Towards precision medicine: Accurate predictive modeling of infectious complications in combat casualties

Authors

Christopher J Dente MD$^{1,2,5}$ cdente@emory.edu
Matthew Bradley MD$^{3,4,5}$ matthew.j.bradley22.mil@mail.mil
Seth Schobel PhD$^{3,5}$ seth.schobel-mchugh.ctr@usuhs.edu
Beverly Gaucher PhD$^{3,5}$ beverly.gaucher.ctr@usuhs.edu
Timothy Buchman MD PhD$^{1,5}$ tbuchma@emory.edu
Allan D. Kirk MD, PhD$^{5,6}$ allan.kirk@duke.edu
Eric Elster MD$^{3,4,5}$ eric.elster@usuhs.edu

Affiliations

1. Emory University, Atlanta GA.
2. Grady Memorial Hospital, Atlanta GA.
3. Uniformed Services University of the Health Sciences, Bethesda, MD.
4. Walter Reed National Military Medical Center, Bethesda, MD.
5. Surgical Critical Care Initiative (SC2i), Bethesda, MD.
6. Duke University, Durham, NC.
Short Title: Predictive modeling of infectious complications

Corresponding Author and Address for Reprints

Christopher J Dente MD FACS
Professor of Surgery
69 Jesse Hill Jr. Dr SE, Suite 311
Dept of Surgery, Emory University
Atlanta, Ga 30303

No conflicts of interest to report

Funding: Research activities leading to the development of this abstract were funded by the Department of Defense’s Defense Health Program – Joint Program Committee 6 / Combat Casualty Care (USUHS HT9404-13-1-0032 and HU0001-15-2-0001).

Paper to be presented as an Oral Presentation at the 30th Annual Meeting of the Eastern Association for the Surgery of Trauma, January 10-14, 2017, Hollywood, Fl.
ABSTRACT

Background: The biomarker profile of trauma patients may allow for the creation of models to assist bedside decision making & prediction of complications. We sought to determine the utility of modeling in the prediction of bacteremia & pneumonia in combat casualties.

Methods: This is a prospective, observational trial of patients with complex wounds treated at Walter Reed National Military Medical Center (2007-2012). Tissue, serum and wound effluent samples were collected during operative interventions until wound closure. Clinical, biomarker & outcome data were used in machine learning algorithms to develop models predicting bacteremia or pneumonia. Modeling was performed on the first operative washout to maximize predictive benefit. Variable selection of dataset variables was performed and the best fitting Bayesian belief network (BBN), using Bayesian information criterion (BIC), was selected for predictive modeling. Random forest was performed using variables from BBN step. Model performance was evaluated using area under the receiver operating characteristic curve (AUC) analysis.

Results: 73 patients (mean age 23, mean Injury Severity Score 25) were enrolled. Patients required a median of 3 (2-13) operations. The incidence of bacteremia & pneumonia was 22% & 12%, respectively. Best fitting variable selected BBNs were max-min parents & children (MMPC) for both bacteremia (BIC-24948) and pneumonia (BIC-17886). Full variable & MMPC random forest models AUC were 0.721 & 0.834 respectively for bacteremia and 0.809 & 0.856 respectively for pneumonia.
Conclusions: We identified a profile predictive of bacteremia and pneumonia in combat casualties. This has important clinical implications and should be validated in the civilian trauma population. This and similar tools will allow for increasing precision in the management of critically ill and injured patients.

Level of Evidence: III (Prognostic)

Keywords: Prediction, pneumonia, bacteremia, clinical decision support
BACKGROUND

Nosocomial infections are common occurrences in critically ill patients. In fact, patients requiring intensive care unit (ICU) level of care have a three to five fold increase in these morbid complications.\(^1\) Also, they remain the leading cause of late death after traumatic injury.\(^2\) Two of the most common complications that inflict critically ill and injured patients are pneumonia and blood stream infections. At least 25\% of infectious complications in the modern ICU are thought to be pulmonary in origin and up to 10\% of ICU patients experience a blood stream infection, many of which are related to indwelling catheters.\(^3\) Early diagnosis and treatment of these infectious complications, most notably pneumonia has been associated with improved outcomes in the modern ICU.\(^1\)

While much of the focus of the late care of the ICU patient revolves around the diagnosis and management of infections, less work has been done around prediction and risk stratification. While preventative strategies and guidelines are now widely published, much of the care of the patient who develops a nosocomial infection remains reactive. The predictive models that exist generally use serial measurements of common clinical data that compile into a scoring system. Tools like the APACHE score and SOFA score are relatively resource intensive to collect and provide gross estimates of the patient’s disease severity and prediction of in-hospital survival. Tools like NSQIP and the DS3 complication calculator are more usable on the individual patient level but also focus on common clinical data. More complex and individualized tools that use a combination of real time data collection and individualized physiologic information in the form of molecular biomarkers may be more precise.\(^4\) Indeed, having precise, easy to use and highly
individualized tools may allow a bedside clinician to identify the patients at highest risk for a variety of infectious complications and therefore could allow for more proactive and directed preventative strategies. Indeed, recent emphasis on precision medicine and a recent Institute of Medicine report on the current rate of diagnostic error suggest that there is a great need to improve the timeliness and accuracy of our predictive and diagnostic methods in our ICU patients.\textsuperscript{4}

The care of causalities in recent military conflicts has led to an interest in the creation of data and biomarker based clinical decision support tools (CDST). Several such tools are in various stages of development and use Bayesian Belief Networks (BBNs) to create highly accurate models to predict clinical outcomes.\textsuperscript{5,6} A military-civilian collaborative called the Surgical Critical Care Initiative (SC2i) has partnered military and civilian centers with a goal to further this area of research.

Because of the need to better understand the risk of developing common nosocomial infections in our ICU populations, we sought to create a predictive model for blood stream infection and pneumonia in a critically injured group of service people using the advanced statistical techniques described above. We hypothesized that a model could be created that would accurately predict these complications and could be used as a basis to create a novel CDST for clinicians caring for these patients.
METHODS

Patients with complex wounds cared for at Walter Reed National Military Medical Center (WRNMMC) had data collected prospectively in this observational study. This patient cohort had been previously used in the creation of a clinical decision support tool to enhance the accuracy of wound closure and provided a well-described and accurate data set and represented a highly regimented treatment algorithm. It also provided a combination of global clinical data, global biomarker data and biomarker data associated with the chronic inflammatory process (i.e. open wound) for each patient and was thus felt to provide an excellent physiologic data set for prediction of infections. This study was approved by the institutional review board at the primary institution under the auspices of the SC2i collaborative. Tissue, serum and wound effluent samples were collected at all relevant operative interventions from time of consent until wound closure. All wounds were managed via the Joint Theater Trauma System Clinical Practice Guideline, a highly regimented and well established protocol for wound management and all were managed with negative pressure dressings allowing molecular assessment of local wound dynamics. Serum measurements allowed for molecular assessment of total body dynamics. At each of the time points both clinical and biomarker data were collected. Clinical data points included: gender; age; date, location and mechanism of injury; requirement for transfusion and total number of blood products; injury severity score (ISS) and Acute Physiology and Chronic Health Evaluation II (APACHE II) score; wound surface area and depth; associated injuries; type of, timing and success of wound closure; Glasgow Coma Scale (GCS) score; presence and severity of traumatic brain injury; ICU and hospital length of stay; ventilator days; number of wound debridements; development of nosocomial infections and disposition from the
hospital. The biomarker data included Luminex proteomic, quantitative PCR (QPCR) transcriptomic, and quantitative bacteriology data. This data was gathered on both serum and wound effluent samples for QPCR and Luminex, whereas quantitative bacteriology assessments were conducted on wound tissue and effluent samples. To extract the most predictive and clinical value and the earliest possible diagnosis and risk prediction of onset of bacteremia and pneumonia in our patient cohort we created a subset of the dataset with just the first available time point. Because data collection began after transfer of the patient to Walter Reed in most cases, data from the immediate post injury phase of treatment was not available.

Techniques of blood collection as well as serum and wound inflammatory biomarker analysis have been published elsewhere. In brief, blood was collected, fractionated immediately using a centrifuge and plasma supernatant was flash frozen in liquid nitrogen and stored at -70°C. Serum was then analyzed using a beadlyte® Human 22-Plex Multi-Cytokine Detection System on the Luminex® 100 IS xMAP Bead Array Platform (Millipore Corp., Ontario, Canada). Twenty-two cytokines were quantified in pg/mL according to manufacturer’s instructions. Effluent samples from negative pressure containers were handled similarly.

For this study, we defined bacteremia as the presence of bacteria in the blood at any point during the study period that was treated with antibiotics. The number of positive cultures was not used to ascertain clinical importance of the event, rather the physician determination to provide the patient with a course of therapy. Similarly pneumonia was defined as a confirmed lung infection diagnosed through chest radiographic examination showing infiltrates, cavitation, pleural effusion or consolidation in addition to isolation of a pathogen from quantitated respiratory
culture \((10^5\text{ organisms by bronchial wash, }10^4\text{ by bronchial brush})\). Both clinical end points were determined through chart reviews of enrolled patients.

**Variable Selection and Statistical Modeling**

**Bayesian Variable Selection** – To perform variable selection on the entire set of serum and effluent Luminex variables as well as available clinical variables we used the “bnlearn” R package. We utilized several supervised machine learning algorithms to search the input dataset for the nodes of Bayesian networks to construct a reduced variable set that best represents the underlying distribution of all variables with respect to our infectious complication outcomes. We used a feature selection algorithm to choose the reduced variable set. These algorithms included the inter.iamb, fast.iamb, iamb, gs, mmpc, and si.hiton.pc algorithms. While all algorithms were examined for testing we primarily selected either the Maximum Minimum Parents Children (mmpc) or the semi-interleaved Hiton-PC (si.hiton.pc) algorithm to choose the nodes of the corresponding Bayesian network as our reduced variable set. Once each algorithm selected variables for inclusion, a Bayesian network was constructed for illustration, the nodes of which are the reduced variable set.

**Random Forest Modeling** – We constructed a random forest model using the full set of variables pulled from the raw data as our baseline. To handle process samples with missing data, we used the R packages rfImpute(). We plotted the total, positive class and negative class out-of-bag (OOB) error estimates of the model and calculated the Accuracy and Kappa scores. This was done with the “randomForest” R package. This full set of variables was the same full set from which we select variables in the previous step. Next we constructed a random forest model with
the Bayesian network-selected variables pulled from the raw data and assessed the random forest performance with OOB error plots, Accuracy and Kappa scores. We choose the model with the smallest OOB errors and the highest Accuracy and Kappa scores. Both random forest models were constructed using 10001 classification and regression trees and square root of p variables randomly sampled as candidates at each split, where p is the number of variables in the model. Once these two models were produced we compared the shape of their Receiver Operator Characteristic Curves (ROC) and respective Areas Under Curve (AUC) and sensitivities. We examined the confusion matrices and further assessed model performance using Vickers and Elkins’ Decision Curve Analysis (DCA). We plotted the decision curves of both the full variable random forest model and the reduced variable random forest model. DCA was used to assess the net benefit of the null model—or the “treat none” model—and the “treat all” model against our full variable set model and our reduced variable set model.

RESULTS

Patient and wound characteristics demonstrate a highly injured military population at risk for infectious complications.

73 patients were enrolled. Demographic information has been published elsewhere. Patients required a median of 3 (2-13) operations subsequent to enrollment. The patients had an average Injury Severity Score (ISS) of 24.9 (8-59), received 58.9 (0-519) units of blood products, with 32 of 73 having experienced injuries to named vascular structures. The average hospital and ICU
The dataset included 116 wounds and 399 data collection time points. Although an individual patient may have more than one wound tracked in the study, all systemic variables were utilized for modeling purposes only once. All modeling results were generated using the first available time point of data, a median of 5 days (3-13). Models were also generated using systemic and clinical markers on a per patient basis. Both a full data set model and a reduced variable model were created as is described below.

A combination of clinical and biomarker variables can identify patients at risk for bacteremia and pneumonia with great accuracy.

Bacteremia – The statistical modeling results described below demonstrate that a combination of clinical and biomarker variables can be used to identify trauma patients at risk for bacteremia. The variable selection process resulted in choosing the nodes of several Bayesian networks. Several methods were used to select variables of interest; those sets were then evaluated based on their ROC, AUC, Accuracy, Kappa, sensitivity, and specificity using random forest models. We then selected the model that displayed the most robust predictive ability. The Bayesian network with the best properties as described above included the following variables: Blood Bethesda, RBC (Red Blood Cells) Bethesda (both measures of volume of blood product received at WRNMMC), Critical Colonization (presence of greater than $10^6$ colony forming units per gram of tissue or per mcl of wound effluent), serum IL2R (soluble Interleukin 2 receptor) and serum MIG (Monokine induced by gamma interferon) (Figure 1). The random forest models
developed herein were demonstrably predictive models capable of identifying patients with or at risk for developing bacteremia. The evaluation of the random forest modeling via ROC/AUC analyses shows the full variable model with an AUC of 0.721, and show the reduced variable model with an AUC of 0.834 (Figure 2). These results indicate that the variable reduced model was superior to the full variable model. Furthermore, the sensitivity of the reduced variable model was .500, whereas the specificity was .912. This demonstrates that 50% of the bacteremia positive cases were predicted with this model. The DCA was performed to evaluate the net benefit to the patient of using the models developed here in comparison to “treat none” and “treat-all” paradigms. The DCA curves also show a measurable net benefit of both the full and reduced variable models, however the reduced variable model at most thresholds showed greater net benefit in identifying or predicting risk of bacteremia (Figure 3).

Pneumonia – The same statistical procedure was used to construct and validate a clinical and biomarker driven model capable of identifying trauma patients destined for pneumonia. The Bayesian network with the best properties as described above included the following variables: serum IL7, ISS (Injury Severity Score) Head, ISS Chest, and Critical Colonization (Figure 4). The results of random forest modeling and ROC/AUC analyses show the full variable model with an AUC of 0.809, and show the reduced variable model with an AUC of 0.856 (Figure 5). Similar to the bacteremia models, these results also indicated that the reduced variable pneumonia model performed better than the full variable model. The sensitivity of the reduced variable model was .556, whereas the specificity was .953. This suggests that greater than 55% of the pneumonia positive cases would be anticipated by this model. Again, we performed DCA to evaluate the net benefit to the patient of using the pneumonia models developed above. The
DCA curves also show a measurable net benefit of both the full and reduced variable models; however again, as was the case for bacteremia, the reduced variable model showed greater net potential benefit to trauma patients at most thresholds in identifying pneumonia or pneumonia risk (Figure 6).

**DISCUSSION**

Assisting bedside clinicians in timely decision making and reducing errors are important in many settings, particularly in the surgical ICU. Much of our ICU care, however, remains reactive and supportive. Therefore, there has been increasing interest in both prediction and early, precise diagnosis of nosocomial complications in an effort to shorten hospital stays and improve overall outcomes. The concept of “clinical decision support” is not a new one. These support systems were proposed more than fifty years ago by numerous authors. Many systems used automated measurements and rules as a potential “substitute” for judgment in critically ill patients.\(^\text{19-22}\) Indeed, clinical decision support tools have taken many forms over the years and have been reported to improve safety and quality of care in both intensive care unit and outpatient settings.\(^\text{23,24}\) Philosophically, it is now recognized that many of these tools are probably best used to support rather than replace clinical judgment, and newer tools were designed with this philosophy in mind.\(^\text{25-27}\) Indeed, the complexity of decisions made in this environment likely justifies the retention of human judgment rather than a fully automated system.\(^\text{28,29}\)
The use of biomarkers as predictive markers for nosocomial complications is increasingly more common. Indeed, procalcitonin has been used by multiple groups and rising levels have been shown to correlate with the development of sepsis in post-operative ICU patients. Many novel biomarkers are being studied, some of which may become useful predictive tools either alone or in combination with other clinical or biomarker data. The complications seen in casualties wounded during recent military conflicts have stimulated great interest in the creation of these types of tools to guide care and predict outcome. Indeed, a Bayesian Belief Network tool was created in a set of war victims that accurately predicted the occurrence of hospital acquired infection, ICU length of stay and the success of wound healing. Further refinement led to a tool that used a combination of clinical data with systemic and wound effluent biomarker measurements to accurately predict the success of attempted wound closure in a set of 73 military patients. This allowed military surgeons to make more precise decisions on the timing of closure in these patients. More accurate decisions have the potential to both shorten time to closure in marginal appearing wounds that have a high predicted rate of success while at the same time avoiding the morbidity of a failed closure attempt and the subsequent need for further operative debridement and longer lengths of stay.

The current study looks to extend this thought process by using data from the same patients to predict the occurrence of two common nosocomial infectious complications: bacteremia and pneumonia. Both of these entities are associated with significant patient morbidity and cost to the healthcare system. To some degree, the prediction of these complications is currently possible using a variety of tools. The CPIS tool, for example, uses six clinical variables to predict occurrence and outcome of ventilator associated pneumonia but was shown to have the
poorest AUC of the predictive models tested in a recent meta-analysis.\textsuperscript{34} The tools in this study, even with their limitations summarized below, seem to be much more accurate. It is felt that this and similar tools, could allow for physicians at the bedside to be less reactive in patients at high risk for these complications. Resources could be allocated precisely for targeted pulmonary strategies and even more aggressive attempts to remove invasive lines and catheters could be undertaken if a patient could be designated as high risk for these occurrences. Resource utilization is an increasingly important reality in our healthcare environment and preventative strategies targeted at the highest risk patients may provide increased efficiencies.

Interestingly, the reduced variable data sets suggested that a variety of factors seem to be of special important in the development of these infections. For example, anatomic degree of injury, in the form of Chest and Head AIS were important in the subsequent development of pneumonia. This is not surprising. However, additional predictive information was obtained from the serum biomarker IL-7 which is a cytokine important in the differentiation of B, T and NK cells. Conversely, IL2R, the level of which played an important role in the prediction of subsequent bacteremia acts as an inhibitor of the resolution of the T cell response to infection and MIG is a chemokine involved in the recruitment of T cells to sites of infection. This work thus suggests that adding specific biomarker information from a patient’s own unique response to critical illness may be of benefit in predicting outcomes.

Our statistical analysis technique involved two main steps: variable reduction and binary classification. The strengths of variable selection is that they are designed to search for a smaller dimension set of variables that seek to represent the underlying distribution of the full set of
variables. Since our datasets are relatively small, computational time was not a consideration. Since the variable selection was based on a better representation of the underlying distribution of the full variables set, in theory, they should be more generalizable and less susceptible to over fitting.

For our binary classification algorithm, we choose random forest.\textsuperscript{13} The two chief draws of the random forest are (1) it does not require the data to be either normally distributed or transformed and (2) the algorithm requires little tuning, which is advantageous when updating data sets, and its numerical process includes cross validation precluding the need for post model-building cross validation.\textsuperscript{13-15} Comparisons of the variable selected models to the full variable models showed better performance in the former. This is a key strength of these methods as over-parameterization frequently leads to model underperformance. In the variable selected models, the ROC curves and their respective AUCs showed the models have good predictive ability. Similarly these models have higher Accuracy and Kappa statistics than the full variable models. The reduced variable set models have higher Accuracy statistics than their no information rates. One method of evaluating modeling performance is the Out-of-Bag (OOB) error analysis.\textsuperscript{13} OOB is a way of measuring prediction error in random forest and other machine learning techniques that rely on bootstrapping to sub-sample training data.\textsuperscript{13} The OOB error analysis showed the variable selected models greatly improved the OOB error (predictive performance) for the positive class (pneumonia or bacteremia). Both performed well demonstrating that the random forest models picked up true positive signals, our ultimate goal. Furthermore the Decision Curve Analysis (DCA) showed the variable selected models outperformed the full variable models.\textsuperscript{16,17} This was also supported in the respective confusion matrices showing various model
performance metrics. DCA is used to compare various predictive and diagnostic paradigms in terms of net-benefit to the patient. A typical DCA analysis will compare the null model, treat no one, to various alternative models, such as “treat-all” or treat according to the guidance of models built on biomarker predictors. DCA analysis can be interpreted as showing positive net-benefit to the patient if the decision curve for a particular model is above the null model (x axis), and to the right of the “treat-all” model. Net-benefit is defined mathematically as a summation of model performance (for instance propensity to predict false positive or false negatives) over a series of predictive threshold cutoffs and the respective sensitivity/specificity at those thresholds. The threshold cutoffs could be thought of as the point at which a decision to treat would be made given the relative harms and benefits of treating given the uncertainty of the prediction at that threshold. This analysis should demonstrate the threshold cutoffs where the predictive models are most useful to the patient.

In total, we have shown that it is possible to anticipate the occurrence of common nosocomial infections in a cohort of injured warriors with great accuracy. This particular model uses a combination of clinical and biomarker data within a Bayesian belief network. It is hoped that many of the complex decisions that clinicians make in the critical care arena may, in the future, be supported by mathematical models that will allow for more precise and accurate prognostication. Moving forward this and other models will need to be validated and potentially modified for use in a civilian setting. Indeed the goal of our current military/civilian SC2i collaborative is in the creation, validation and promulgation of such tools. This particular model could give clinicians early and profound insight into the care of critically ill and injured patients in their care.
There are certainly limitations to the current tool. First, it is based on the care of a small (albeit inclusive) subset of wounded combat casualties. This patient population is, in many ways, less heterogenic and certainly healthier than the vast number of patients cared for in a civilian trauma center. Also, the data generated by this patient population was the result of a highly regimented and homogenous post injury management protocol in a military hospital. Overall, information from this group of patients and data derived from their management may not be widely generalizable especially in a civilian hospital. Further study to validate or modify this tool will be necessary before it can be recommended for wide use. Second, this is a modeling done retrospectively on previous collected data. Third, we did not have accurate information on the placement of central venous access in this patient population and thus it was not included in the modelling. Prospective validation of these tools will include data on central line usage. Fourth, we did not use serial measurements to any predictive benefit. This may be possible in future models. Finally, all complex mathematical modeling has the potential for overfitting. The reason we do not feel that overfitting is likely here is described above but certainly any tool with such high ROC curves has to come with the concern for overfit. As already mentioned, prospective validation will be an important next step.

In conclusion, we have been able to very accurately identify a profile predictive of two common nosocomial infections in a group of combat casualties. This tool, and tools like it, should allow for greater precision in the care of critically ill and injured patients in both civilian and military settings.
REFERENCES


FIGURE LEGENDS

**Figure 1:** Bayesian network representation representing conditionally dependent relationships between the bacteremia outcome and mmpc-selected variables.

**Figure 2:** ROC plot and AUC analysis of random forest model with mmpc-selected variables.

**Figure 3:** DCA curves comparing full variable, mmpc-selected, and treat-all models showing net benefit to the patient to use these various models to guide treatment. The full and mmpc models show a measurable net benefit over the treat-all model.

**Figure 4:** Bayesian network representation representing conditionally dependent relationships between the pneumonia outcome and inter.iamb-selected variables.

**Figure 5:** ROC plot and AUC analysis of random forest model with inter.iamb-selected variables.

**Figure 6:** DCA curves comparing full variable, inter.iamb-selected, and treat-all models showing net benefit to the patient to use these various models to guide treatment. The full and inter.iamb models show a measurable net benefit over the treat-all model.
Figure 1

Bayesian Belief Network

[Diagram showing a network with nodes labeled 'outcome', 'Blood_Retinoid', 'CC', 'SeZn_IL3R', and 'SeZn_MK4', connecting with arrows indicating relationships.]
Figure 2

WoundVAC (Bacteremia Positive) ROC plot
Reduced Biomarker Variable Set—MMPC

Random Forest Markov Blanket Variables AUC = 0.83443
Figure 3

Decision Curve Analysis Bacteremia
Full and Reduced Dimension Models

Net benefit

Threshold probability

Proteomic Reduced Model
Figure 4

Bayesian Belief Network

outcome

ISO_head
ISO_cheal
Ser2s_IL7

CC
Figure 5

WoundVAC (Pneumonia Positive) ROC plot
Reduced Biomarker Variable Set---inter.iamb

Random Forest Markov Blanket Variables AUC = 0.8559
Figure 6

Decison Curve Analysis Pneumonia
Full and Reduced Dimension Models

Net benefit

Threshold probability