed in a backward stepwise logistic regression analysis, which identified ICU day, presence of a central venous catheter, requirement for mechanical ventilation, and presence of fever or leukocytosis as independent predictors of a culture (Table 6). The presence of fever or leukocytosis had the highest odds ratio of the variables remaining in the logistic regression analysis, suggesting the major impact of the fever workup on our STICU culture practice. Therefore, in the remainder of this analysis, the fever workup will be used as a surrogate marker for our STICU culture practice when evaluating the mathematical models.

Evaluation of mathematical models for predicating infections

A random 243 non-infected patient-days were chosen to combine with the 243 infected-days in a 1:1 ratio. These 486 patient-days were then split into a 70% model training set ($n = 345$) and a 30% testing set ($n = 141$)(Fig. 2). Table 7 compares the clinical and demographic variables in the training and testing sets, showing no significant differences. The results of model testing are displayed in Table 8.

Fever or leukocytosis alone had an accuracy of 66% with a sensitivity of 78% for identifying infectious complications. Both the neural network and the decision tree modeling techniques increased the accuracy of identifying infections compared with the fever workup; however, the neural network demonstrated a slight decrease in sensitivity, whereas the decision tree increased the sensitivity. Compared with the fever workup, logistic regression analysis had a decrease in both sensitivity and accuracy for identifying infections. All models showed an increase in discrimination, with areas under the receiver operating characteristic (ROC) curves between 0.81 and 0.91. The H-L calibration statistic was 58.28,

TABLE 7. COMPARISON OF TRAINING AND TESTING SET OF THE TEN MODELING VARIABLES

	Training set (N = 345)	Testing set (N = 141)	P value
Age (years)	47.5 ± 1.1 ^a	48.6 ± 1.7	0.600
Max daily temperature (°C)	38.3 ± 0.05	38.2 ± 0.07	0.812
Max daily leukocyte count (×10 ³ /mL ³)	11.2 ± 0.3	12.0 ± 0.4	0.065
Injury Severity Score	22.7 ± 0.5	21.8 ± 1.0	0.394
Intensive care unit day	6.2 ± 0.2	5.7 ± 0.3	0.137
Males (%)	71	69	0.583
Blunt injury (%)	95	91	0.134
Patient-days with central venous catheter (%)	50	50	0.920
Patient-days with ventilator requirement (%)	69	75	0.253
Use of antibiotics the day before cultures obtained (%)	65	65	0.971

^aMean ± standard error of the mean per patient-day or percent per patient-day.

158.24, and 194.23 for the logistic regression, decision tree, and neural network, respectively.

Each of the models listed in Table 8 was created using the definition of non-infected patient-day as any day with negative microbiologic cultures or on which the clinical suspicion of infection was so low that no cultures were obtained. Realizing the potential of missing infected patient-days in those patients without cultures, we performed an identical analysis using the same 243 infected patient-days and only the non-infected patient-days associated with negative microbiologic cultures (Fig. 3). This second analysis was done to remove some of the selection bias and potential false negatives that may be found in our initial models.

For this second study sample, a comparison of the training and testing set variables are listed in Table 9, and the model results are presented in Table 10. There was no significant difference between the training and testing set for any of the variables. Fever or leukocytosis was 47% accurate in identifying infections in patients having cultures. Mathematical modeling increased this accuracy to 83% in the decision tree model. The neural network and the decision tree analysis increased the sensitivity of identifying infections to 79% and 86%, respectively, but logistic regression lost sensitivity, to 66%. Decision tree analysis again had the best discrimination of all the models, with an area under the ROC curve of 0.87.

Discussion

Our analysis demonstrated that the presence of fever, leukocytosis, or both dominates our STICU trauma patient culture practice. In a population believed to be infected, as demonstrated by the physician obtaining a microbiologic culture, fever or leukocytosis alone identified 72% (sensitivity) of those with an infection. However, on patient-days on which

cultures were obtained and there was fever or leukocytosis, only 50% (PPV) of the cultures were positive. Mathematical modeling utilizing ten variables associated with infectious complications can improve the sensitivity, specificity, PPV, NPV, accuracy, and discrimination of diagnosing infections in a critically ill trauma population.

Statistical and mathematical models in the trauma and surgical critical care population have been utilized in the past. As performance comparisons between hospitals and physicians increase, models have been used to predict outcomes for both trauma (adult and pediatric) and critical care patients [29–34]. A common modeling technique used in the trauma and critical care literature continues to be logistic regression. This technique was used to create the Trauma and Injury Severity Score (TRISS), which calculates the probability of death on the basis of anatomic and physiological severity indicators [35]. More recently, other mathematical models have been gaining momentum in the medical community. For example, artificial neural networks have been used in mammography algorithms, discerning the prognosis of patients experiencing a myocardial infarction, predicting outcomes after liver transplantation, and assisting in the diagnosis of pulmonary emboli [36–39]. However, models other than logistic regression analysis have been slow to be adopted in the medical community because of their complexity and conflicting data when comparing performance directly using logistic regression analysis [29–31, 34].

The two additional modeling techniques we evaluated were artificial neural networks and decision tree analysis. Neural networks are best suited for situations that are unstructured and data-intensive, with high uncertainty, hidden relationships, and noise. We believe they are potentially useful for predicting infections in critically ill trauma patients. Predicting infections in such patients is data-intensive and

TABLE 8. ANALYSIS OF MATHEMATICAL MODELING COMPARED WITH FEVER WORKUP^a

	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)	Discrimination ^b
Fever workup	78	55	63	72	66	0.67
Neural network	76	78	79	76	77	0.84
Decision tree	83	80	79	85	82	0.91
Logistic regression	62	65	69	58	63	0.81

^aAll values are reported for the model testing set (n = 141).

Fever workup = cultures of blood, urine, or respiratory secretions because of fever, leukocytosis, or both.

^bDiscrimination: Area under the receiver operating characteristic curve.

PPV = positive predictive value; NPV = negative predictive value.

FIG. 3. Flow diagram of patient selection to evaluate mathematical models in patient-days with a microbiologic culture workup. Sample was created by identifying the 236 non-infected patient-days (negative culture) and combining them with the 243 infected patient-days. This sample was then split into a 70% model training set and a 30% testing set.

highly uncertain, and the data can be noisy. However, developing a good neural network model requires a substantial amount of time for trial and refinement, and interpreting the resulting model is a serious challenge. Decision trees have key merits, including their ability to handle missing data and the ease of explaining and deploying the resulting models. However, they are based on step (not smooth) functions, and hence may provide only coarse estimators for relationships.

Statistical modeling of infectious complications has been performed by Leibovici et al. in their development of TREAT, a clinical decision support system for antibiotic selection in inpatients with common bacterial infections [40–43]. This system was evaluated in a prospective study in three countries, and the investigators concluded that appropriate empirical antibiotic treatment increased in the intervention group whereas length of hospital stay, costs related to future resistance, and total antibiotic costs all were decreased [42]. Our models differ from TREAT in that we are using variables identified in the literature as being associated with infections and then our own data to generate the models instead of relying on probabilities identified by other authors. Use of our

own data will allow our system to continue to “learn” and adapt as new prospective data are gathered within the SIC-IR system.

There are several limitations of this study that must be addressed in the context of the overall objective of our analysis. Our goal was not to describe when an “infectious workup” should be initiated, but rather to determine if mathematical modeling using multiple variables associated with infections can improve the predictive power over the fever workup. The database used for this analysis was collected in a retrospective fashion and has the inherent weakness of this type of data. The database also lacked specific important variables that may help predict infections such as vital signs and hemodynamic support requirements. However, for our analysis, we believe the ten variables used were relevant to the goal of proving mathematical modeling may help identify infections in trauma patients. The current SIC-IR system offers prospective data collected at the point of care and includes multiple other variables associated with infection, which will improve the robustness of future models.

The assumption that patients in whom no cultures were obtained were not infected was addressed by performing our second modeling analysis looking only at patients from whom cultures were obtained. This second analysis also proved that modeling can increase the accuracy and discrimination without sacrificing sensitivity in a population believed to be clinically infected. Finally, all analyses were limited to the three infections most commonly evaluated in the fever workup: UTIs, bacteremia, and respiratory tract infections. This may introduce a low signal-to-noise ratio in the models, which results in variables that are unimportant becoming more significant. However, the recommended number of infections per variable in the model should be at least ten, creating the requirement for 100 infected-days. Our data had 243 infected-days, which yielded just over 20 infected-days per variable entered in the model.

Our STICU practice of obtaining microbiologic cultures in critically ill trauma patients is dominated by fever and leukocytosis, individually or together. This practice results in sensitivities between 72% and 80% and accuracies of 47–66% in identifying infectious complications. Mathematical models beyond logistic regression analysis improved sensitivity, specificity, PPV, NPV, accuracy, and discrimination of infection. Decision tree analysis performed best, with

TABLE 9. COMPARISON OF TRAINING AND TESTING SET OF TEN MODELING VARIABLES IN PATIENT-DAYS WITH MICROBIOLOGIC CULTURE WORKUP

	Training set (N = 340)	Testing set (N = 139)	P value
Age (years)	46.4 ± 1.0 ^a	45.8 ± 1.6	0.764
Max daily temperature (°C)	38.6 ± 0.04	38.6 ± 0.06	0.981
Max daily leukocyte count (×10 ³ /mL ³)	12.3 ± 0.3	11.5 ± 0.4	0.164
Injury Severity Score	23.6 ± 0.6	22.7 ± 1.0	0.416
Intensive care unit day	6.7 ± 1.2	6.8 ± 0.3	0.800
Males (%)	79	78	0.814
Blunt injury (%)	91	94	0.222
Patient-days with central venous catheter (%)	54	56	0.690
Patient-days with ventilator requirement (%)	83	82	0.747
Presence of antibiotics the day before cultures obtained (%)	80	77	0.507

^aMean ± standard error of the mean per patient-day.

TABLE 10. ANALYSIS OF MATHEMATICAL MODELING TECHNIQUES IN PATIENT-DAYS WHERE CULTURES WERE OBTAINED^a

	<i>Sensitivity (%)</i>	<i>Specificity (%)</i>	<i>PPV (%)</i>	<i>NPV (%)</i>	<i>Accuracy (%)</i>	<i>Discrimination^b</i>
Fever workup	72	18	50	38	47	0.50
Neural network	79	67	65	80	72	0.76
Decision tree	86	80	81	85	83	0.87
Logistic regression	66	58	58	66	62	0.74

^aAll values are reported for the testing set (n = 139).

Fever workup = cultures of blood, urine, or respiratory secretions because of fever, leukocytosis, or both.

^bDiscrimination = area under the receiver operating characteristic curve.

PPV = positive predictive value; NPV = Negative predictive value.

sensitivities between 83% and 86% and an accuracy of approximately 82% for identifying infections in critically ill trauma patients. Thus, to create an infection workup to assist with the differentiation between infectious and non-infectious etiologies of the SIRS response, decision trees should be considered for designing computerized clinical decision support within our SIC-IR.

Conclusion

The results of this study alone have not altered the way we utilize antibiotics in the ICU. However, these results add to our continued efforts to determine a more efficient method of diagnosing infection. This study demonstrates our failings in this endeavor. Furthermore, this study was performed to investigate which modeling technique might be best to incorporate behind the scenes to assist with clinical decision support.

There are several things that need to happen before we can use clinical decision support reliably to assist the clinician in improving the diagnosis of infection. First, we need more data, especially on patient-days that an infection is not suspected. Eventually, we would then conduct a trial of the clinical decision support in the hospital. The clinical decision support system would recommend a practice to the physician regarding whether to obtain certain cultures on the basis of the suspected probability of infection on a given day.

Acknowledgments

This work and JAC were supported in part by Grant Number 1KL2RR024990 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH), and the NIH Roadmap for Medical Research. Its contents are solely the responsibility of the authors and do not necessarily represent the official view of NCRR or NIH.

This work was also partially funded by a Junior Faculty Fellowship awarded to JAC by the Surgical Infection Society Foundation.

Author Disclosure Statement

No conflicting financial interests exist.

References

1. Edwards JR, Peterson KD, Andrus ML, et al. National Healthcare Safety Network (NHSN) Report: Data summary for 2006. *Am J Infect Control* 2007;35:290–301.
2. Hurr H, Hawley HB, Czachor JS, et al. APACHE II and ISS scores as predictors of nosocomial infections in trauma patients. *Am J Infect Control* 1999;27:79–83.
3. Jarvis WR. Selected aspects of the socioeconomic impact of nosocomial infections: Morbidity, mortality, cost, and prevention. *Infect Control Hosp Epidemiol* 1996;17:552–557.
4. Freischlag J, Busuttill RW. The value of postoperative fever evaluation. *Surgery* 1983;94:358–363.
5. Hoover L, Bochicchio GV, Napolitano LM, et al. Systemic inflammatory response syndrome and nosocomial infection in trauma. *J Trauma Inj Infect Crit Care* 2006;61:310–316.
6. Miller PR, Munn DD, Meredith JW, Chang MC. Systemic inflammatory response syndrome in the trauma intensive care unit: Who is infected? *J Trauma Inj Infect Crit Care* 1999;47:1004–1008.
7. Papias G, McLellan BA, El-Helou P, et al. Infection in hospitalized trauma patients: Incidence, risk factors, and complications. *J Trauma Inj Infect Crit Care* 1999;47:923–927.
8. Schey D, Salom EM, Papadia A, et al. Extensive fever workup produces low yield in determining infectious etiology. *Am J Obstet Gynecol* 2005;192:1729–1734.
9. Wallace WC, Cinat M, Gornick WB, et al. Nosocomial infections in the surgical intensive care unit: A difference between trauma and surgical patients. *Am Surg* 1999;65:987–990.
10. Claridge JA, Sando MJ, Golob JF, et al. The “fever workup” and respiratory culture practice in critically ill trauma patients. *J Crit Care*. 2009.
11. Golob JF, Claridge JA, Fadlalla A, et al. Fever and leukocytosis in critically ill trauma patients: It’s not the blood. *Am Surg*. 2009;75:405–410.
12. Golob JF, Claridge JA, Sando MJ, et al. Fever and leukocytosis in critically ill trauma patients: It’s not the urine. *Surg Infect* 2008;9:49–56.
13. Wagenlehner FM, Loibl E, Vogel H, et al. Incidence of nosocomial urinary tract infections on a surgical intensive care unit and implications for management. *Int J Antimicrob Agents* 2006;28(Suppl 1):S86–S90.
14. Murray CK, Hoffmaster RM, Schmit DR, et al. Evaluation of white blood cell count, neutrophil percentage, elevated temperature as predictors of bloodstream infection in burn patients. *Arch Surg* 2007;142:639–642.
15. Theuer CP, Bongard FS, Klein SR. Are blood cultures effective in the evaluation of fever in perioperative patients? *Am J Surg* 1991;162:615–618.
16. Fadlalla AMA, Golob JF, Claridge JA. The Surgical Intensive Care–Infection Registry (SIC-IR): A research registry with daily clinical support capabilities. *Am J Med Qual* 2009;24:29–34.
17. Golob JF, Fadlalla MA, Kan JA, et al. Validation of SIC-IR[®]: A medical informatics system for intensive care unit research, quality of care improvement, and daily patient care. *J Am Coll Surg* 2008;207:164–173.
18. Charalambous C, Swoboda SM, Dick J, et al. Risk factors and clinical impact of central line infections in the surgical intensive care unit. *Arch Surg* 1998;133:1241–1246.

19. Pawar M, Kapoor P, Sharma J, et al. Central venous catheter-related blood stream infections: Incidence, risk factors, outcome, and associated pathogens. *J Cardiothorac Vasc Anesth* 2004;18:304–308.
20. Angstwurm MW, Gaertner R, Schopohl J. Outcome in elderly patients with severe infection is influenced by sex hormones but not gender. *Crit Care Med* 2005;33:2786–2793.
21. Bochicchio GV, Joshi M, Bochicchio K, et al. Incidence and impact of risk factors in critically ill trauma patients. *World J Surg* 2006;30:114–118.
22. Bochicchio GV, Joshi M, Knorr KM, et al. Impact of nosocomial infections in trauma: Does age make a difference? *J Trauma Inj Infect Crit Care* 2001;50:612–617.
23. Crabtree TD, Pelletier SJ, Gleason TG, et al. Gender-dependent differences in outcome after the treatment of infection in hospitalized patients. *JAMA* 1999;282:2143–2148.
24. Jamulitrat S, Narong MN, Thongpiyapoom S. Trauma severity scoring systems as predictors of nosocomial infection. *Infect Control Hosp Epidemiol* 2002;23:268–273.
25. Coopersmith C, Kollef M. ACS surgery online. In Dale D, Federman D, eds. *Critical Care, Postoperative and Ventilator-Associated Pneumonia*. WebMD Inc., 2000.
26. Bleeker SE, Moll HA, Steyerberg EW, et al. External validation is necessary in prediction research: A clinical example. *J Clin Epidemiol* 2003;56:826–832.
27. Terrin N, Schmid CH, Griffith JL, et al. External validity of predictive models: A comparison of logistic regression, classification trees, and neural networks. *J Clin Epidemiol* 2003;56:721–729.
28. Vergouwe Y, Steyerberg EW, Eijkemans JC, et al. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol* 2005;58:475–483.
29. Becalick DC, Coats TJ. Comparison of artificial intelligence techniques with UKTRISS for estimating probability of survival after trauma. *UK Trauma and Injury Severity Score. J Trauma Inj Infect Crit Care* 2001;51:123–133.
30. Clermont G, Angus DC, DiRusso SM, et al. Predicting hospital mortality for patients in the intensive care unit: A comparison of artificial neural networks with logistic regression models. *Crit Care Med* 2001;29:291–296.
31. DiRusso SM, Chahine AA, Sullivan T, et al. Development of a model for prediction of survival in pediatric trauma patients: Comparison of artificial neural networks and logistic regression. *J Pediatr Surg* 2002;37:1098–1104.
32. DiRusso SM, Sullivan T, Holly C, et al. An artificial neural network as a model for prediction of survival in trauma patients: Validation for a regional trauma area. *J Trauma Inj Infect Crit Care* 2000;49:212–220.
33. Gabbe BJ, Cameron PA, Wolfe R, et al. Prehospital prediction of intensive care unit stay and mortality in blunt trauma patients. *J Trauma Inj Infect Crit Care* 2005;59:458–465.
34. Wolfe R, McKenzie DP, Black J, et al. Models developed by three techniques did not achieve acceptable prediction of binary trauma outcomes. *J Clin Epidemiol* 2006;59:26–35.
35. Boyd CR, Tolson MA, Copes WS. Evaluating trauma care: The TRISS method. *J Trauma Inj Infect Crit Care* 1988;27:370–378.
36. Baxt WG. Use of an artificial neural network for the diagnosis of myocardial infarction. *Ann Intern Med* 1991;115:843–848.
37. Doyle HR, Dvorchik I, Mitchell S, et al. Predicting outcomes after liver transplantation: A connectionist approach. *Ann Surg* 1994;219:408–415.
38. Tourassi GD, Floyd CE, Sostman HD, et al. Artificial neural network for diagnosis of acute pulmonary embolism: Effect of case and observer selection. *Radiology* 1995;194:889–893.
39. Vyborny CJ, Giger ML. Computer vision and artificial intelligence in mammography. *AJR Am J Roentgenol* 1994;162:699–708.
40. Leibovici L, Gitelman V, Yehezkeili Y, et al. Improving empirical antibiotic treatment: Prospective, nonintervention testing of a decision support system. *J Intern Med* 1997;242:395–400.
41. Leibovici L, Paul M, Nielsen AD, et al. The TREAT project: Decision support and prediction using causal probabilistic networks. *J Antimicrob Chemother* 2007;30(Suppl 1):S93–S102.
42. Paul M, Andreassen S, Tacconelli E, et al. Improving empirical antibiotic treatment using TREAT, a computerized decision support system: Cluster randomized trial. *J Antimicrob Chemother* 2006;58:1238–1245.
43. Paul M, Nielsen AD, Goldberg E, et al. Prediction of specific pathogens in patients with sepsis: Evaluation of TREAT, a computerized decision support system. *J Antimicrob Chemother* 2007;59:1204–1207.
44. Kass G. An exploratory technique for investigating large quantities of categorical data. *Appl Stat* 1980;29:119–127.
45. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and Regression Trees*. New York: Chapman and Hall, 1984.
46. Loh WY, Shih YS. Split selection methods for classification trees. *Stat Sin* 1997;7:815–840.
47. Quinlan JR. *C4.5: Programs for Machine Learning*. San Francisco: Morgan Kaufmann Publishers, 1993.

Address correspondence to:
 Dr. Jeffrey A. Claridge
 MetroHealth Medical Center
 2500 MetroHealth Dr.
 Cleveland, OH 44109

E-mail: jclaridge@metrohealth.org