

Development of a Bayesian model to estimate health care outcomes in the severely wounded

Alexander Stojadinovic¹
John Eberhardt²
Trevor S Brown³
Jason S Hawksworth⁴
Frederick Gage³
Douglas K Tadaki³
Jonathan A Forsberg⁵
Thomas A Davis³
Benjamin K Potter⁵
James R Dunne⁶
E A Elster³

¹Combat Wound Initiative Program,
⁴Department of Surgery, Walter Reed Army Medical Center, Washington, DC, USA; ²DecisionQ Corporation, Washington, DC, USA; ³Regenerative Medicine Department, Combat Casualty Care, Naval Medical Research Center, Silver Spring, MD, USA; ⁵Integrated Department of Orthopedics and Rehabilitation, ⁶Department of Surgery, National Naval Medical Center, Bethesda, MD, USA

Background: Graphical probabilistic models have the ability to provide insights as to how clinical factors are conditionally related. These models can be used to help us understand factors influencing health care outcomes and resource utilization, and to estimate morbidity and clinical outcomes in trauma patient populations.

Study design: Thirty-two combat casualties with severe extremity injuries enrolled in a prospective observational study were analyzed using step-wise machine-learned Bayesian belief network (BBN) and step-wise logistic regression (LR). Models were evaluated using 10-fold cross-validation to calculate area-under-the-curve (AUC) from receiver operating characteristics (ROC) curves.

Results: Our BBN showed important associations between various factors in our data set that could not be developed using standard regression methods. Cross-validated ROC curve analysis showed that our BBN model was a robust representation of our data domain and that LR models trained on these findings were also robust: hospital-acquired infection (AUC: LR, 0.81; BBN, 0.79), intensive care unit length of stay (AUC: LR, 0.97; BBN, 0.81), and wound healing (AUC: LR, 0.91; BBN, 0.72) showed strong AUC.

Conclusions: A BBN model can effectively represent clinical outcomes and biomarkers in patients hospitalized after severe wounding, and is confirmed by 10-fold cross-validation and further confirmed through logistic regression modeling. The method warrants further development and independent validation in other, more diverse patient populations.

Keywords: combat, wounds, probabilistic model, Bayesian belief network, outcomes

Introduction

Blast-related injuries predominate on the modern day battlefields in Iraq and Afghanistan. Improvised bombs and rocket attacks inflict devastating injuries on both civilians and military personnel.¹⁻³ Frequency and severity notwithstanding, more military personnel are surviving these attacks due to a number of advances in care and body protection, albeit with a marked change in the type of injuries most commonly sustained. The development and widespread utilization of effective personal body armor has further shifted injury concentrations to the extremities, as up to 80% of surviving combat-injured personnel sustain extremity injuries.⁴⁻⁶ These advances in personal body armor, coupled with improved vehicular armor, rapid aero-medical evacuation of casualties, and the deployment far-forward of cutting-edge medical technologies and treatments have improved the survival of wounded service members. Likewise, similar emergency medical service and technological advances have improved treatment of civilian casualties.⁷

Correspondence: Colonel Alexander Stojadinovic,
Walter Reed Army Medical Center, 6900 Georgia Avenue, Suite 5C27A, NW Washington, DC 20307, USA
Email alexander.stojadinovic@us.army.mil

The cumulative result of these technological advances is an unprecedented cohort of surviving casualties with devastating war wounds, traumatic amputations, and penetrating and closed traumatic brain injuries. These high-energy blast wounds are characterized by massive zones of injury that combine bone, muscle, and soft tissue loss with gross bacterial and retained metal and composite material wound contamination.^{8–10} Patient management consists of rapid initial stabilization in the theater of operations, inter-continental aero-medical evacuation, and multiple surgical debridement procedures every 24–48 hours at combat casualty care waypoints. A majority of injured service members present to definitive treatment facilities, such as Walter Reed Army Medical Center and National Naval Medical Center within the continental United States, with multiple complex wounds and a profound, ongoing systemic inflammatory response. These patients require additional diagnostic and serial therapeutic interventions, including continued surgical debridement procedures, which results in a prolonged hospital stay and long-term rehabilitation, amounting to resource-intensive protracted health care.

A considerable amount of time and resources is committed to the care of this highly complicated patient population. Unfortunately, patient-specific estimates of resource allocation, critical care utilization, length of hospital stay (LOS), nosocomial morbidity, and individual wound outcomes are difficult to create, or do not exist. Blast injuries pose formidable therapeutic challenges, and occur in the context of multiple local and systemic impediments to healing, making improving health care quality and outcomes extremely challenging. These include systemic and local inflammation, as evidenced by measurable cytokine and chemokine profiles, deficient nutritional status, variable wound debridement adequacy, and the often-compromised vascular status of the injured limb(s). Describing and projecting relational outcomes of these highly complex, interrelated, and time-dependent variables have proven difficult in modern-day health care systems. Additionally, no applicable prognostic model exists for this complex patient population.

Recently, there has been renewed interest in a modeling technique used to evaluate complex relationships such as those that exist in the cohort described above. Bayesian models have been demonstrated to be useful in determining injury severity,^{11,12} intensive care unit (ICU) mortality,¹³ operative risk,¹⁴ and surgical outcomes.^{15–18} As greater emphasis is placed on improving the efficiency and quality of health care, improved methods to facilitate understanding of key factors impacting patient outcomes are needed. We sought to evaluate

the use of Bayesian belief networks (BBNs) as a method for developing networks of associations between clinical factors influencing health care outcomes of particular interest. The use of the BBN model has clinical utility for estimating clinical outcomes, particularly with incomplete clinical data, unlike logistic regression methodology, which requires complete data sets for prognostic variables. The BBN is a graphical modeling methodology that presents associations in a hierarchical, comprehensible, graphical structure through an interactive interface for querying the model at point-of-care. This allows health care providers and clinical researchers a straightforward method to estimate clinical outcomes of interest on an individualized patient basis. Importantly, the BBN provides insights for complex clinical situations, showing how important factors interrelate to impact important health care-related outcomes. This understanding can assist clinicians in developing individualized, targeted therapeutic interventions. We chose to evaluate this method in a severely wounded health care, resource-intensive, blast-related injury population by training a BBN. The BBN model was validated using 10-fold cross-validation, and then compared to logistic regression models trained, in order to gain insights from the BBN model.

In this proof-of-principle study, we explore the utility of BBN model development to help expand our understanding of how biomarkers and health care outcomes associate in a population of war-wounded service members enrolled in a prospective observational clinical trial, through the development of a graph of variable associations. We trained a BBN to evaluate associations between hospital-acquired infection, ICU LOS, hospital LOS, impaired wound healing, serum biomarkers, and clinical data at the time of admission to a tertiary care military medical center. This model effectively provides a model of association and estimation of nosocomial morbidity, hospital resource utilization, and length of stay. As such, it may serve as the basis for further independent validation studies in diverse trauma patient populations, and the development of novel models for health care quality improvement in patient populations with traumatic wounds.

Methods

Study methodology

The institutional review board approved this prospective observational clinical trial in compliance with all applicable Federal regulations governing the protection of human subjects. Study participants were recruited from wounded US service members evacuated from Iraq and Afghanistan. Informed

consent was obtained from each study subject, or legal medical representative. Inclusion criteria for this study were defined as adult, active duty service members who sustained high-energy, penetrating (open) extremity injuries during combat operations abroad. Those with pre-morbid confounding inflammatory conditions, including immune deficiency and connective tissue disorders, or any medical illness requiring immunosuppressive therapy, were excluded a priori.

Demographic and injury-related data were collected prospectively, including: gender and age; date, location and mechanism of wounding; requirement for blood transfusion and total units of packed red blood cells (PRBCs) transfused; injury severity score (ISS) and Acute Physiology and Chronic Health Evaluation II (APACHE II) scores; wound surface area and depth, associated major vascular injury to the affected extremity; and type of wound closure or coverage; Glasgow Coma Scale (GCS) and concomitant traumatic brain injury (defined by a consulting subject matter expert [eg, neurosurgeon, neurologist] in the setting of closed or penetrating intracranial injury resulting from an externally applied force); intensive care unit length of stay, need for mechanical ventilation and ventilator days; number of surgical wound debridements in the operating room and hospital length of stay; development of nosocomial infection during the index hospitalization and impaired wound healing, as defined below.

Surgical debridement, saline irrigation, and negative pressure therapy utilizing vacuum-assisted closure (VAC®; Kinetic Concepts, San Antonio, TX) were repeated every 48 to 72 hours until surgical wound closure or coverage (flap or split-thickness skin graft) occurred according to the current institutional standards of practice and at the discretion of the attending surgeon.

Human biological specimen collection

Peripheral venous whole blood (8 mL) was collected in a Red-Top Serum BD Vacutainer® (Becton Dickinson, Franklin Lakes, NJ) prior to each surgical wound debridement, and immediately fractionated for serum collection using a centrifuge (Thermo-Electron Corp, Waltham, MA) at $2500 \times g$ (4°C) for 10 minutes. Serum supernatant samples were transferred to individually labeled Cryo-Loc™ polypropylene tubes (Lake Charles Manufacturing, Lake Charles, LA) and flash-frozen in liquid nitrogen. All samples were stored at -70°C until analysis.

Serum inflammatory biomarker analysis

Serum was diluted twofold with Beadlyte® Human Serum Sample Diluent (Millipore Corp, Billerica, MA) and as

previously described, serum proteins of interest were quantitated using a Beadlyte® Human 22-Plex Multi-Cytokine Detection System on the Luminex® 100 IS xMAP Bead Array Platform (Millipore Corp).^{11,22} Briefly, this system utilizes analyte-specific monoclonal antibodies covalently linked to uniquely fluorescent beads with subsequent fluorescent-report of analyte binding. Twenty-two cytokines were quantified (pg/mL) according to manufacturer's instructions.

Wound closure, follow-up, and outcomes

All wounds were examined daily following wound closure or coverage until suture removal. All patients were followed clinically for a minimum of 30 days. Impaired wound healing included delayed wound closure or subsequent wound dehiscence. Delayed wound closure was defined as definitive closure occurring two standard deviations outside of the mean normal wound closure time period, or ≥ 21 days after injury. Dehiscence was defined as spontaneous partial or complete disruption of the surgical wound after closure or $> 50\%$ graft loss, necessitating a return to the operating room for treatment. Wounds that progressed to healing at 30 days following closure, without necessitating a return to the operating room, were considered healed. Hospital length of stay was defined from time of admission to the tertiary care Military Treatment Facility to the time of discharge from the same. Intensive care unit length of stay was calculated from time of admission to the ICU until time of transfer out of the ICU. Nosocomial infection was defined as an infection that was a result of treatment in hospital, but secondary to the patient's original condition, that appeared 48 hours or more after admission or within 30 days of hospital discharge. Any culture-positive infection that, based on symptoms, clinical, or radiological signs, required antibiotic therapy in the opinion of an Infectious Disease physician was considered nosocomial.

Statistical analysis

Development of a Bayesian model to estimate outcome in war wounded

Data analysis was conducted using a BBN model. As a data modeling tool, BBNs have unique value in that they provide the user with a graphical representation of how variables associate to estimate outcomes. Multivariate dependence relationships between clinical variables, serum inflammation-related biomarkers measured at time of study enrollment, and outcomes (nosocomial infection, ICU LOS, and impaired wound healing) were identified using FasterAnalytics™ Bayesian modeling software (DecisionQ, Washington, DC). This Bayesian modeling software uses machine-learning

algorithms to dynamically detect complex patterns in multidimensional data sets. The BBN model provided demonstrates how multiple clinical variables associate when estimating outcomes. The BBN modeling software supports a step-wise process in order to develop a robust model, consisting of an iterative process of training, qualitative and quantitative evaluation, parameter tuning, attribute pruning, and re-training. Tenfold intra-set cross-validation was performed, to assess robustness of the final model, by training a BBN model on each of ten randomly-created, nonoverlapping training sets, each of which was used to create a set of case-specific estimates, and each of which is characterized by a receiver operating characteristics curve to establish Bayesian classification accuracy.^{18–20}

The 10-fold cross validation method is an established approach that provides reliable estimates of model accuracy. BBN models were trained both with and without biomarker data to assess the relative contribution of biomarkers to current clinical data.

Development of logistic regression models to estimate outcome in war wounded

Forward stepwise logistic regression was also used to model estimates of impaired wound healing, ICU admission and occurrence of infection to assess whether the findings of the BBN would continue to be robust if transferred into another modeling methodology. Stepwise regression was chosen in the absence of an a priori rationale for ordering entry of predictor variables into the model, and is considered exploratory. Forward stepwise entry uses a likelihood ratio test (chi-square difference), based on maximum likelihood estimation, in order to determine which variables to retain in the model. It determines the forward entry model and then alternates between backward elimination and forward entry until all variables not in the model fail to meet entry or removal criteria. On the first step of each equation, selected clinical variables, determined by the Bayesian analysis, were entered: impaired wound healing was regressed on ICU admission; ICU admission was regressed on PRBCs and the APACHE II score; the occurrence of infection was regressed on PRBCs, the APACHE II score and ISS. In the second step of each equation, four biomarkers were entered: IL-6, IL-8, IL-12 p40, and MCP1, in order to assess the contribution of biomarkers to predictive power. Incomplete records were censored, as logistic regression analysis does not support modeling with incomplete information. The robustness of the LR models was assessed using 10-fold intra-set cross-validation

and receiver operating characteristics (ROC) curve analysis, wherein the Area Under the Curve (AUC) for each of the models, with and without biomarkers, was compared to assess the improvement (curve lift) in the AUC curve from the addition of serum biomarkers. As the Bayesian analysis was used for variable selection in the logistic regression analysis, the regression analysis was conducted to determine if the findings of the BBN analysis would continue to be robust in other modeling methodologies.

Results

Patient and wound characteristics

Thirty-two patients were enrolled in the study; all were males (age 19 to 42 years) with predominantly severe blast-related injuries of war (Table 1). Sixty percent of patients required blood products for initial resuscitation (median 2, range 1–134 units PRBCs), and 40% required ICU admission for an average of 6 days per patient. Mean hospital LOS was one month (range 11–117 days), over which time an average of 6 surgical wound interventions were required until definitive wound closure or coverage. Nosocomial

Table 1 Patient (wound) demographics

Category	n =	%	Mean ± SD
Male	32	100.0%	
Age (years)			24.0 ± 5.9
Mechanism of injury			
Blast	21	65.6%	
Blast and blunt trauma	6	18.8%	
Gunshot wound	4	12.5%	
Fall from height	1	3.1%	
Injury severity			
GCS			14.0 ± 2.3
ISS			15.0 ± 9.7
APACHE			6.4 ± 5.2
ICU admission	13	40.6%	
Mechanical ventilation	5	15.6%	
Transfusion			
Initial PRBC transfusion			10.1 ± 24.3
Initial PRBC ≤ 4 units	11	34.4%	
Initial PRBC > 4 units	21	65.6%	
Surgical interventions			
			6.0 ± 6.2
Length of stay (days)			
On ventilator			2.5 ± 6.8
ICU			5.6 ± 14.5
Hospital			31.0 ± 23.6
Outcomes			
Any infection	14	43.8%	
Impaired wound healing	6	18.8%	
Mortality	1	3.1%	

Abbreviations: GCS, Glasgow Coma Scale; ISS, injury severity score; APACHE, acute physiology and chronic health evaluation; ICU, intensive care unit; PRBC, packed red blood cells.

infections occurred in 44%, and impaired wound healing in 19% of patients. Mean ISS and APACHE II scores at time of admission were 15.0 ± 9.7 (range: 4–36), and 6.4 ± 5.2 (range: 1–22) respectively.

Development of BBN and LR models to estimate outcome in war wounded

Figure 1 depicts the structure of the BBN network developed to estimate likelihood of nosocomial infection (Infection Cat), requirement for ICU care (ICU Admit Cat), and impaired wound healing (Impaired Wound Healing Cat) without the use of serum biomarkers. This figure not only shows that there are extensive associations between outcomes and between clinical variables and outcomes, but also allows us to understand how these different variables associate when estimating outcomes, allowing us to develop hypotheses about mechanisms that can support clinical intervention. Figure 1 is a directed acyclic graph, wherein the arcs (lines) represent conditional dependence associations between otherwise independent factors. Thus, likelihood of nosocomial infection can be estimated *a priori* using serum albumin, injury severity score (ISS), and initial transfusion requirement; impaired wound healing can be estimated *a priori* using ICU admission; and likelihood of ICU admission can be estimated using initial transfusion requirement and APACHE II score. Figures 2–5 shows the revised model, including serum biomarker data. The structure of the network in Figure 2 shows that likelihood of nosocomial infection can be estimated by ISS and initial transfusion requirement. Likelihood of ICU admission can be estimated using initial transfusion requirement, APACHE II score, and serum IL-6. Finally, impaired wound healing can

be estimated *a priori* using ICU admission, serum IL-8, and MCP-1 through estimated hospital LOS.

Table 2 details the ten-fold cross-validation results from each train-and-test pair for the summarized posterior probability estimates of nosocomial infection, ICU admission, and impaired wound healing for both the BBN and LR Models. Both the trained BBN models and LR models are robust, with and without serum biomarkers, with ROC AUCs ranging from 0.72 to 0.97. The table shows, for example, that the LR model can estimate nosocomial infection with an AUC of 0.81, both with and without biomarkers, while the addition of biomarkers improves the AUC of the BBN model from 0.76 to 0.79. Sensitivity and specificity statistics for LR and BBN models also show promise, with the ability to detect 83% of impaired healing wounds with 67% specificity using the BBN with biomarkers, while the same BBN can detect 93% of nosocomial infections with 63% specificity. These performance characteristics are consistent with many current clinical diagnostic tests, and with a larger sample size, there is opportunity for further improvement. The BBN receiver operating characteristics curve lift column calculates the improvement (or degradation) of performance statistics of the BBN compared to the LR models, while the biomarker lift column does the same calculation to evaluate the addition of serum biomarkers. The results for both methods are robust and remarkably similar. However, each method and feature set can be shown to outperform in a specific area. Importantly, one element that was missing from the cross-validation results was the assessment of the impact of partial information on robustness. Often, decisions are made on partial (incomplete or inaccurate) clinical information,

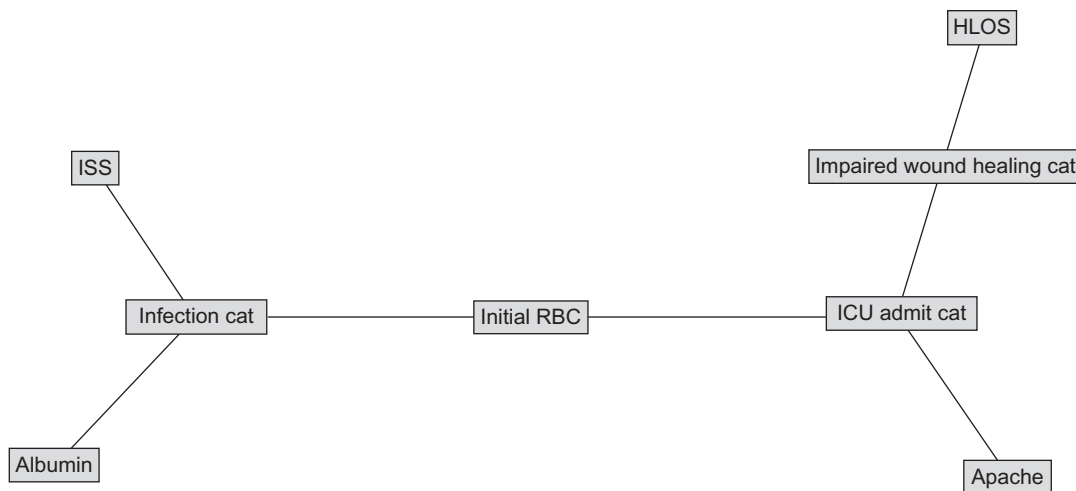


Figure 1 Bayesian belief network of outcomes in war wounded, excluding biomarker data. The network demonstrates the hierarchical structure of the conditional dependence relationships between clinical parameter study variables.

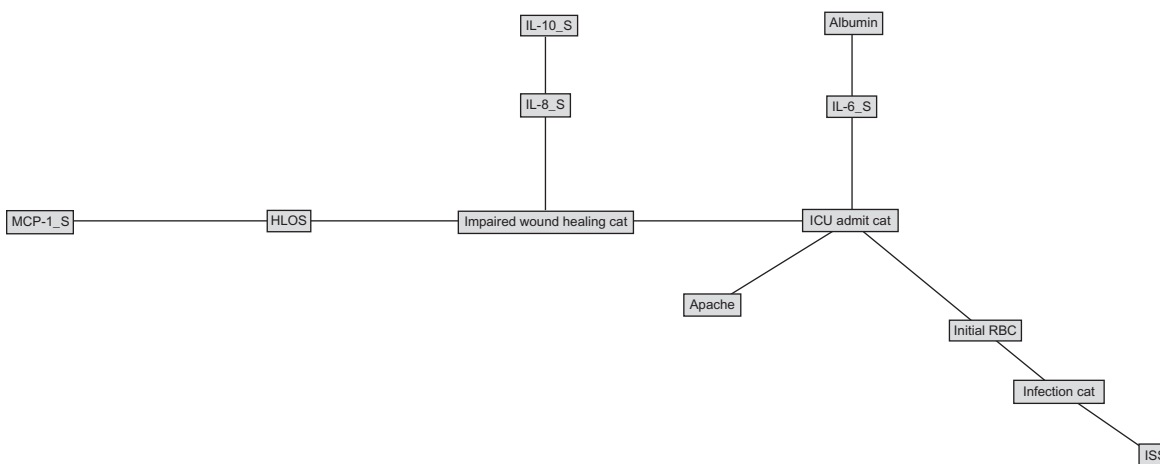


Figure 2 Bayesian belief network of outcomes in war wounded, including biomarker data.

and our decision support algorithms must incorporate this level of uncertainty. Logistic regression models cannot be calculated using partial records. Hence, LR models produce positively biased results. Conversely, BBN are designed to reflect uncertainty. Therefore, they can be trained on partial records, producing a negatively biased result. The cross-validation statistics for the BBN models developed in this

study reflect the inclusion of partial information, while the cross-validation statistics of the LR models do not. In order to assess the potential impact of partial information in the clinical setting, we used the nine records censored from the LR modeling (missing serum albumin results [n = 3], and missing results of some of the serum biomarkers [incomplete data set in 6 patients]) to perform independent set validation

Table 2 Summary of 10-fold intra-set cross-validation exercises describing area-under-the-curve, sensitivity, specificity, accuracy, negative, and positive predictive values for key study outcomes. The first two column groups describe the performance of each of the BBN and LR models, with and without biomarkers. The BBN Lift column group describes the lift, in points, of the BBN model compared to the LR model in each test category, both excluding and including biomarkers, while the Biomarker Lift column group describes the lift of adding biomarker data to both the LR and BBN models. Degradations in performance are described in parentheses

	Without biomarkers		With biomarkers		BBN lift		Biomarkers lift	
	LR	BBN	LR	BBN	Without	With	LR	BBN
Impaired wound healing								
AUC	0.91	0.79	0.91	0.72	(0.12)	(0.19)	0.00	(0.07)
Percent correct	85.7%	76.7%	85.7%	70.0%	(0.09)	(0.16)	0.00	(0.07)
Positive predictive value	100.0%	40.0%	100.0%	38.5%	(0.60)	(0.62)	0.00	(0.02)
Negative predictive value	81.0%	95.0%	81.0%	94.1%	0.14	0.13	0.00	(0.01)
Sensitivity	61.0%	80.0%	62.0%	83.3%	0.19	0.21	0.01	0.03
Specificity	100.0%	76.0%	100.0%	66.7%	(0.24)	(0.33)	0.00	(0.09)
ICU admission								
AUC	0.97	0.76	0.97	0.81	(0.21)	(0.16)	0.00	0.05
Percent correct	93.0%	80.0%	95.0%	80.0%	(0.13)	(0.15)	0.02	0.00
Positive predictive value	87.0%	72.7%	91.0%	72.7%	(0.14)	(0.18)	0.04	0.00
Negative predictive value	99.0%	84.2%	97.0%	84.2%	(0.15)	(0.13)	(0.02)	0.00
Sensitivity	98.0%	72.7%	95.0%	72.7%	(0.25)	(0.22)	(0.03)	0.00
Specificity	93.0%	84.2%	95.0%	84.2%	(0.09)	(0.11)	0.02	0.00
Nosocomial infection								
AUC	0.81	0.76	0.81	0.79	(0.05)	(0.02)	0.00	0.03
Percent correct	73.0%	70.0%	73.0%	76.7%	(0.03)	0.04	0.00	0.07
Positive predictive value	55.0%	75.0%	55.0%	68.4%	0.20	0.13	0.00	(0.07)
Negative predictive value	84.0%	68.2%	84.0%	90.9%	(0.16)	0.07	0.00	0.23
Sensitivity	67.0%	53.8%	67.0%	92.9%	(0.13)	0.26	0.00	0.39
Specificity	75.0%	88.2%	75.0%	62.5%	0.13	(0.13)	0.00	(0.26)

Abbreviations: BBN, Bayesian belief network; LR, logistic regression; AUC, area under the curve; ICU, intensive care unit.

on the LR models. When we conducted this independent set validation on the LR models, the percent correct for impaired wound healing dropped from 86% to 78%, ICU admission from 95% to 56%, and ability to estimate nosocomial infection from 73% to 44%.

BBN models may be useful at the point of clinical care. This point is illustrated in Table 3, which represents an inference table calculated using the BBN model developed in this study. It shows how knowing ICU admission requirement (Yes or No), and the serum MCP-1 (pg/mL), and IL-8 (pg/mL) concentrations for an individual patient allows one to estimate case-specific likelihood of wound healing.

Discussion

The current conflicts in Iraq and Afghanistan have produced a patient population with multiple traumatic injuries for which definitive treatment requires the mobilization of large amounts of health care system resources to support protracted multidisciplinary care. We have previously determined that the systemic inflammatory response to injury in these patients can be measured and is related to subsequent outcomes such as wound failure, the development of heterotopic ossification, and the need for blood transfusions.^{21,22} However, we still lack a clear understanding of how health care outcomes and quality factors are related to one another, and how biomarkers can help us to better understand these quality indicators. We have

developed a robust BBN model based on these observations using machine-learned BBNs, and then validated these findings through both cross-validation and further analysis with additional, previously nonexistent logistic regression models. By analyzing both clinical variables and serum protein measurements (biomarkers) with the BBN models, we have developed a series of robust graphical models that can be used not only to estimate individual risk of wound failure, hospital-acquired infection, and intensive care admission, but also to give us a better understanding of the associations between these factors, from which we can begin to design changes in clinical practice directed at improving outcomes and quality.

While both methods (BBN and LR) produced robust statistical results using intra-set cross-validation testing, there are some subtle but critical differences. Cross-validation allows us to estimate model performance in novel populations and to develop robust estimates of how we would expect a model to perform in a de novo population. Estimates of sensitivity, specificity, and positive and negative predictive values are consistent with the performance of many current clinical chemistry tests. With a larger study population and additional modeling, further improvements and optimization should be achievable. First, the selection of covariates was accomplished for both methods using the machine-learned BBN. The reason for this is that our initial

Table 3 Inference table calculated using the Bayesian model developed in this study for selected cases of ICU admission, serum MCP-1, and IL-8 concentrations

Probability of case	Independent variables			Dependent variable	
	ICU admit cat	IL-8_S	MCP-I_S	Impaired wound healing cat	
				No	Yes
17.396%	No	Up to 17.9	Up to 38.6	98.0	2.0
7.721%	Yes	Up to 17.9	Up to 38.6	74.0	26.0
3.496%	No	17.9 to 19.6	Up to 38.6	97.5	2.5
1.644%	Yes	17.9 to 19.6	Up to 38.6	69.4	30.6
1.399%	No	19.6 plus	Up to 38.6	81.2	18.8
1.889%	Yes	19.6 plus	Up to 38.6	20.1	79.9
16.416%	No	Up to 17.9	38.6 to 72.4	96.7	3.3
8.394%	Yes	Up to 17.9	38.6 to 72.4	63.4	36.6
3.31%	No	17.9 to 19.6	38.6 to 72.4	96.0	4.0
1.832%	Yes	17.9 to 19.6	38.6 to 72.4	58.1	41.9
1.46%	No	19.6 plus	38.6 to 72.4	72.5	27.5
2.66%	Yes	19.6 plus	38.6 to 72.4	13.3	86.7
14.896%	No	Up to 17.9	72.4 plus	95.9	4.1
8.285%	Yes	Up to 17.9	72.4 plus	57.8	42.2
3.009%	No	17.9 to 19.6	72.4 plus	94.9	5.1
1.831%	Yes	17.9 to 19.6	72.4 plus	52.2	47.8
1.409%	No	19.6 plus	72.4 plus	67.6	32.4
2.943%	Yes	19.6 plus	72.4 plus	10.8	89.2

data set included clinical data as well as a 22-cytokine/chemokine panel, resulting in complex multivariate associations that are extremely difficult to assess and characterize using classical frequentist statistical approaches. Further, the BBN methodology incorporates all data elements into a single hierarchical graphical network, while the LR models are trained individually for each outcome of interest using a linear construct. While this difference in method allows the LR models to more closely fit the data, producing better cross-validation statistics in most cases, this most likely comes at the cost of reduced robustness, as shown when accuracy is assessed using clinical records with only partial information, where the accuracy statistic degrades from 85.7% to 77.8% in impaired wound healing, 95.0% to 55.6% in ICU admission, and 73.0% to 44.4% in infection. By comparison, all of the cross-validation statistics for the BBN models are inclusive of missing information. We further assessed the robustness of the LR models using partial information records that were censored in the LR model training but were included in the BBN model training. When this analysis was performed, the accuracy statistics of the LR models degraded significantly, suggesting that the LR results may be positively biased by method, and may not be as robust as indicated when applied to actual (incomplete) clinical data. One of the critical differences between the BBN and LR methods, and the reason why BBNs can incorporate partial information without degraded robustness, is that the BBN is a nonlinear parametric classification method that demonstrates multivariate relationships between attributes that inform

each other. LR is a linear classification method which does not allow for features to inform one another. Hence, partial information significantly degrades the robustness of the LR methodology.

The critical improvement of using a graphical modeling methodology such as the BBN method is its ability to present associations in a hierarchical, user-friendly graphical structure with an interactive interface for querying the model at point-of-care (Figure 1). This allows clinicians and researchers to not only arrive at posterior estimates of outcomes of interest while applying the model to individual patients, but also to begin to develop a better understanding of how factors combine to create important health care-related outcomes, which in turn allows the clinician to develop novel, targeted interventions. A clinician can enter patient-specific observations (evidence) into the model to derive estimates of hospital-acquired infection, requirement for intensive care, and likelihood of uncomplicated wound healing. This then permits clinicians to anticipate or estimate an outcome or complication of interest and intervene accordingly in a timely and cost-effective manner. For example, using the BBN, a patient with no prior transfusion requirement and an $ISS \leq 9$ has a low (7%) likelihood of nosocomial infection (Figure 3). If the same patient required four or more units of blood transfusions, the likelihood of hospital-acquired infection increases dramatically to 58% (Figure 4). This same patient's information can simultaneously be utilized to calculate estimates of need for critical care (ICU admission, which is increased from 13% to 71%) and likelihood

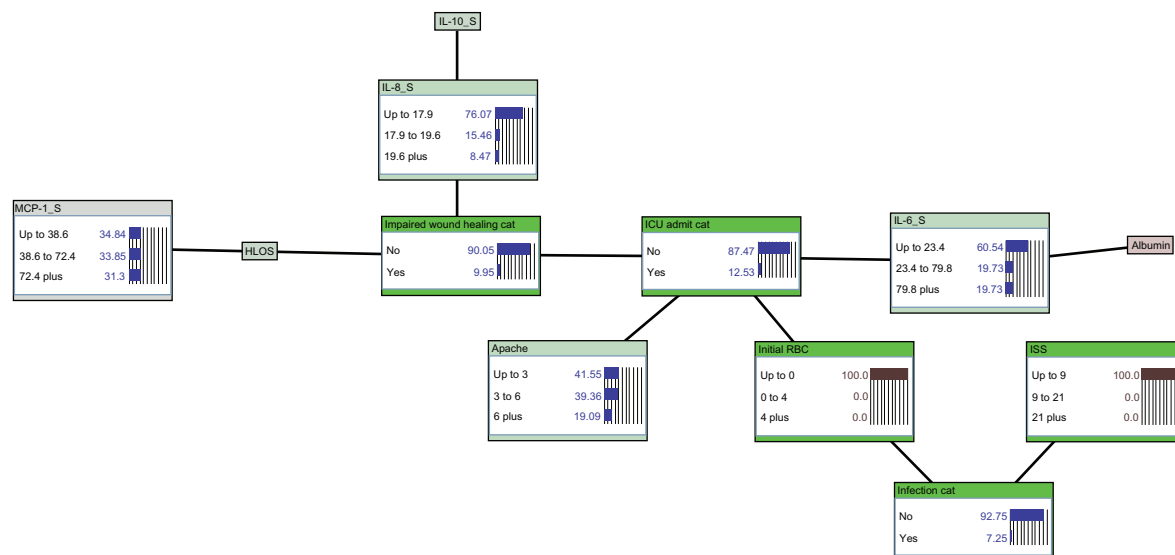


Figure 3 Estimate of infectious morbidity outcome derived from prior knowledge of initial transfusion requirement and ISS. Likelihood of infection is 7% with evidence of no prior blood product transfusion and lowest ISS category ($ISS \leq 9$). In addition, no initial transfusion requirement also results in a relatively low likelihood of ICU admission (13%). **Abbreviations:** ISS, injury severity score; ICU, intensive care unit.

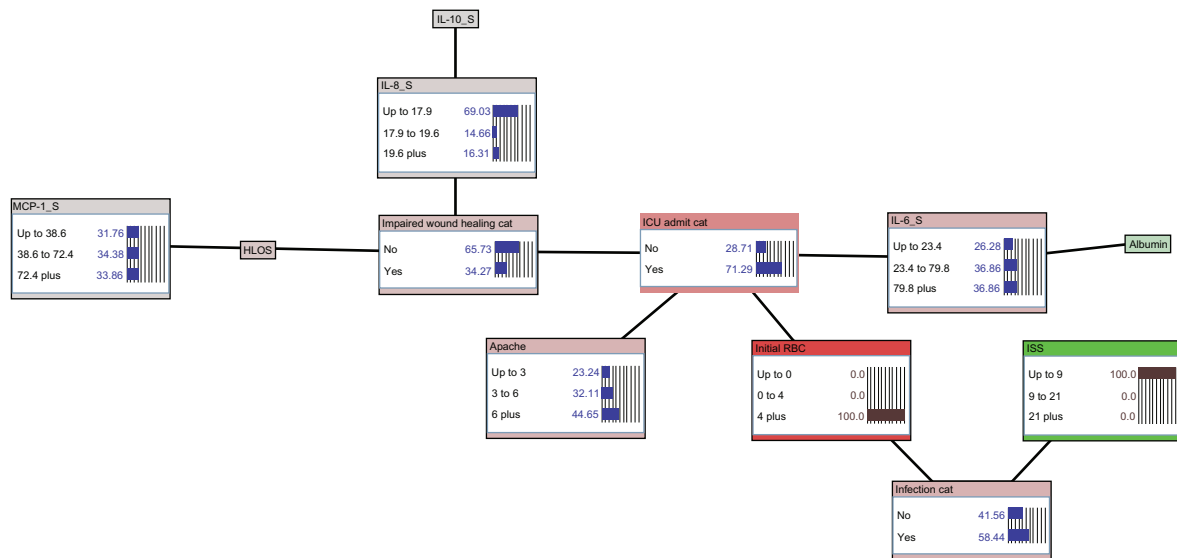


Figure 4 Estimate of infectious morbidity and ICU care outcome derived from prior knowledge of initial transfusion requirement and ISS. Likelihood of infection is 58% based on initial transfusion of 4 or more units of PRBCs and lowest ISS category (ISS ≤ 9). Likelihood of ICU admission increases to 71%. **Abbreviations:** ICU, intensive care unit; ISS, injury severity score; PCRb, packed red blood cells.

of adverse wound healing outcome (risk of wound failure increased from 10% to 34%; Figure 5).

The results of this study are promising. Tenfold cross-validation indicated that our BBN model is indeed robust, as measured by the receiver operating characteristics AUCs ranging from 0.72–0.81, and these results were re-affirmed by step-wise LR modeling using features identified by the BBN, which produce AUCs ranging from 0.81 to 0.97. Expanding the scope of this study is warranted to further

assess model robustness, and to determine whether or not this model can improve clinical practice and patient outcomes in widespread application. Further study of additional data features may also improve the model, not only in terms of its robustness but also in terms of its clinical utility. The potential clinical utility of our model is demonstrated in that physicians applying ICU admission information, serum MCP-1, and IL-8 concentrations can develop case-specific estimates of resource-intensive wound care due to impaired

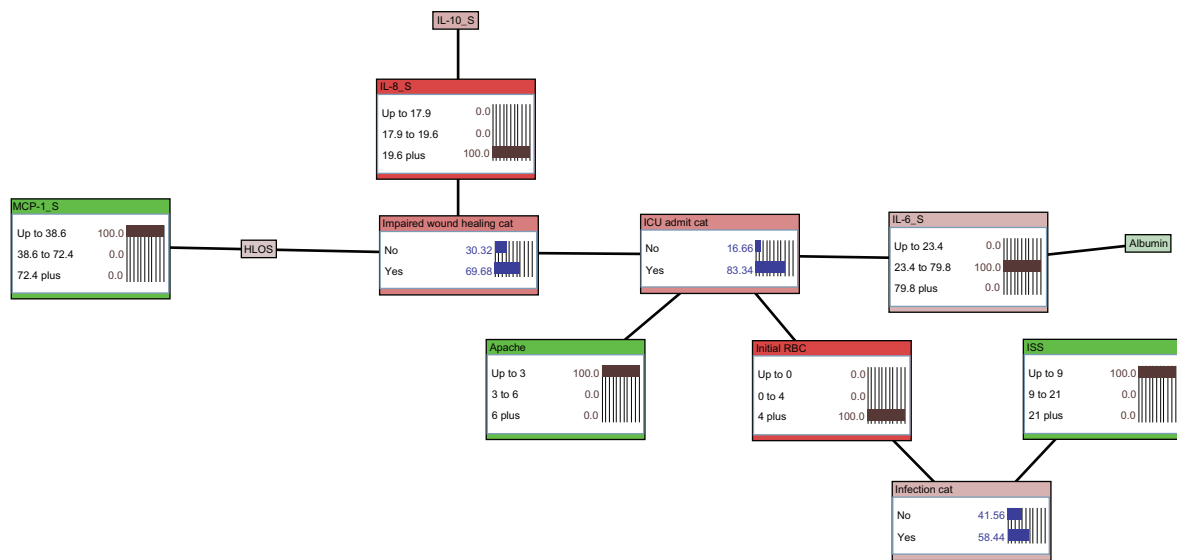


Figure 5 Estimate of infectious morbidity, ICU care, and wound healing outcome derived from prior knowledge of initial transfusion requirement, ISS, APACHE II score, and serum concentrations of IL-6, IL-8, and MCP-1. Adding prior knowledge of an APACHE II of less than 3, IL-6 concentration of 23.4 to 79.8, MCP-1 concentration of less than 38.6, and IL-8 concentration of greater than 19.6 leaves the posterior estimate of infection unchanged at 58%, while increasing likelihood of ICU admission to 83%, and yielding a likelihood of impaired wound healing of 70%. **Abbreviations:** ICU, intensive care unit; ISS, injury severity score; APACHE II, acute physiology and chronic health evaluation II.

wound healing (Table 3). Further, the clinician can query the graphical user interface to understand why the estimate of outcome is being derived the way it is and what patient- and health care-related factors might be influenced to change it. One of the strengths of the methodology pilot-tested in this study is that it accounts for dimensionality and uncertainty and has the ability to codify complex clinical problems into straightforward, intuitive, robust classification models. For example, our simple pilot model has 729 distinct rule sets that can be applied using a priori information and statistically significant associations. These types of complex rule sets are difficult to translate into user-friendly systems. However, the BBN model interface allows the user to input clinical and laboratory data into a complex statistical model in a graphical, “user friendly” output (Figures 2–5).

While our early work is promising, our current model has distinct limitations given the complexity of our data. First, this model was developed using a limited patient population for the purposes of this pilot proof-of-concept study. While our cross-validation demonstrated that the model is robust, there is expected variance between testing exercises due to this small sample size. In order to reduce this variance, population sample size expansion is indicated to further refine our BBN model. Second, while our current model shows promising receiver operating characteristics for health care outcomes that are reasonable, further improvement is warranted. Finally, while the biomarkers studied were not additive, they were indeed illustrative of the utility of the modeling process applied in this study. These limitations aside, we believe that this approach outlines a significant step forward in the care of severely wounded casualties requiring complex multidisciplinary, resource-intensive care, with possible application far beyond the military population, in better understanding health care outcomes and quality. When compared with conventional, frequentist statistics, the BBN-defined model was equivalent and demonstrated increased robustness when missing variables (often encountered in clinical practice) were evaluated. Further, this approach was able to define relationships not readily apparent with standard methods, and to codify them into a graphical interface that allows the clinician to understand how patterns of association relate to outcomes and quality on a patient-specific basis. This ability to analyze complex datasets consisting of clinical data, standard laboratory values, and molecular biomarkers in an inter-dependent fashion represents a meaningful advance in our toolset for improving health care quality and outcomes. Demonstrating the ability to estimate outcomes is the first step among many towards individualized care, as models such as these are coupled to medical treatment

interventions. Using a combat-wounded population, we have demonstrated the potential of such an approach, that warrants expansion into a larger, broader trial, as well as to other aspects of medical and surgical care.

Conclusion

We developed a BBN model utilizing numerous clinical factors and biomarkers. This integrated proof-of-principle model is robust in estimating clinical outcomes in patients hospitalized with combat wounds, as determined through cross-validation. We further re-affirmed our findings with additional cross-validated logistic regression modeling, trained using insights developed from the BBN. We chose a homogeneous study cohort to limit the number of confounding variables in this pilot study. The model, in its current form, is not generalizable to a broad trauma population. From this initial development of a BBN for the estimation of clinically relevant outcomes, we intend to expand the study to a larger trauma population, and introduce additional variables in a step-wise manner in order to make the model applicable to a wide-ranging trauma population. The methods developed in this study have the potential to fundamentally change the quality of health care by empowering the clinician to reason uncertainty in complex high-dimensionality datasets, and reduce the information into usable and personalized networks of associations that allow the clinician to both better understand those factors influencing outcome as well as to estimate the likelihood of a case-specific outcome.

Disclosure

This effort was supported, in part, by the US Navy Bureau of Medicine and Surgery under the Medical Development Program (PE 0604771N), Office of Naval Research work unit number 604771N.0933.001.A0812 and the Combat Wound Initiative Program.

This study was approved by the National Naval Medical Center Institutional Review Board in compliance with all Federal regulations governing the protection of human subjects.

We are military service members (or employees of the US Government). This work was prepared as part of our official duties. Title 17 U.S.C. 105 provides the “Copyright protection under this title is not available for any work of the United States Government.” Title 17 U.S.C. 101 defines a US Government work as a work prepared by a military service member or employee of the US Government as part of that person’s official duties.

We certify that all individuals who qualify as authors have been listed; each has participated in the conception and design of this work, the analysis of data (when applicable), the writing of the document, and the approval of the submission of this version; that the document represents valid work; that if we used information derived from another source, we obtained all necessary approvals to use it and made appropriate acknowledgements in the document; and that each takes public responsibility for it.

The views expressed in this manuscript are those of the authors and do not reflect the official policy of the Department of the Army, Department of the Navy, the Department of Defense, or the United States Government.

Acknowledgments

We have a government-sponsored translational research program, which has partnered with DecisionQ Corporation to develop predictive models to advance personalized medicine within the Department of Defense.

We acknowledge Edward R Utz, MS for his assistance, supported in part by the Alpha Omega Alpha Carolyn L Kuckein Student Research Fellowship.

The multidisciplinary care of these patients would not have been possible without the dedicated efforts of everyone at National Naval Medical Center and Walter Reed Army Medical Center. Both civilian and military personnel have rendered skilled and compassionate care for these casualties. All of our efforts are dedicated to those who have been placed in harm's way for the good of our nation.

References

- Montgomery SP, Swiecki CW, Shriver CD. The evaluation of casualties from Operation Iraqi Freedom on return to the continental United States from March to June 2003. *J Am Coll Surg*. 2005;201:7–12, discussion 12–13.
- Peoples GE, Jezior JR, Shriver CD. Caring for the wounded in Iraq – a photo essay. *N Engl J Med*. 2004;351:2476–2480.
- Owens BD, Kragh JF Jr, Wenke JC, Macaitis J, Wade CE, Holcomb JB. Combat wounds in operation Iraqi Freedom and operation Enduring Freedom. *J Trauma*. 2008;64:295–299.
- Owens BD, Kragh JF, Macaitis J, Svoboda SJ, Wenke JC. Characterization of extremity wounds in Operation Iraqi Freedom and Operation Enduring Freedom. *J Orthop Trauma*. 2007;21:254–257.
- Kosashvili Y, Hiss J, Davidovic N, et al. Influence of personal body armor on distribution of entry wounds: lessons learned from urban-setting warfare fatalities. *J Trauma*. 2005;58:1236–1240.
- Peleg K, Rivkind A, Aharonson-Daniel L, et al. Does body armor protect from firearm injuries? *J Am Coll Surg*. 2006;202:6430–6438.
- Kragh JF Jr, Walters TJ, Baer DG, et al. Survival with emergency tourniquet use to stop bleeding in major limb trauma. *Ann Surg*. 2009 Jan;249(1):1–7.
- Covey DC. Blast and fragment injuries of the musculoskeletal system. *J Bone Joint Surg Am*. 2002;84-A:1221–1234.
- DePalma RG, Burris DG, Champion HR, Hodgson MJ. Blast injuries. *N Engl J Med*. 2005;352:1335–1342.
- Petersen K, Riddle MS, Danko JR, et al. Trauma-related infections from battlefield casualties from Iraq. *Ann Surg*. 2007;245:803–811.
- Hawksworth JS, Broewen TS, Keiser P, et al. War wound infection is associated with a systemic and wound tissue inflammatory profile. *J Surg Res*. 2009;151(2):300.
- Burd RS, Ouyang M, Madigan D. Bayesian logistic injury severity score: a method for predicting mortality using international classification of disease-9 codes. *Acad Emerg Med*. 2008;15(5):466–475.
- Ho KM, Knuiman M. Bayesian approach to predict hospital mortality of intensive care readmissions during the same hospitalisation. *Anaesth Intensive Care*. 2008;36(1):38–45.
- Fazio VW, Tekkis PP, Remzi F, Lavery IC. Assessment of operative risk in colorectal cancer surgery: the Cleveland Clinic Foundation colorectal cancer model. *Dis Colon Rectum*. 2004;47(12):2015–2024.
- Edwards FH, Peterson RF, Bridges C, Ceithaml EL. Use of a Bayesian statistical model for risk assessment in coronary artery surgery. *Ann Thorac Surg*. 1995;59(6):1611–1612.
- Biagioli B, Scolletta S, Cevenini G, Barbini E, Giomarelli P, Barbini P. A multivariate Bayesian model for assessing morbidity after coronary artery surgery. *Crit Care*. 2006;10(3 R94):17.
- Lenihan CR, O'Kelly P, Mohan P, et al. MDRD-estimated GFR at one year post-renal transplant is a predictor of long-term graft function. *Ren Fail*. 2008;30(4):345–352.
- Hoot N, Aronsky D. Using Bayesian networks to predict survival of liver transplant patients. *AMIA Annu Symp Proc*. 2005;345–349.
- Swets JA. *Signal Detection Theory and ROC Analysis in Psychology and Diagnostics: Collected papers*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1995.
- Dawson B, Trapp RG. *Basic and Clinical Biostatistics*. New York, NY: Lange Medical Books; 2004.
- Dunne JR, Riddle MS, Danko J, Hayden R, Petersen K. Blood transfusion is associated with infection and increased resource utilization in combat casualties. *Am Surg*. 2006;72(7):619–625; discussion 625–626.
- Forsberg JA, Elster EA, Andersen RC, et al. Correlation of procalcitonin and cytokine expression with dehiscence of wartime extremity wounds. *J Bone Joint Surg Am*. 2008;90(3):580–588.

Journal of Multidisciplinary Healthcare

Publish your work in this journal

The Journal of Multidisciplinary Healthcare is an international, peer-reviewed open-access journal that aims to represent and publish research in healthcare areas delivered by practitioners of different disciplines. This includes studies and reviews conducted by multidisciplinary teams as well as research which evaluates the results or conduct of such teams or

Submit your manuscript here: <http://www.dovepress.com/journal-of-multidisciplinary-healthcare-journal>

healthcare processes in general. The journal covers a wide range of areas and welcomes submission from practitioners at all levels, from all over the world. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.