World J Surg (2019) 43:2734–2739
https://doi.org/10.1007/s00268-019-05087-8

ORIGINAL SCIENTIFIC REPORT

Predictors of Death in Necrotizing Skin and Soft Tissue Infection

Emanuel Eguia[1] · Vincent Vivirito[1] · Adrienne N. Cobb[1] · Haroon Janjua[2] · Matthew Cheung[1] · Paul C. Kuo[2]

Published online: 16 July 2019

Abstract

Background Necrotizing skin and soft tissue infection (NSTI) is a surgical emergency that is associated with high morbidity and mortality. This study aims to identify predictors of in-hospital death following a NSTI.

Material and methods We queried the Healthcare Cost and Utilization Project (HCUP) State Inpatient Database (SID) for California between 2006 and 2011. We used conventional and advanced statistical methods to identify predictors of in-hospital mortality, which included: logistic regression, stepwise logistic regression, decision trees, and *K*-nearest neighbor (KNN) algorithms.

Results A total of 10,158 patients had a NSTI. The full and stepwise logistic regression models had a ROC AUC in the validation dataset of 0.83 (95% CI [0.80, 0.86]) and 0.81 (95% CI [0.78, 0.83]), respectively. The KNN and decision tree model had a ROC AUC of 0.84 (95% CI [0.81, 0.85]) and 0.69 (95% CI [0.65, 0.72]), respectively. The top predictors of in-hospital mortality in the KNN and stepwise logistic model included: (1) the presence of in-hospital coagulopathy, (2) having an infectious or parasitic diagnoses, (3) electrolyte disturbances, (4) advanced age, and (5) the total number of beds in a hospital.

Conclusion Patients with a NSTI have high rates of in-hospital mortality. This study highlights the important factors in managing patients with a NSTI which include: correcting coagulopathy and electrolyte imbalances, treating underlying infectious processes, providing adequate resources to the elderly population, and managing patients in high-volume centers.

Abbreviations

MDC	Major diagnosis criteria
HCUP SID	Healthcare Cost and Utilization Project State Inpatient Database
ICD-9	International Classification Disease Version 9 Diagnosis Code
CDC	Center for Disease Control and Prevention

Presented as an oral poster presentation at the Central Surgical Association and Midwest Surgical Association 2017 Annual Meeting, July 30, 2017, in Chicago, IL.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00268-019-05087-8) contains supplementary material, which is available to authorized users.

✉ Emanuel Eguia
emanuel.eguia@lumc.edu

[1] Department of Surgery, Burn Shock Trauma Research Institute, Loyola University Medical Center, 2160 S. 1st Avenue, Maywood IL 60153, USA

[2] Department of Surgery, University of South Florida, Tampa FL USA

Introduction

Necrotizing skin and soft tissue infection (NSTI) is a surgical emergency that requires early diagnosis and treatment. There are approximately 500–1500 cases of NSTI per year, and a large proportion of them are due to group A Streptococcus [1]. Studies have shown that many patients who develop a NSTI have immune-suppressing comorbidities, such as diabetes mellitus, cirrhosis, chronic corticosteroid use, alcohol abuse, and malnutrition [2]. Despite the advances in caring for critically ill patients and diagnostic tools, NSTI still portends high morbidity and mortality [3]. Therefore, there is a critical need to understand predictors that can mitigate in-hospital death.

To date, there is a limited number of studies that have used machine learning techniques to predict in-hospital mortality in patients with a NSTI. Machine learning is a category of algorithms that allows software applications to become more accurate in predicting outcomes without being explicitly programmed. Traditionally, logistic regression is used for predictive modeling. Previous comparison studies have suggested that machine learning methods can be more accurate than traditional logistic regression [4, 5]. This study aims to identify patient and hospital characteristics predictive of in-hospital mortality following a NSTI.

Methods

Study design

We conducted a retrospective cohort study using the Healthcare Cost and Utilization Project (HCUP) State Inpatient Database (SID) for California between the years 2006 and 2011. As cited in our previous studies [6, 7], HCUP is an administrative dataset composed of a family of healthcare databases developed through a Federal-State-Industry partnership sponsored by the Agency for Healthcare Research and Quality (AHRQ). Each SID captures all inpatient discharges at non-federal facilities for the respective state, regardless of primary payer.

Study population

The analysis cohort included adult patients 18 years or older. Cases of NSTI were identified based on a primary diagnosis *International Classification of Diseases, Ninth Revision* (ICD-9) code of 728.86 during their index admission. This ICD-9 code was exclusively used to reduce inaccuracies in diagnosing patients with a NSTI. Only the first listed ICD-9 code 728.86 was used to identify patients who were admitted for this particular principal diagnosis. The principal diagnosis is defined in HCUP as the condition mainly responsible for causing the admission of the patient to the hospital for care. The subsequent diagnosis codes are those associated with the principal diagnosis. These additional codes were also used for our analysis.

Analytic approach

Baseline characteristics were presented as means and standard deviations, medians, and interquartile ranges, or counts and percentages. Unadjusted comparisons of two or more proportions between those who died and lived were performed using a Chi-squared test, and continuous variables were compared using *t*-tests or Wilcoxon rank sum tests as appropriate. The primary outcome of interest was predicting in-hospital death by analyzing the predictive ability of several variable selection techniques, such as logistic regression, stepwise logistic regression, decision tree models, and *K*-nearest neighbor (KNN) model. Candidate variables for evaluation included the following: (1) age, (2) sex, (3) total hospital beds, (4) presence of a full-time nursing staff, (5) residents, (6) advanced imaging capabilities, (7) comorbidities, (8) major diagnosis category (MDC), (9) diagnosis ICD-9 codes, (10) procedures ICD-9 codes, and (11) intensive care unit (ICU) capacity. Comorbidities used in the analysis included the following: AIDS, alcohol abuse, chronic anemia, arthritis, congestive heart failure, chronic obstructive pulmonary disorder, history of coagulopathy, depression, diabetes, drug abuse, hypertension, hypothyroidism, cirrhosis, metastatic cancer, obesity, peripheral vascular disease, history of weight loss, renal failure, heart valve disease, peptic ulcer, pulmonary hypertension, and neurological disorders. The MDC and diagnosis-related group (DRG) variables are provided in HCUP to help categorize other principal diagnoses into exclusive diagnosis areas during the patient's hospitalization. Please refer to supplemental Table 1 for a full list of variables used in our analysis.

The data were split into training (80%) and testing (20%) subsets for model development and performance evaluation. Dividing the dataset into training and testing subsets gives the ability to test the consistency of the model performance. The best performing model was chosen by considering discrimination (predictive ability) with the receiver operating characteristic (ROC) area under the curve (AUC) and the number of predictors (model fit). The variable importance function was used to extract a list of the most critical predictors impacting in-hospital death.

A single cart classification decision tree model was used for this analysis, and a classic machine learning methodology was used where the dataset was split into training and testing cohorts. Model performance was evaluated on

the training set, and consistency in performance was validated on the testing dataset. For the logistic regression model, all predictors were included in the model and a stepwise logistic regression was used to create a more parsimonious model. The value of K for the KNN model was based on the modeling results (KNN plot) returned by the KNN algorithm, which gives a graphical representation of the values of K plotted against the accuracy of the KNN model. We choose the value of K ($K = 21$) for which the accuracy was highest. We chose decision tree models and KNN given the versatility of these methods to be used for classification problems, the number of data points that were being evaluated, easy interpretability, calculation time, and predictive power. This study was evaluated and approved by the Institutional Review Board at Loyola University, Chicago. All statistical analyses were performed using R statistical analysis software, version 3.4.3, and R Studio version 1.1.419.

Tree-based and *K*-nearest neighbor models

As described in previous studies [8], decision tree models partition data by splitting the variables at discrete cut-points. For our study, the decision tree algorithm determined the best variable and location for each split using the Gini Index. Also, a cost complexity parameter, using ten-fold cross-validation, was used to control the size of the final tree. A KNN model uses local geographic information in the predictor space to predict the outcomes of a new sample [5].

Results

Patient characteristics

The demographics and characteristics of patients who had an NSTI and either survived or died are summarized in Table 1. A total of 10,158 patients met our inclusion criteria. 8850 (87%) patients survived, and 1307 (13%) died in the hospital. On univariate comparison, compared to NSTI patients that survived, patients that died tended to be older (58 vs. 52, $p < 0.01$), had a higher frequency of drug abuse (24 vs. 16%, $p < 0.01$), and had a higher frequency of renal failure (25 vs. 16%, $p < 0.01$). There were no significant differences in sex, race, and ethnicity between the two groups (Table 1).

Predictive probability of statistical techniques

The predictive probabilities of the statistical methods used in our analysis are summarized in Fig. 1. We hypothesized that machine learning techniques (KNN and decision trees) would be better at predicting in-hospital mortality than conventional statistical methods (logistic regression). The full and stepwise logistic regression models had a ROC AUC in the validation dataset of 0.83 (95% CI [0.80, 0.86]) and 0.81 (95% CI [0.78, 0.83]), respectively (Fig. 1, Supplemental Tables 2, 3). The KNN and decision tree model had a ROC AUC of 0.84 (95% CI [0.81, 0.85]) and 0.69 (95% CI [0.65, 0.72]), respectively (Fig. 1, Supplemental Fig. 1).

Predictors of in-hospital mortality in patients with a NSTI

To better understand the most important factors in predicting in-hospital death, in patients with a NSTI, we examined the common predictors of the KNN and stepwise logistic regression models. These two models had similar predictive abilities and fit. The common predictors between the two models included the following: coagulopathy, infectious or parasitic disease, electrolyte abnormalities, advanced age, iron deficiency anemia, psychoses, drug abuse, weight loss, depression, and total hospital beds (Fig. 2 and Supplemental Table 3).

Discussion

The treatment of NSTI rests upon aggressive surgical exploration and debridement of necrotic tissue in addition to broad-spectrum empiric antibiotic therapy. Despite aggressive surgical and pharmacological management of NSTI, it is still associated with very high mortality [9–11]. The objective of our study was to examine hospital- and patient-level factors associated with in-hospital death. Very few studies have applied predictive analytic techniques such as machine learning to identify risk factors that may be associated with in-hospital mortality in patients who present with a NSTI. The advantage of using machine learning algorithms is the ability to develop predictive models which can generate hypotheses and supplement conventional statistical methods. We had several interesting findings. We found that KNN performed similarly to the stepwise logistic regression model in predicting in-hospital mortality in patients with a NSTI. Also, we identified several common patient and hospital level variables between our models that are predictive of in-hospital mortality, which included in-hospital coagulopathy, electrolyte abnormalities, advanced age, and the total number of hospital beds. The consistency between the two models adds credibility and importance to the factors identified in the two models.

One of the main findings in our study was that the total number of hospital beds was predictive of in-hospital

Table 1 Demographics of patients with a NSTI

	Survived		Died		*p*
No. of patients, *n* (%)	8850	87%	1307	13%	
Age, mean (sd)	52	15	58	15	<.01
Male, *n* (%)	5343	60.4%	764	58.5%	0.68
Race/ethnicity, *n* (%)					
White	4290	48.5%	649	49.7%	0.81
Black	830	9.4%	114	8.7%	
Hispanic	2811	31.8%	399	30.5%	
Other	919	10.4%	145	11.1%	
Comorbidities, *n* (%)					
Alcohol abuse	741	8.4%	32	2.4%	<.01
Diabetes	2421	27.4%	298	22.8%	
Drug abuse	1406	15.9%	311	23.8%	
Hypertension	4331	48.9%	565	43.2%	
Obese	2030	22.9%	207	15.8%	
Renal failure	1448	16.4%	331	25.3%	
Chronic anemia	3215	36.3%	340	26.0%	
Coagulopathy	828	9.4%	446	34.1%	
Electrolyte abnormalities	3510	39.7%	807	61.7%	
Psychoses	442	5.0%	94	7.2%	
Weight loss	1391	15.7%	249	19.1%	
Depression	838	9.5%	66	5.0%	
Infectious or parasitic disease, *n* (%)	2270	25.6%	759	58.1%	
Hospital beds, mean (SD)	312	173	333	182	<.01

NSTI Necrotizing soft tissue infection

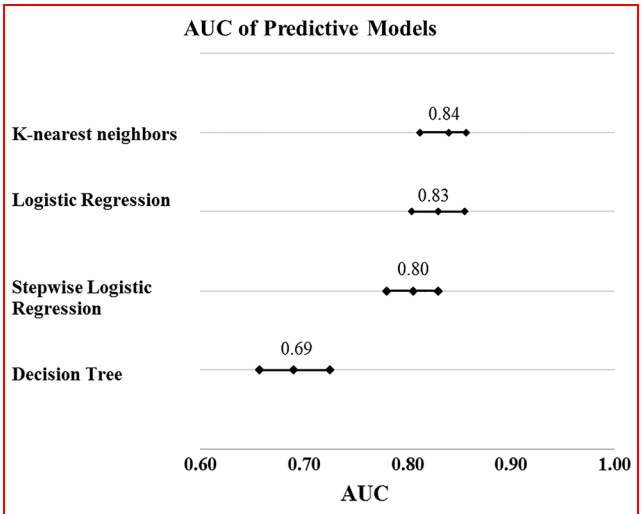

Fig. 1 The area under receiver operator characteristic curve (AUC) of the compared methods for in-hospital death in the necrotizing fasciitis in the validation cohort

death. The findings showed that hospitals with a lower number of hospital beds had a higher rate of in-hospital death among patients with a NSTI. Therefore, smaller hospitals or healthcare systems may not have adequate resources in caring for patients with a NSTI, which may lead to poor patient outcomes. A retrospective cohort study of patients with a NSTI revealed that as the time between admission and surgical intervention increased, mortality rate also increased [12, 13]. The association between time lapse in treatment and increased mortality may be in part due to smaller hospitals not having the necessary resources to promptly diagnose and treat patients with a NSTI. A study by Ogola et al. examined the 2010 national inpatient sample data and found that emergency general surgery patients treated in healthcare systems with a higher volume of critically ill patients experience lower mortality rates [14]. This leads us to believe that patients who are suspected of having a NSTI should be transferred to centers that have the appropriate infrastructure and experience caring for patients with suspected NSTI.

In our study, we were also able to identify patient-level predictors that contribute to in-hospital death. Some of

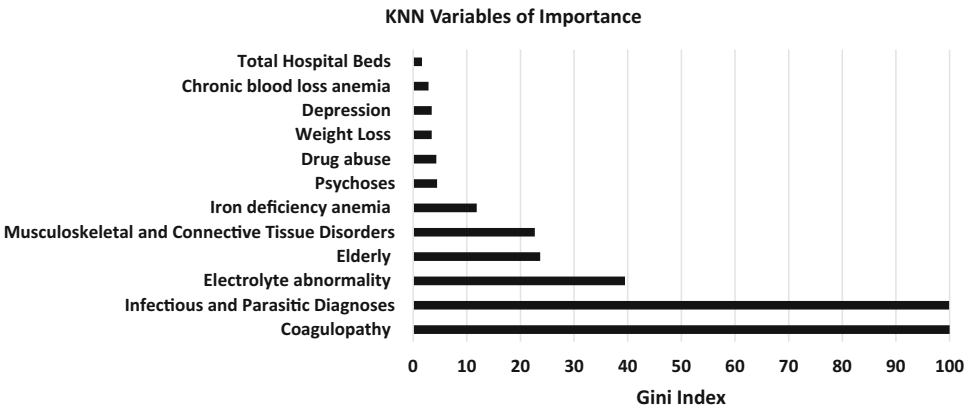

Fig. 2 A variable importance plot for the *K*-nearest neighbor model. Variable importance is computed using the mean decrease in Gini index and expressed relative to the maximum of 100

these factors include coagulopathy which may lead to bleeding, electrolyte imbalances due to fluid shifts, anemia due to chronic disease, and history of drug abuse which may lead to increased exposure to pathogens via unsterile intravenous injections. Prior studies have found that electrolyte derangements, lactic acid levels, and comorbidities may portend to poor outcomes in patients with a NSTI [15]. In a retrospective cohort study of patients with a NSTI, variables such as admission and serum sodium were found to be predictive of mortality [16]. Furthermore, Lee et al. identified risk factors associated with increased mortality in patients with a NSTI, such as liver cirrhosis, chronic liver dysfunction, chronic renal failure, serum albumin less than 3 g/dl, thrombocytopenia, serum creatinine greater than 1.6 mg/dl, postoperative mechanical ventilation, and increasing days in the intensive care unit [17].

There are a limited number of studies which have used machine learning techniques in conjunction with conventional statistical methods to identify common factors that lead to in-hospital mortality in patients with a NSTI. In our study, we found that the KNN and stepwise logistic regression models had similar predictive abilities and common predictors in identifying patients with a NSTI who are at risk of in-hospital death. Our study reveals the advantages of using machine learning techniques and conventional statistical methods when analyzing Big Data. Using these two methods of analysis reveals the consistency between the two models and adds credibility to the importance of the factors that were identified. Future studies using more robust machine learning algorithms, neural networks, natural language processing, or hierarchically well-formulated logistic models are needed to compare advanced statistical methods. Also, we believe further studies are needed using electronic healthcare record data to develop real-time predictive models that can help mitigate poor outcomes in patients with a NSTI.

This study has several limitations, including the use of administrative datasets that do not contain detailed clinical information such as laboratory results and physiologic measures, which may contribute to in-hospital mortality. Furthermore, there may be inherent under or over-coding of ICD-9 codes. Despite some of these limitations, the use of randomly sampled training and validation datasets with the available data helps mitigate some of these issues. Due to limited availability by the HCUP California SID, the study only includes data from 2006 to 2011 and therefore cannot be extrapolated to the entire US population or more recent trends. Given these limitations, future studies are needed using more granular electronic healthcare record data to improve our prediction model.

Conclusion

Patients with a NSTI have high rates of in-hospital mortality. Attention should be given to the correction of in-hospital coagulopathy, electrolyte imbalances, control of infectious processes, the elderly population, and the number of beds in a hospital. Also, transferring patients with a suspected NSTI to hospitals that provide higher levels of acute care should be considered.

Author contributions Cheung and Cobb acquired data. Analysis and interpretation of data were performed by Eguia, Cobb, Cheung, Janjua, and Kuo. Drafting of the manuscript was done by Eguia, Cheung, Janjua, Cobb, and Kuo. Lastly, the critical revision was performed by Eguia, Cobb, Cheung, Janjua, and Kuo.

Funding This work was supported by the National Institute of Health 5 T32 GM008750-18.

Compliance with ethical standards

Conflict of interest No conflicts of interest to disclose among the authors.

References

1. Hussein QA, Anaya DA (2013) Necrotizing soft tissue infections. Crit Care Clin 29:795–806
2. Angoules AG, Kontakis G, Drakoulakis E et al (2007) Necrotising fasciitis of upper and lower limb: a systematic review. Injury 38:S25
3. May AK (2009) Skin and soft tissue infections. Surg Clin North Am 89:403–420
4. Fernández-Delgado M, Cernadas E, Barro S et al (2014) Do we need hundreds of classifiers to solve real world classification problems? J Mach Learn Res 15:3133–3181
5. Churpek MM, Yuen TC, Winslow C et al (2016) Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. Crit Care Med 44:368
6. Eguia E, Cobb A, Kothari A et al (2018) Impact of the Affordable Care Act (ACA) Medicaid expansion on cancer admissions and surgeries. Ann Surg 268(4):584–590
7. Eguia E, Cobb AN, Kirshenbaum EJ et al (2018) Racial and ethnic postoperative outcomes after surgery: the Hispanic paradox. J Surg Res 232:88–93
8. Kuhn M, Johnson K (2013) Applied predictive modeling. Springer, Berlin
9. Darenberg J, Luca-Harari B, Jasir A et al (2007) Molecular and clinical characteristics of invasive group A streptococcal infection in Sweden. Clin Infect Dis 45:450–458
10. StevensDl T, Winship J, SwartsR RKM et al (1989) Severe group A streptococcal infections associated with a toxic shock-like syndrome and scarlet fever toxin A. N Engl J Med 321(1):7
11. Kaul R, McGeer A et al (1997) Population-based surveillance for group A streptococcal necrotizing fasciitis: clinical features, prognostic indicators, and microbiologic analysis of seventy-seven cases. Am J Med 103:18–24
12. Wong C-H, Chang H-C, Pasupathy S et al (2003) Necrotizing fasciitis: clinical presentation, microbiology, and determinants of mortality. JBJS 85:1454–1460
13. Chao W-N, Tsai C-F, Chang H-R et al (2013) Impact of timing of surgery on outcome of *Vibrio vulnificus*-related necrotizing fasciitis. Am J Surg 206:32–39
14. Ogola GO, Haider A, Shafi S (2017) Hospitals with higher volumes of emergency general surgery patients achieve lower mortality rates: a case for establishing designated centers for emergency general surgery. J Trauma Acute Care Surg 82:497–504
15. Wong C-H, Khin L-W, Heng K-S et al (2004) The LRINEC (Laboratory Risk Indicator for Necrotizing Fasciitis) score: a tool for distinguishing necrotizing fasciitis from other soft tissue infections. Crit Care Med 32:1535–1541
16. Yaghoubian A, de Virgilio C, Dauphine C et al (2007) Use of admission serum lactate and sodium levels to predict mortality in necrotizing soft-tissue infections. Arch Surg 142:840–846
17. Lee C-Y, Kuo L-T, Peng K-T et al (2011) Prognostic factors and monomicrobial necrotizing fasciitis: gram-positive versus gram-negative pathogens. BMC Infect Dis 11:5