Seeing the Forest Beyond the Trees: Predicting Survival in Burn Patients with Machine Learning

Adrienne N. Cobb, M.D[a,b], **Witawat Daungjaiboon, M.S.**[c], **Sarah A. Brownlee, B.A.**[b], **Anthony J. Baldea, M.D.**[a], **Arthur P. Sanford, M.D.**[a], **Michael M. Mosier, M.D., FACS**[a], and **Paul C. Kuo, M.D.**[a,b]

[a]Loyola University Medical Center, Department of Surgery, 2160 S. 1st Avenue Maywood, IL 60153, USA

[b]One:MAP Section of Surgical Analytics, Department of Surgery, Loyola University, Chicago, 2160 S. 1st Avenue Maywood, IL 60153, USA

[c]DePaul University,College of Computing and Digital Media, Department of Predictive Analytics, Chicago, IL, 243 South Wabash Avenue Chicago, IL 60604

Abstract

Background—This study aims to identify predictors of survival for burn patients at the patient and hospital level using machine learning techniques.

Methods—The HCUP SID for California, Florida and New York were used to identify patients admitted with a burn diagnosis and merged with hospital data from the AHA Annual Survey. Random forest and stochastic gradient boosting (SGB) were used to identify predictors of survival at the patient and hospital level from the top performing model.

Results—We analyzed 31,350 patients from 670 hospitals. SGB (AUC .93) and random forest (AUC .82) best identified patient factors such as age and absence of renal failure (p<0.001) and hospital factors such as full time residents (p<0.001) and nurses (p=0.004) to be associated with increased survival.

Conclusions—Patient and hospital factors are predictive of survival in burn patients. It is difficult to control patient factors, but hospital factors can inform decisions about where burn patients should be treated.

Corresponding Author: Paul C. Kuo, MD, MS, MBA, FACS, Chairman, Department of Surgery, Loyola University Medical Center, 3rd Floor EMS Building, Health Sciences Campus, 2160 S. First Avenue, Maywood, IL 60153, paul.kuo@luhs.org, Phone: 708-327-2705, Fax: 708-327-2852.

Keywords

INTRODUCTION

Burn patients require complex care involving a delicate balance among resuscitation, stabilization and rehabilitation. Their injuries can range from superficial burns only requiring local wound care to more severe burns that may require surgery and are potentially complicated by respiratory failure or sepsis. The predictors of outcome in burn patients are well established, correlating well with the Baux index of age and percent surface area burned as well as the revised Baux index which also takes into account inhalation injury.[1–4] Therefore, multiple factors must be considered when caring for these patients. Traditionally, the focus has been on preventing mortality by controlling or treating as many patient specific or disease specific factors as possible. Scoring tools such as the Baux Score or the Abbreviated Burn Severity Index (ABSI) have been designed to predict morbidity and mortality based on total body surface area (TBSA) age, sex and the presence of inhalation injury.[5] While patient characteristics are key, there are other factors that contribute to patient outcomes.

Studies suggest that system characteristics within institutions such as staffing and technology can impact the ability of hospitals to provide optimal care for patients. [6–7] Since many studies examining the prediction of survival in burn patients are completed in a single center, there is little data surrounding what systems characteristics may have contributed to the survival of burn patients. The use of a large all payer, administrative database linked to hospital level data provides a different perspective. Additionally, the use of machine learning allows us to uncover patterns or associations not otherwise elucidated with traditional linear statistical techniques. A study utilizing artificial neural networks to predict survival in burn patients found non-linear techniques better suited to address complex questions regarding prognosis due to their ability to "observe" the real events or facts then evaluate the relative influence of variables on each other and the whole case.[8]

The present study addresses two understudied topics: factors that predict the survival of burn patients beyond traditional burn specifications and the contribution of systems infrastructure by examining hospital characteristics that predict survival. We identified a heterogeneous group of burn patients, created various models to predict survival based on patient and hospital characteristics, and chose the model that performed best with the goal of informing clinical decision-making.

METHODS

Data Source and Patient Selection

The Healthcare Cost and Utilization Project (HCUP) State Inpatient Database (SID) for California 2006–2011, Florida 2009–2013 and New York 2009–2013 was used to identify adult patients admitted to the hospital with a burn diagnosis. Burns of varying severity and

location were included. The SID is an administrative, all-payer data set aggregated by the Agency for Healthcare Research and Quality (AHRQ) to inform health related decisions.[9] The diagnosis codes were identified by *International Classification of Disease, 9th Revision, Clinical Modification*, (ICD9) codes for burn injury (941.20–59, 942.20–35, 942.39–45, 942.49–59, 943.20–26, 943.29–36, 943.39–46, 943.49–56, 943.59, 944.20–28, 944.30–38, 944.40–48, 944.50–58, 945.20–26, 945.29–36, 945.39–46, 945.49–56, 945.59, 946.2–5, 947.0–4, 947.8–9). These data were then merged with the 2011 American Hospital Association (AHA) Annual Survey to provide hospital level data associated with the selected burn patients. This nationwide database contains information categorizing an institution's organizational structure, facility and service lines, operation expenses, and staffing. [10] The Institutional Review Board at our institution deemed the study exempt from review as the data are de-identified, protected and publically available.

Data Pre-Processing

Patient level data were pre-processed to provide uniform variable formats across states using the dplyr package in R. It provides a flexible grammar of data manipulation and focuses on tools for working with data frames. Variables selected for analysis included various comorbidities, age, mortality, hospital state, insurance type, procedure codes, race, admission type, and median income quartile. Several variables were generated in the hospital level AHA dataset including Joint Commission designation, Commission on Cancer, Council of Teaching Hospitals, Level 1 Trauma Center, Nurse to Bed Ratio, surgical volume, GI intensity, radiology intensity, and ICU beds.

Missing values for both groups were replaced with column means for numerical variables and Random forest algorithm was used to impute categorical variables using the caret package. Data were split into training (66%) and test (34%) sets. Our target variable DIED (mortality=1, survival=0) was extremely imbalanced in both datasets. To avoid not detecting the minority class, we used Synthetic Minority Over-sampling Technique (SMOTE) to balance both groups by up-sampling the minorities (DIED=1) and down sampling majorities (DIED=0).

Statistical Analysis

Descriptive statistics of the study population were calculated using arithmetic means with standard deviations or median with interquartile range for continuous variables and proportions for categorical variables. Population unadjusted mortality was obtained using a simple proportion of number of inpatient mortalities by the total population. Age categories by 7-year intervals were created and plotted against mortality rates as seen in Figure I. Additionally, mortality rates were calculated by age range and burn type (Table II) as seen in Figure II. Multiple models were built to determine the model best suited to predict variables that impact survival in burn patients. Accuracy, sensitivity and specificity were used to evaluate the models for completeness. We also used receiver operator curve (ROC) for model comparison. The ROC demonstrates how well models separate both classes while the area under the curve (AUC) can be interpreted as the accuracy of the models. The AUC ranges from poor class separation at 0.51 to perfect class separation at 1.

We used tree-based ensemble models, such as stochastic gradient boosting (SGB) and random forest (RF), as we are able to use a variable importance measure to determine those factors that affect patient survival. The variable importance measure was used to indicate how well each variable split our target class. The stochastic gradient boosting algorithm was run with 650 trees, an interaction depth of 9, shrinkage of 0.1 and a minimum of ten observations per node. Random forest models were run with a weight class of 1:3. All analysis were completed using various packages in RStudio including randomForest, party, caret, and pROC and as mentioned above.

Gradient Boosting Machine (GBM) uses a boosting method to build ensemble trees by iteratively adding a weak classifier one at a time (in this case tree stump). In each iteration, a new tree tries to correct errors in the model from the previous iteration. New trees are added until we reach the goal of prediction. Unlike the GBM, the random forest algorithm uses bootstrap aggregation, also known as bagging, to construct a model by creating trees from sampling data from a training set with replacement and subsequently combines the trees together.

RESULTS

We analyzed 31,350 patients from 670 hospitals across the three states included. The mean patient age was 40.5 years. The study population was largely male, Caucasian, and had Medicaid insurance. Hypertension was the most common comorbidity with 24.6% of the study population affected. Baseline patient and hospital characteristics are summarized in Table I. The age distribution varied in the study population. The overall mortality rate was 2.86%, and the rate of mortality increased with age (Figure I). Patients with burns from multiple specified sites (ICD-9 946) had the highest mortality rate among burn patients, while unspecified sites and burns to the eyes carried no mortality (Figure II).

At the patient level, we found that the stochastic gradient boosting model performed slightly better than the random forest algorithm with an area under the ROC curve of .93 and .90 respectively. Although, the weighted random forest model tended to pick up the minority class better than the SGB as evidenced by its superior specificity value (0.74 v. 0.71). The top five patient characteristics, as evidenced by their variable importance score, predicting survival in burn patients according to the SBG model were younger age, absence of electrolyte imbalance or coagulopathy, admission on a weekend, and absence of renal failure (Figure III). The top five patient characteristics that predicted survival in the random forest model were absence of electrolyte imbalance or coagulopathy, younger age, absence of congestive heart failure, and presence of weight loss. All were predicted with p<0.001.

At the hospital level, the random forest algorithm far outperformed SGB with an AUC of 0.82 compared to AUC 0.62. With a specificity of 0.61, the random forest model is able to predict survival correctly for more than 60% of the test set, while SGB only predicts survival correctly for 1% of the observations with a specificity of 0.006. Given its poor performance, the variables deemed important by the SGB model were not included here. The top five hospital characteristics predicting survival in burn patients per the RF model, were the presence of full-time residents (p<0.001) and nursing staff (p=0.004), availability of

inpatient MRI, high surgical volume (p=0.022), and increased number of medical/surgical intensive care unit beds (p=0.0002). Their variable importance is graphically represented in Figure IV.

DISCUSSION

Due to steady improvements in the management of burn patients, mortality after burn injury has decreased over time, even for those with severe burn injuries. Predictors of morality such as the extent of the burn, increasing age, and more recently the presence of inhalation injury are well studied.[11–13] However, in this era of quality improvement and patient centered outcomes, there is pressure for hospitals to find novel methods to improve healthcare quality.

Patient Level

This study demonstrated that patient characteristics such as age, admission on a weekend, absence of coagulopathy, electrolyte imbalance, and renal failure predict survival in burn patients using stochastic gradient boosting modeling. Age is a well know predictor of mortality, with increasing age leading to increased mortality. However, age likely does not tell the whole story. Moving forward, as the assessment of frailty becomes commonplace, we may see the absence of frailty emerge as a predictor of survival. Romanowski et al conducted a study on elderly burn patients and found that a low frailty score was an independent predictor of survival.[14] It was surprising to note that admission on a weekend would lead to survival as the "weekend effect" has been well established.[15–17] Avoiding coagulopathy and correcting electrolyte imbalances in burn patients is a constant battle. The results of this study provide further evidence of the need for protocol driven treatment plans not only in the realm of wound care, physical therapy, and nutrition, but also in the fundamental critical care areas like fluid balance and temperature regulation.

Hospital Level

The random forest algorithm performed best at determining hospital characteristics such as the presence of full time residents and registered nurses, access to MRI, number of medical/surgical ICU beds and overall surgical volume predicted survival in burn patients. Components of hospital infrastructure can play an important role in improving the outcomes of burns patients beyond that which is accomplished by optimizing patient level factors. There is literature to support the impact of resource availability, adequate staffing and surgical volume on patient outcomes, independent of patient factors.[18–21]

A strength of our study is the use of predictive modeling to identify factors associated with survival in burn patients. These models allow for the evaluation of other factors that contribute to whether patients survive or die, i.e. the systems characteristics that we elucidate. Traditional scoring systems tend to focus on a few distinct patient level factors (age, %TBSA, inhalation injury) that likely don't completely account for a patient's subsequent outcome. Additionally, predictive models are derived using techniques that account for non-linear effects that may not have been revealed using traditional linear statistical techniques.

Limitations and Future Work

This study has several limitations, including use of an administrative dataset and the lack of granular information regarding burn severity. While there is information within ICD9 codes to determine TBSA, the information is often inconsistent across hospitals and states as coding practices vary greatly. However, we purposely did not narrow our study population to a particular TBSA because we wanted to have a diverse population of burn patients. The administrative database used here does not contain reliable information regarding burn depth, inhalation injury or other predictors of mortality such as the modified Baux score. Additionally, there is always a degree of selection bias in these types of studies, but we account for this by using rigorous model validation through training and test sets. Despite this, we assert that our results would likely persist even with the stratification of patients by burn severity and inhalation injury, particularly for those patients with burns of greater severity. Many of the patient factors are likely to remain the same, though it is certainly reasonable to think that both %TBSA as well as inhalation injury would have been selected by the model with high variable importance as previous research has cited these factors as contributing to increased mortality in burn patients. Furthermore, the use of randomly sampled training and testing sets provides a platform for the model to identify strong predictors in the data as is, regardless of stratification. Lastly, the AHA Annual survey does not have the ability to designate hospitals as certified burn centers. Knowing this designation would have been useful to include in the model as we know that these centers have a set of characteristics that make them unique to addressing burn injuries.

Going forward, this study should be repeated with a particular subset in mind, such as patients with burns greater than 20% TBSA as this tends to be the inflection point at which mortality increases. It would also be prudent to use another dataset, such as the ABA National Burn Repository that contains more burn specific information, although that may limit our ability to investigate the hospital characteristics that we believe also play an important role in predicting the survival of burn patients.

CONCLUSION

Our findings demonstrate that there are patient level, but more importantly hospital level factors that predict survival in patients with varying burn severity. While advances in critical care have improved outcomes for burn patients thus far, focusing on hospital resources and quality improvement may be what is needed to overcome the current plateau in burn patient survival.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Jeschke MG, Pinto R, Kraft R, Nathens AB, Finnerty CC, Gamelli RL, Gibran NS, Klein MB, Arnoldo BD, Tompkins RG, Herndon DN. Morbidity and survival probability in burn patients in modern burn care. Critical care medicine. 2015 Apr.43(4):808. [PubMed: 25559438]

2. Klein MB, Goverman J, Hayden DL, Fagan SP, McDonald-Smith GP, Alexander AK, Gamelli RL, Gibran NS, Finnerty CC, Jeschke MG, Arnoldo B. Benchmarking outcomes in the critically injured burn patient. Annals of surgery. 2014 May.259(5):833. [PubMed: 24722222]

3. Curreri PW, Luterman A, Braun DW Jr, Shires GT. Burn injury. Analysis of survival and hospitalization time for 937 patients. Annals of surgery. 1980 Oct.192(4):472. [PubMed: 7425694]

4. Osler T, Glance LG, Hosmer DW. Simplified estimates of the probability of death after burn injuries: extending and updating the Baux score. Journal of Trauma and Acute Care Surgery. 2010 Mar 1; 68(3):690–7.

5. Hussain A, Choukairi F, Dunn K. Predicting survival in thermal injury: a systematic review of methodology of composite prediction models. Burns. 2013 Aug 31; 39(5):835–50. [PubMed: 23384617]

6. Kothari AN, Blanco BA, Brownlee SA, Evans AE, Chang VA, Abood GJ, Settimi R, Raicu DS, Kuo PC. Characterizing the role of a high-volume cancer resection ecosystem on low-volume, high-quality surgical care. Surgery. 2016 Oct 31; 160(4):839–49. [PubMed: 27524432]

7. Funk LM, Gawande AA, Semel ME, Lipsitz SR, Berry WR, Zinner MJ, et al. Esophagectomy outcomes at low-volume hospitals: the association between systems characteristics and mortality. Ann Surg. 2011; 253:912–7. [PubMed: 21422913]

8. Estahbanati HK, Bouduhi N. Role of artificial neural networks in prediction of survival of burn patients—a new approach. Burns. 2002 Sep 30; 28(6):579–86. [PubMed: 12220917]

9. Healthcare Cost and Utilization Project (HCUP). Agency for Healthcare Research and Quality; Rockville, MD: HCUP Databases. Available at: www.hcup-us.ahrq.gov/sidoverview.jsp [Accessed March 15, 2017]

10. American Hospital Association. The AHA Annual Survey Database. Washington, DC: American Hospital Association; 2005.

11. Bhatia AS, Mukherjee BN. Predicting survival in burned patients. Burns. 1992 Oct 1; 18(5):368–72. [PubMed: 1445625]

12. Roberts G, Lloyd M, Parker M, Martin R, Philp B, Shelley O, Dziewulski P. The Baux score is dead. Long live the Baux score: a 27-year retrospective cohort study of mortality at a regional burns service. Journal of Trauma and Acute Care Surgery. 2012 Jan 1; 72(1):251–6. [PubMed: 22310134]

13. Taylor SL, Lawless M, Curri T, Sen S, Greenhalgh DG, Palmieri TL. Predicting mortality from burns: the need for age-group specific models. Burns. 2014 Sep 30; 40(6):1106–15. [PubMed: 24846014]

14. Romanowski KS, Barsun A, Pamlieri TL, Greenhalgh DG, Sen S. Frailty score on admission predicts outcomes in elderly burn injury. Journal of Burn Care & Research. 2015 Jan 1; 36(1):1–6. [PubMed: 25383979]

15. Sharp AL, Choi H, Hayward RA. Don't get sick on the weekend: an evaluation of the weekend effect on mortality for patients visiting US EDs. Am J Emerg Med. 2013; 31:835–837. [PubMed: 23465873]

16. Kostis WJ, Demissie K, Marcella SW, et al. Weekend versus weekday admission and mortality from myocardial infarction. N Engl J Med. 2007; 356:1099–1109. [PubMed: 17360988]

17. Zapf MA, Kothari AN, Markossian T, Gupta GN, Blackwell RH, Wai PY, Weber CE, Driver J, Kuo PC. The "weekend effect" in urgent general operative procedures. Surgery. 2015 Aug 31; 158(2):508–14. [PubMed: 26013983]

18. Blecker S, Goldfeld K, Park N, et al. Electronic health record use, intensity of hospital care, and patient outcomes. Am J Med. 2014; 127:216–221. [PubMed: 24333204]

19. Kothari AN, Zapf MA, Blackwell RH, Markossian T, Chang V, Mi Z, Gupta GN, Kuo PC. Components of hospital perioperative infrastructure can overcome the weekend effect in urgent general surgery procedures. Annals of surgery. 2015 Oct.262(4):683. [PubMed: 26366549]

20. Yasunaga H, Hashimoto H, Horiguchi H, et al. Variation in cancer surgical outcomes associated with physician and nurse staffing: a retrospective observational study using the Japanese Diagnosis Procedure Combination Database. BMC Health Serv Res. 2012; 12:129–136. [PubMed: 22640411]

21. Chukmaitov A, Harless DW, Bazzoli GJ, et al. Delivery system characteristics and their association with quality and costs of care: Implications for accountable care organizations. Health Care Manage Rev. 2015; 40:92–103. [PubMed: 24566250]

Summary sentences

This study aims to identify predictors of survival at the patient and hospital level using advanced machine learning techniques. We found that both patient and hospital factors are predictive of survival in the burn population such as age, absence of electrolyte imbalance, coagulopathy, or number of residents, full time nursing and surgical volume respectively. It is often difficult to control patient level factors, hospital factors can inform decisions about where burn patients should be treated.

Highlights

- Patient factors predicting survival included age, absence of renal failure, electrolyte abnormalities, coagulopathy and weekend admission.

- Hospital factors predicting survival included having full-time residents and nursing staff, access to advanced imaging, number of intensive care unit beds and overall surgical volume.

- Hospital factors in addition to patient clinical factors can be used to inform decisions on where burn patients should be treated.
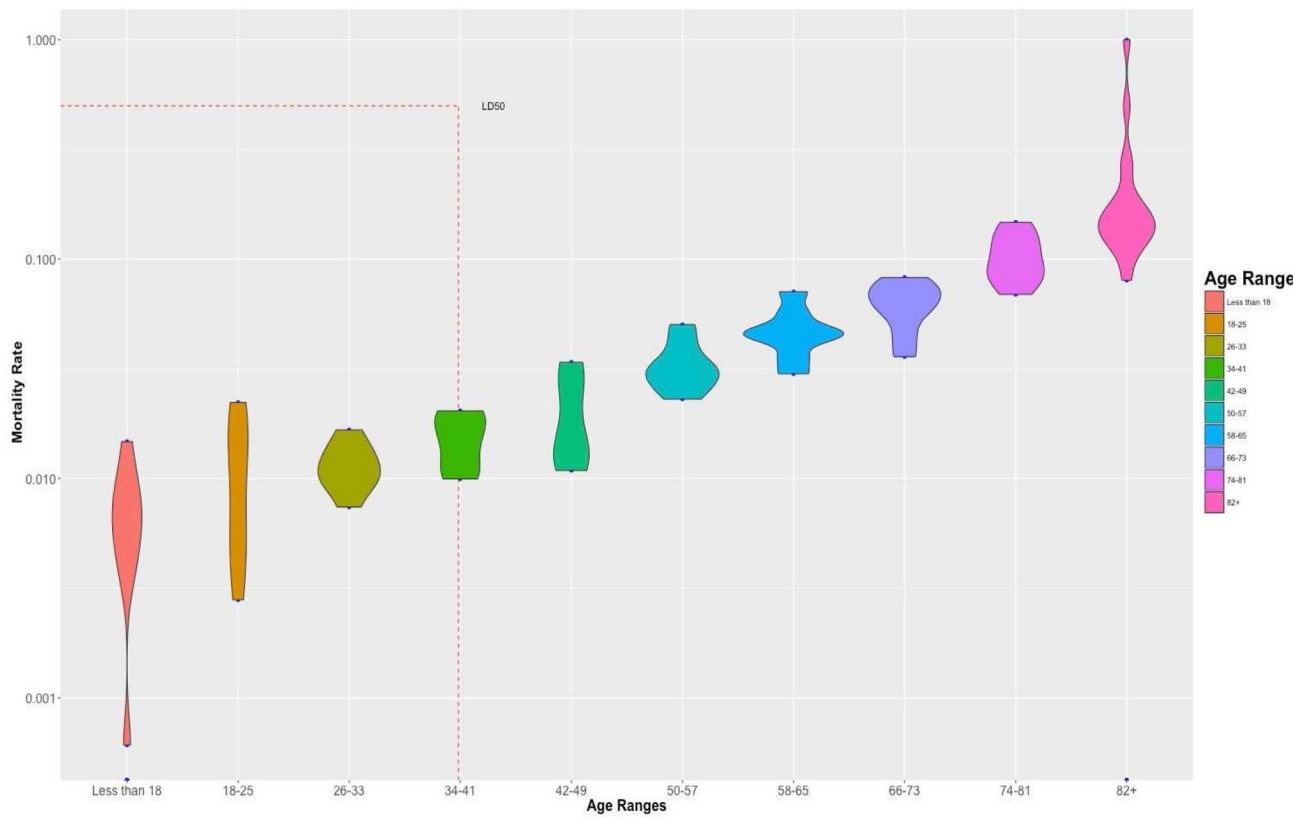
Figure I.
Violin plot with age distribution and mortality by age ranges with LD50. The shape of the violin plot shows each age range distribution. LD50=39 indicating patients at age more than 39 have 50% change of death.

*Figure to be used in color in online version and grayscale in print.

Age ranges	Burn CI 940	941	942	943	944	945	946	947	948	949	Avg. Mortality Rate
Less than 18	0.0000	0.0012	0.0036	0.0020	0.0015	0.0020	0.0435	0.0250	0.0022	0.0000	0.0000 — 0.8000
18-25	0.0000	0.0266	0.0179	0.0032	0.0015	0.0058	0.1935	0.0000	0.0000		
26-33	0.0000	0.0163	0.0261	0.0074	0.0040	0.0044	0.1463	0.0263	0.0048	0.0000	
34-41	0.0000	0.0415	0.0209	0.0087	0.0031	0.0021	0.2286	0.0000	0.0164		
42-49	0.0000	0.0332	0.0471	0.0159	0.0074	0.0072	0.1633	0.0152	0.0105	0.0000	
50-57	0.0000	0.0460	0.0812	0.0175	0.0124	0.0142	0.2941	0.0577	0.0144	0.0000	
58-65	0.0000	0.0751	0.1048	0.0242	0.0160	0.0147	0.3793	0.0256	0.0352	0.0000	
66-73	0.0000	0.0720	0.1250	0.0183	0.0273	0.0214	0.2727	0.1463	0.0769		
74-81	0.0000	0.1590	0.1644	0.0903	0.0000	0.0442	0.8000	0.1429	0.1296	0.0000	
82+	0.0000	0.2754	0.1933	0.0988	0.0357	0.0730	0.5556	0.2143	0.1918	0.0000	

Figure II.

Mortality rates by age ranges and burn type. As age increases mortality from burn injury increases. Moralities per burn type are listed vertically.

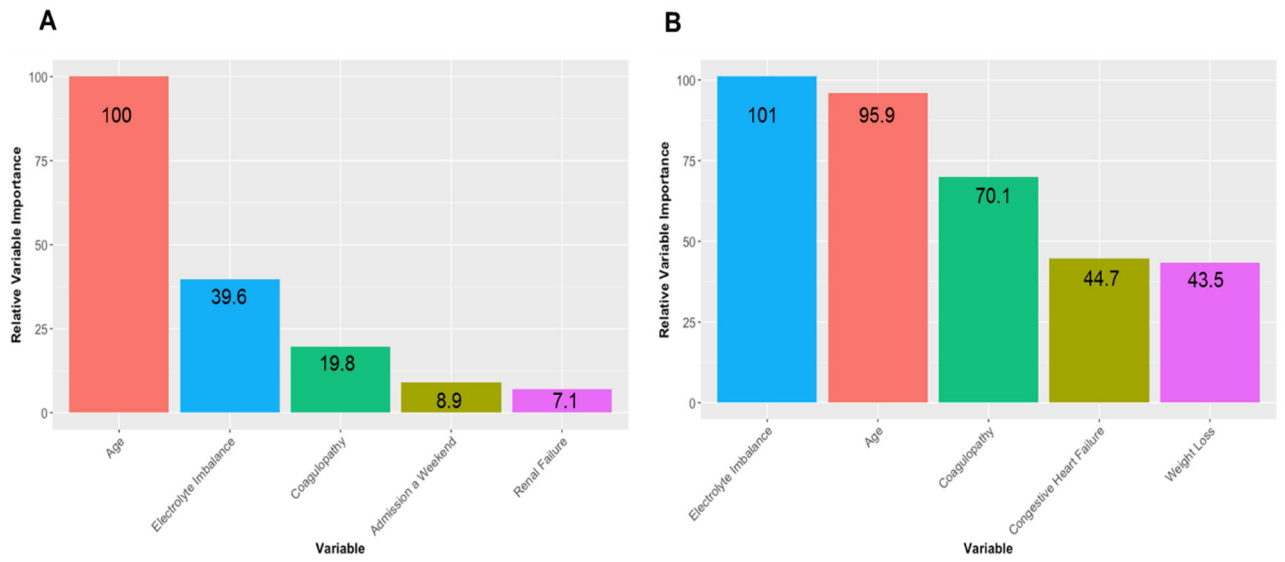*Figure to be used in color in online version and grayscale in print.

Figure III.
Variable importance for (A) stochastic gradient boosting and (B) random forest algorithms at the patient level. Variable importance is without units, but is normalized to a scale of 1 to 100, with the larger number indicating increased importance.

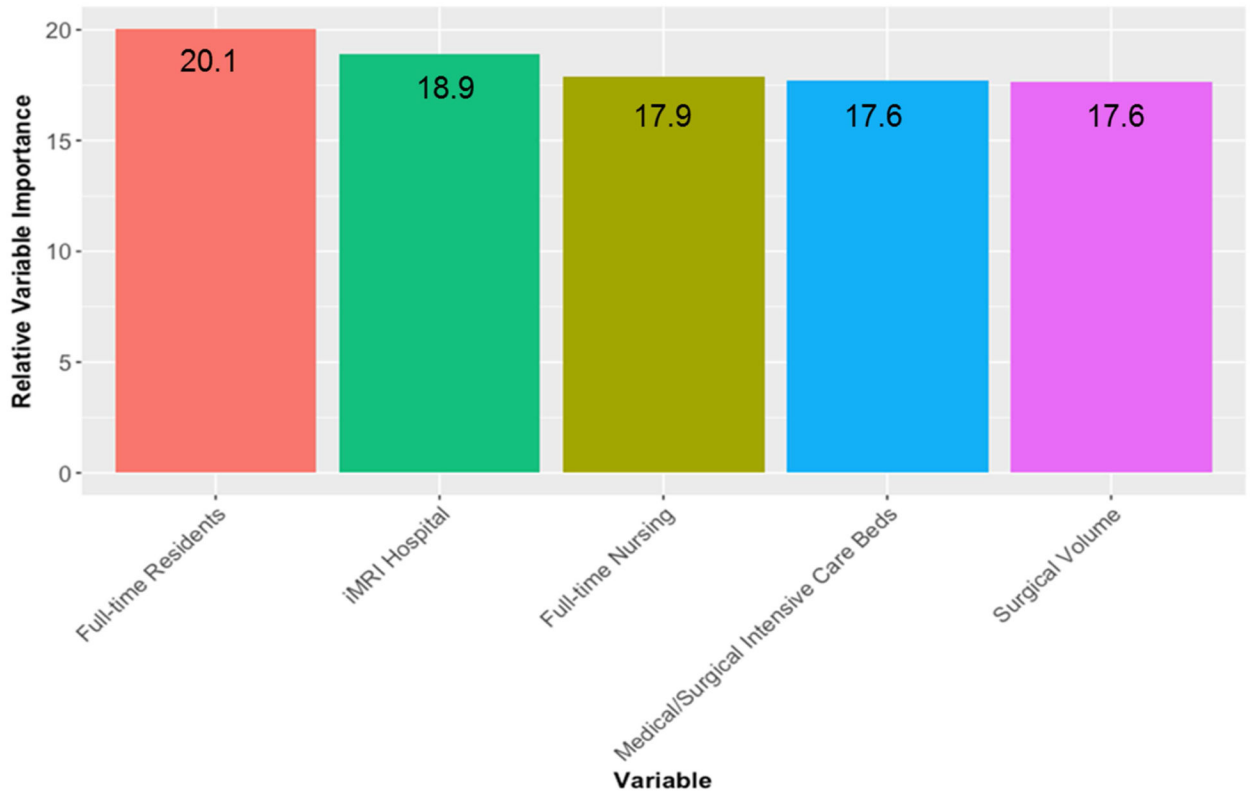*Figure to be used in color in online version and grayscale in print.

Figure IV.
Random forest variable importance at the hospital level. Variable importance is without units, but is normalized to a scale of 1 to 100, with the larger number indicating increased importance.

*Figure to be used in color in online version and grayscale in print.

Table 1

Baseline Patient and Hospital Characteristics

Characteristic	Frequency(%),Mean (SD)
PATIENT CHARACTERISTICS (n=31,350)	
Age (mean)	40.52 (22.9)
Sex	
Male	20,667 (65.9%)
Female	10,683 (34.1%)
Race	
White	17,085 (54.5%)
Black	4,239 (13.5%)
Hispanic	6,591 (21.1%)
Other	3,435 (10.9%)
Insurance	
Medicare	5,943 (19%)
Medicaid	8,447 (26.9%)
Private Insurance	7,594 (24.2%)
Self-Pay	3,342 (10.7%)
Other	5,736 (10.7%)
Comorbidities	
Diabetes	3,106 (9.9%)
Hypertension	7,719 (24.6%)
Alcohol Dependence	2,245 (7.2%)
Congestive Heart Failure	945 (3.0%)
Obesity	1,329 (4.2%)
Overall Mortality	896 (2.86%)
Length of Stay (median)	5 (2–11)
HOSPITAL CHARACTERISTICS (n=670)	
Total Admissions	36,768.9 (28892)
Inpatient Surgeries	10,464 (8875.2)
Full time residents	416 (484.9)
Full time nurses	1,536 (1355.8)
ICU beds	48.1 (39.2)

Table II

Burn type by ICD9 Code

ICD9 Code	Burn Type
940	Burn confined to eye and adnexa
941	Burn of face head and neck
942	Burn of trunk
943	Burn of upper limb except wrist and hand
944	Burn of wrist(s) and hand(s)
945	Burn of lower limb(s)
946	Burns of multiple specified sites
947	Burn of internal organs
948	Burns classified according to extent of body surface involved
949	Burn unspecified site