Received: 29 June 2021 | Revised: 25 September 2021 | Accepted: 1 October 2021

DOI: 10.1002/osp4.571

ORIGINAL ARTICLE

Obesity Science and Practice Open Access WILEY

Machine learning to predict in-hospital mortality among patients with severe obesity: Proof of concept study

Shelly Soffer[1,2] | Eyal Zimlichman[3,4,5] | Matthew A. Levin[6,7] | Alexis M. Zebrowski[8,9] | Benjamin S. Glicksberg[10,11] | Robert Freeman[6] | David L. Reich[7] | Eyal Klang[4,5,12]

[1]Internal Medicine B, Assuta Medical Center, Ashdod, Israel

[2]Ben-Gurion University of the Negev, Be'er Sheva, Israel

[3]Hospital Management, Sheba Medical Center, Tel Hashomer, Israel

[4]Sackler Medical School, Tel Aviv University, Tel Aviv, Israel

[5]Sheba Talpiot Medical Leadership Program, Tel Hashomer, Israel

[6]Department of Population Health Science and Policy, Institute for Healthcare Delivery Science, Icahn School of Medicine at Mount Sinai, New York, New York, USA

[7]Department of Anesthesiology, Perioperative and Pain Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA

[8]Department of Emergency Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA

[9]Department of Population Health Science and Policy, Institute for Translational Epidemiology, Icahn School of Medicine at Mount Sinai, New York, New York, USA

[10]Hasso Plattner Institute for Digital Health at Mount Sinai, Icahn School of Medicine at Mount Sinai, New York, New York, USA

[11]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York, USA

[12]Department of Diagnostic Imaging, Sheba Medical Center, Tel Hashomer, Israel

Correspondence
Shelly Soffer, Internal Medicine B, Assuta Medical Center, Ha-Refu'a St 7, Ashdod 7747629, Israel.
Email: soffer.shelly@gmail.com

Abstract

Objectives: Hospitalized patients with severe obesity require adapted hospital management. The aim of this study was to evaluate a machine learning model to predict in-hospital mortality among this population.

Methods: Data of unselected consecutive emergency department admissions of hospitalized patients with severe obesity (BMI ≥ 40 kg/m^2) was analyzed. Data was retrieved from five hospitals from the Mount Sinai health system, New York. The study time frame was between January 2011 and December 2019. Data was used to train a gradient-boosting machine learning model to identify in-hospital mortality. The model was trained and evaluated based on the data from four hospitals and externally validated on held-out data from the fifth hospital.

Results: A total of 14,078 hospital admissions of inpatients with severe obesity were included. The in-hospital mortality rate was 297/14,078 (2.1%). In univariate analysis, albumin (area under the curve [AUC] = 0.77), blood urea nitrogen (AUC = 0.76), acuity level (AUC = 0.73), lactate (AUC = 0.72), and chief complaint (AUC = 0.72)

were the best single predictors. For Youden's index, the model had a sensitivity of 0.77 (95% CI: 0.67–0.86) with a false positive rate of 1:9.

Conclusion: A machine learning model trained on clinical measures provides proof of concept performance in predicting mortality in patients with severe obesity. This implies that such models may help to adopt specific decision support tools for this population.

KEYWORDS
big data, in-hospital mortality, machine learning, obesity, severe

1 | INTRODUCTION

A high percentage of patients admitted to the hospital for general medical conditions are diagnosed with obesity.[1] These patients differ in age and comorbidities from hospitalized patients with non-clinically severe obesity.[2] Additionally, patients with obesity have different adaptive biological, physiological, and immunological mechanisms to cope up with illnesses.[3,4]

The association between BMI and in-hospital mortality is conflicting and not straightforward. Several studies have demonstrated that in-hospital mortality is lower among patients with obesity. This has been termed as "the obesity paradox".[5-7] On the contrary, several studies have shown that obesity elevates death risk.[8,9] It has been suggested that mortality is characterized by a U- or J-shaped curve, whereby patients at the extreme ends of weight have worse survival rates.[10,11] In the current COVID-19 pandemic, the mortality risk is higher for hospitalized patients with severe obesity.[12] Thus, hospitalized patients with severe obesity represent a complex population in which mortality prediction is challenging. Predictive models for hospitalized patients with severe obesity are at risk for possible deterioration may be beneficial.

In the last few years, artificial intelligence research has rapidly advanced. The use of machine learning has had a significant impact on the health system.[13-15] The wealth of medical data, stored in electronic health records, has enabled the use of cutting-edge big data technologies. Such technologies can be used to build decision support tools. Models for in-hospital mortality have been applied to several chronic and acute diseases. Examples include chronic heart failure, myocardial infarction, and acute kidney disease.[16-19] In the emergency department (ED) setting, models were also developed for the prediction of different outcomes, such as mortality, ICU hospital admission, and resource utilization.[20,21]

Today, commonly used machine learning algorithms for tabular data classification tasks are tree-based models. The most frequently used tree-based models are random forest (RF) and gradient boosting (GB).[22,23] Tree-based algorithms have the ability to capture the nonlinear relationships that are present in a dataset. In cases where higher order relationships exist in the data, nonlinear methods often outperform linear models.

RF is an ensemble learning method most commonly used for classification problems.[23,24] The model operates by constructing multiple decision trees at training. The output of the RF is selected by a majority vote of all the trees. RF has the ability to correct the common challenge of overfitting which is apparent in decision trees. GB is another machine learning model commonly used for tabular data classification tasks.[22] The model predicts an outcome by constructing an ensemble of multiple weak prediction models. The weak models are usually decision trees.

GB differs from RF in a number of ways. While in RFs multiple trees are created at the same time using bootstrapping of data, in GB, trees are learned sequentially and based on the performance of all previous trees. Each new decision tree created in GB is learned to correct the errors made by the previously created trees. The GB model has surpassed other models in a number of data challenges.[22,25,26]

In the last few years, tree-based models have shown promising results for mortality prediction.[14,17,19,21,27] Patients with severe obesity are a unique hospitalized population that requires adapted clinical management and may benefit from machine learning decision support tools.[5] The aim of this study was to evaluate machine learning algorithms to predict in-hospital mortality among patients with severe obesity. The two tree-based models, RF and GB, were evaluated and compared.

2 | MATERIALS AND METHODS

2.1 | Study design

This retrospective multi-site study was performed in five hospital campuses serving different geographic populations: Mount Sinai Hospital (MSH), Mount Sinai Brooklyn (MSB), Mount Sinai Queens (MSQ), Mount Sinai Morningside (MSM), and Mount Sinai West (MSW). Records were retrieved of all ED admissions of hospitalized patients with severe obesity. Severe obesity was defined as BMI $\geq$ 40 kg/m^2.[28]

Nine years of data (1 January 2011 to 31 December 2019) was extracted from the electronic medical records (EMR). All the hospitals use a unified Epic EMR system (Epic Systems Corporation). Data

included demographics, comorbidities, arrival mode (walk-in, ambulance, or intensive care ambulance), chief complaints, vital signs, laboratory results, and acuity level, also called Emergency Severity Index (ESI), a five-level acuity score assigned by the triage nurse. For each variable, the first measurement that was taken at the ED was used. BMI measurements were acquired from the EMR for all patients.

2.2 | Inclusion and exclusion criteria

Adult patients (≥18 years old) with severe obesity were included. Only ED patients who were hospitalized were included. A documented hospital measurement of BMI ≥ 40 kg/m^2 was required. Exclusion criteria included patients younger than 18 and patients with BMI < 40 kg/m^2.

2.3 | Machine learning model

Data was split into training and internal validation sets (hospitals MSH, MSB, MSM, and MSW) and an external validation set (MSQ). Machine learning models were trained to predict in-hospital mortality. Variables in the models included: demographics, comorbidities, arrival mode, chief complaints, vital signs measurements at admission, ESI, and laboratory results obtained at admission. Table S1 shows the full list of predictors used in the model. The variables that were chosen for the model were based on the data that was available to the physician at the point of entry to the hospitalization ward. These variables included data that was accumulated during the patient's ED stay. Additionally, clinical variables were explored and selected according to the analysis performed in previous publications.[14,29]

2.3.1 | Data preprocessing

Categorical variables were factorized. Continuous variables were not normalized since tree-based methods are not affected by linear transformations (the models' "cut" above and below the desired value). Comorbidities were coded according to the International Classification of Diseases (ICD-9 and ICD-10) records. The ICD codes were grouped using the diagnostic Clinical Classification Software. The first vital signs and laboratory measurements upon arriving to the ED were used as features in the models.

2.3.2 | Machine learning

Two machine learning models were compared: GB and RF. The GB model was implemented using the XGBoost library. The RF algorithm was implemented using the scikit-learn library. For the GB model, null imputation was used. For the RF model, median imputation was used. Model hyper-parameters were tuned in the training/internal validation cohort, using an 80/20 five-cross validation split. An 80/20 split was chosen as it is commonly used in similar studies.[27,30,31]

(GB—number of estimators: 50, eta: 0.3, max depth: 3; RF: number of estimators: 500, criterion: "gini," max depth: "None"). Data balancing techniques using scale weighting did not improve the models' accuracies and thus were not employed. The final GB and RF models were trained on the entire internal validation cohort and tested on the external validation cohort. SHapley Additive exPlanations (SHAP) summary plots were constructed to assess the final GB model feature importance. Programming was done with Python (Version 3.6.5; 64 bits).

2.4 | Statistical analysis

Categorical variables were compared using the χ^2 test. Continuous variables were compared using Student's *t*-test. Statistical significance was established at a two-sided *P*-value of $P < 0.05$. The rate of missing data was analyzed for each variable in the model. The area under the curve (AUC) metric was used to assess the models' performance on the test data. Single variable AUCs were also evaluated. This was conducted by testing each individual variable in the GB model, one variable at a time. Cutoff values of prominent laboratory tests were evaluated for 5%, 10%, and 20% mortality rates. AUCs of two clinical scores were also assessed: The Aged Shock Index (ASI = Age × Pulse/Systolic blood pressure) and Lactate/albumin ratio score.[32,33]

For the GB model, Youden's index was used to find an optimal sensitivity-specificity cut-point on the receiver operating characteristic (ROC) curve. Different metrics were also evaluated for fixed specificities of 90%, 95%, and 99%, respectively. Metrics included sensitivity, specificity, false-positive rate, negative predictive value (NPV), positive predictive value (PPV), and *F*1 score. Bootstrapping validations (100 bootstrap resamples) were used to calculate 95% confidence intervals (CI).

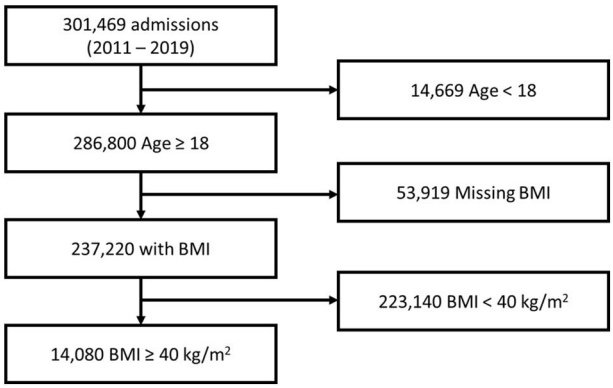

FIGURE 1 Study flow chart

TABLE 1 Baseline characteristics of the study cohort comparing the mortality group to the survival group

	Total (n = 14,078)	Non-survivors (n = 297, 2.1%)	Survivors (n = 13,781, 97.9%)	P-value
Demographics				
Age, median (IQR), years	57.0 (46.0–67.0)	65.0 (58.0–74.0)	57.0 (46.0–66.0)	<0.001
Male, N (%)	4819 (34.2)	114 (38.4)	4705 (34.1)	0.144
White, N (%)	3769 (26.8)	86 (29.0)	3683 (26.7)	0.428
African American, *N* (%)	5040 (35.8)	79 (26.6)	4961 (36.0)	0.001
Arrival status				
BLS, *N* (%)	2278 (16.2)	96 (32.3)	2182 (15.8)	<0.001
EMS, *N* (%)	4717 (33.5)	139 (46.8)	4578 (33.2)	<0.001
ESI, median (IQR) level (1–5)	3.0 (2.0–3.0)	2.0 (2.0–3.0)	3.0 (2.0–3.0)	<0.001
Vital signs				
SBP, median (IQR), mmHg	141.0 (124.0–160.0)	123.0 (105.0–144.2)	141.0 (125.0–161.0)	<0.001
DBP, median (IQR), mmHg	77.0 (67.0–88.0)	69.0 (56.0–81.0)	77.0 (67.0–88.0)	<0.001
Heart rate, median (IQR), bpm	89.0 (77.0–102.0)	92.5 (79.0–108.0)	89.0 (77.0–102.0)	0.042
Temperature, median (IQR), Fahrenheit	98.0 (97.3–98.6)	98.0 (97.0–98.8)	98.0 (97.3–98.6)	0.478
Respirations, median (IQR), num/min	19.0 (18.0–20.0)	20.0 (18.0–22.0)	19.0 (18.0–20.0)	<0.001
O_2 saturation, median (IQR), %	97.0 (96.0–99.0)	97.0 (93.0–99.0)	97.0 (96.0–99.0)	<0.001
Pain scale, median (IQR) (0–10)	4.0 (0.0–8.0)	0.0 (0.0–5.0)	4.0 (0.0–8.0)	<0.001
Comorbidities				
BMI, median (IQR), kg/m^2	44.3 (41.8–48.9)	44.3 (41.7–50.6)	44.3 (41.8–48.9)	0.072
CAD, *N* (%)	2840 (20.2)	68 (22.9)	2772 (20.1)	0.268
CHF, *N* (%)	3581 (25.4)	87 (29.3)	3494 (25.4)	0.140
DM, *N* (%)	6178 (43.9)	132 (44.4)	6046 (43.9)	0.891
HTN, *N* (%)	7180 (51.0)	146 (49.2)	7034 (51.0)	0.559
CKD, *N* (%)	2614 (18.6)	63 (21.2)	2551 (18.5)	0.267
COPD, *N* (%)	2343 (16.6)	59 (19.9)	2284 (16.6)	0.153
Smoking, *N* (%)	6217 (44.2)	120 (40.4)	6097 (44.2)	0.208
Malignancy, *N* (%)	2842 (20.2)	77 (25.9)	2765 (20.1)	0.016
Laboratory				
WBC, median (IQR), $\times$ 10^9 per L	9.3 (7.2–12.1)	10.5 (7.6–16.2)	9.2 (7.2–12.0)	<0.001
LYMPH, median (IQR), $\times$ 10^9 per L	1.7 (1.1–2.6)	1.6 (0.9–4.1)	1.7 (1.1–2.6)	<0.001
NEUT, median (IQR), $\times$ 10^9 per L	6.5 (4.7–9.2)	7.9 (5.4–13.1)	6.5 (4.7–9.1)	<0.001
HGB, median (IQR), g/dL	12.2 (10.6–13.6)	10.9 (9.2–12.5)	12.2 (10.7–13.6)	<0.001
PLT, median (IQR), $\times$ 10^9 per L	244.0 (195.0–303.0)	210.0 (138.0–289.5)	245.0 (196.0–303.0)	<0.001
CR, median (IQR), mg/dL	0.9 (0.7–1.4)	1.5 (0.9–2.5)	0.9 (0.7–1.3)	<0.001
BUN, median (IQR), mg/dL	16.0 (11.0–26.0)	28.0 (19.0–49.8)	16.0 (11.0–25.0)	<0.001
Albumin, median (IQR), g/dL	3.5 (3.1–3.9)	2.8 (2.2–3.3)	3.5 (3.2–3.9)	<0.001
GLC, median (IQR), mg/dL	118.0 (98.0–163.5)	131.0 (104.0–183.0)	118.0 (98.0–163.0)	0.013
Troponin-I, median (IQR), ng/dL	0.0 (0.0–0.0)	0.0 (0.0–0.1)	0.0 (0.0–0.0)	<0.001

(Continues)

TABLE 1 (Continued)

	Total (n = 14,078)	Non-survivors (n = 297, 2.1%)	Survivors (n = 13,781, 97.9%)	P-value
NA, median (IQR), mmol/L	138.0 (135.0–141.0)	136.0 (132.0–139.0)	138.0 (136.0–141.0)	<0.001
CL, median (IQR), mmol/L	103.0 (100.0–105.0)	101.0 (97.0–104.0)	103.0 (100.0–106.0)	<0.001

Abbreviations: BLS, basic life support; BMI,– body mass index; BUN, blood urea nitrogen; CAD, coronary artery disease; CHF, congestive heart failure; CKD, chronic kidney disease; CL, chloride; COPD, chronic obstructive pulmonary disease; CR, Creatinine; DBP, diastolic blood pressure; DM, diabetes mellitus; EMS, emergency medical services; ESI, Emergency Severity Index; GLC, glucose; HGB, hemoglobin; HTN, hypertension; IQR, interquartile range; LYMPH, lymphocytes; NA, natrium; NEUT, neutrophils; PLT, platelets; SBP, systolic blood pressure; WBC, white blood cells.

TABLE 2 The variables with the highest area under the curves (AUCs) to predict in-hospital mortality among patients with severe obesity

Feature	AUC	95% CI
Patient variables		
Chief complaints	0.72	0.65–0.77
Age	0.65	0.59–0.71
African American	0.57	0.52–0.63
White	0.52	0.46–0.57
Sex	0.50	0.44–0.56
Arrival status		
ESI	0.73	0.67–0.79
BLS	0.59	0.52–0.65
EMS	0.57	0.51–0.63
Vital signs		
Respiration rate	0.69	0.62–0.76
SBP	0.67	0.59–0.75
DBP	0.66	0.59–0.73
Pain scale	0.64	0.58–0.69
Pulse oximetry	0.63	0.55–0.70
Temperature	0.60	0.51–0.67
Pulse	0.56	0.49–0.63
Laboratory tests		
Albumin	0.77	0.70–0.83
BUN	0.76	0.71–0.81
Lactate	0.72	0.64–0.80
CR	0.71	0.64–0.77
AST	0.69	0.60–0.77
CL	0.69	0.62–0.76
NA	0.69	0.63–0.75
LYMPH	0.66	0.59–0.74
NEUT	0.65	0.57–0.73
WBC	0.64	0.56–0.73
CA	0.64	0.57–0.70
Troponin	0.63	0.56–0.71
ALK PHOS	0.63	0.55–0.71
HGB	0.62	0.54–0.70
PLT	0.62	0.53–0.70
ALT	0.57	0.48–0.66
CPK	0.57	0.42–0.72
Comorbidities		
BMI	0.55	0.49–0.63
CVD	0.53	0.49–0.59
Neoplastic	0.52	0.49–0.56
HTN	0.52	0.48–0.55
DM	0.52	0.49–0.54
CKD	0.52	0.48–0.56
Smoking	0.51	0.44–0.57
CVA	0.51	0.50–0.51
COPD	0.51	0.49–0.52

Abbreviations: ALK, PHOS alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate transaminase; BLS, basic life support; BMI, body mass index; BUN, blood urea nitrogen; CA, calcium; CKD, chronic kidney disease; CL, chloride; COPD, chronic obstructive pulmonary disease; CPK, creatine phosphokinase; CR, creatinine; CVA, cerebrovascular accident; CVD, cardiovascular disease; DM, diabetes mellitus; EMS, emergency medical services; ESI, Emergency Severity Index; GLC, glucose; HGB, hemoglobin; HTN, hypertension; LYMPH, lymphocytes; NA, natrium; NEUT, neutrophils; PLT, platelets; SBP, systolic blood pressure; WBC, white blood cells.

The Mount Sinai Institutional Review Board (IRB) approved this study. The IRB committee waived informed consent.

3 | RESULTS

A total of 237,197 hospital admissions had documented measurements of BMI. Of those, 14,078 (5.9%) admissions were of patients with severe obesity (BMI $\geq$ 40 kg/m^2). Figure 1 shows the inclusion flow diagram. The in-hospital mortality rate was 297/14,078 (2.1%). The median time to death was 10 days (IQR: 5–22 days). Table 1 presents patient characteristics for the mortality and survival groups. Patients

who died were significantly older. They were more likely to arrive by basic life support or emergency medical services. Significant differences were also observed in vital signs and laboratory results. Table S2 presents how the populations of the five hospitals are different. Of the 14,078 admissions, 10,909 were used for the training/internal validation set, and 3169 for the external validation set.

3.1 | Data exploration

Figure S1A presents the chief complaints associated with the highest mortality rate. Cardiac arrest was the most associated with mortality (63%), followed by hypotension (10.8%), altered mental status (9.9%), and respiratory distress (7.2%). Figure S1B demonstrates the primary diagnosis associated with the highest mortality rate. Cardiac arrest had the greatest association with mortality (8.2%), followed by septic shock (6.5%), shortness of breath (4.4%), and congestive heart failure (3.7%). Single ED variables with the highest AUCs were albumin (AUC = 0.77), blood urea nitrogen (BUN) (AUC = 0.76), ESI (AUC = 0.73), lactate (AUC = 0.72), and chief complaint (AUC = 0.72). The univariate analysis is presented in Table 2. Table S3 shows cutoff values of the prominent laboratory tests for 5%, 10%, and 20% mortality rates.

3.2 | Machine learning model outcome

Table S4 presents the rate of missing data for each variable. The GB model and the RF model showed an AUC of 0.90 (95% CI: 0.87–0.94 and 95% CI: 0.86–0.93, respectively), in the external validation set.

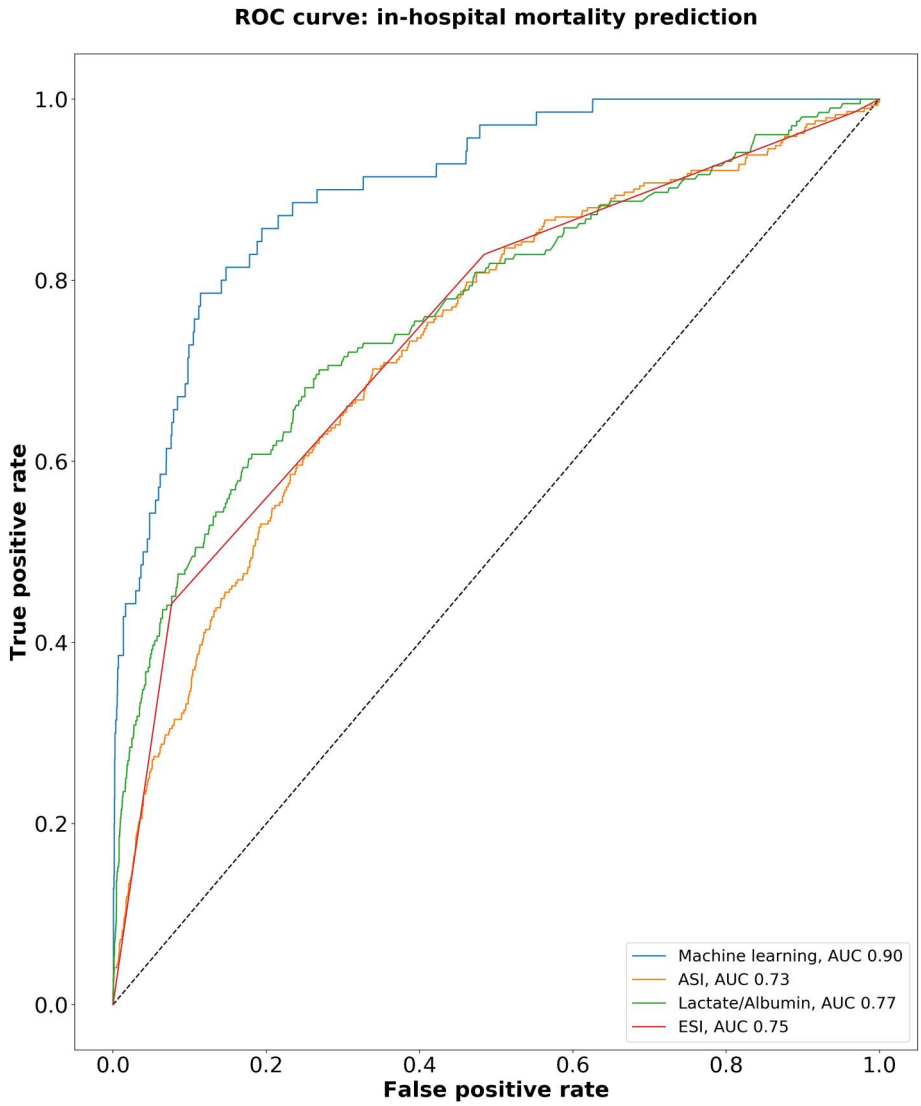

FIGURE 2 Performance characteristic curves for the machine learning model compared to existing clinical scores. ASI, Aged Shock Index; ESI, Emergency Severity Index

TABLE 3 Performance of the final model

Fixed specificity	Sensitivity	Specificity	PPV	NPV	F1[a]
Youden's index	0.77 (95% CI: 0.67–0.86)	0.89 (95% CI: 0.87–0.90)	0.13 (95% CI: 0.10–0.17)	0.99 (95% CI: 0.99–1.00)	0.23 (95% CI: 0.18–0.28)
90%	0.73 (95% CI: 0.62–0.83)	0.90	0.14 (95% CI: 0.10–0.17)	0.99 (95% CI: 0.99–1.00)	0.24 (95% CI: 0.19–0.29)
95%	0.54 (95% CI: 0.42–0.65)	0.95	0.20 (95% CI: 0.14–0.26)	0.99 (95% CI: 0.99–0.99)	0.29 (95% CI: 0.22–0.36)
99%	0.39 (95% CI: 0.28–0.50)	0.99	0.47 (95% CI: 0.33–0.58)	0.99 (95% CI: 0.98–0.99)	0.42 (95% CI: 0.31–0.53)

Abbreviations: NPV, negative predictive value; PPV, positive predictive value.

[a]*F*1 score is defined as the harmonic mean of the sensitivity and positive predictive value.

Figure 2 presents the ROC curve of the GB model compared to clinical scores. In comparison, ASI had an AUC of 0.73 (95% CI: 0.70–0.76), albumin/lactate ratio had an AUC of 0.77 (95% CI: 0.73–0.81), and ESI had an AUC of 0.75 (95% CI: 0.69–0.80). Figure S2 presents the SHAP values plot of the GB model. For Youden's index, the GB model had a sensitivity of 0.77 (95% CI: 0.67–0.86) and a specificity of 0.89 (95% CI: 0.87–0.90), which corresponds to a false positive rate of 1:9 (Table 3). The NPV was 0.99 (95% CI: 0.99–1.0) and PPV was 0.13 (95% CI: 0.10–0.17).

4 | DISCUSSION

This research presents a machine learning tool that can predict in-hospital mortality in patients with severe obesity at the time of admission. Mortality prediction among hospitalized patients with severe obesity is challenging. Predictive models for patients with severe obesity at risk may be beneficial.

The study utilized a large cohort of patients with severe obesity that were hospitalized. An extensive number of data points was collected from the time of ED admission. These included demographics, comorbidities, laboratory results, and ED assessment. Albumin was a strong predictor of mortality. Hypoalbuminemia represents illness severity and was shown to be associated with mortality.[34] Abnormal albumin levels may also be an indicator of nutrition status. In patients with obesity, a high caloric intake does not necessarily correlate with proper nutrition. They may even suffer from malnutrition.[35] Although this test was the best single mortality predictor, it was limited to an AUC of 0.77.

Big data has the potential to help us understand obesity and to address the challenges of this 21st-century epidemic.[36] In recent years, there is an increased usage of machine learning in obesity research.[37-40] Studies that have applied machine learning to obesity have helped predict the disease itself as well as understand the underlying biological and psychological mechanisms of this disease. In this study, machine learning was utilized to predict in-hospital mortality in patients with severe obesity. Since the mortality rate is low (~2%), this may be considered a needle in a haystack problem. For Youden's index, the model achieved a sensitivity of 77% with a false positive rate of 1:9. While not perfect, this proof of concept suggests that a targeted system for patients with severe obesity may alert physicians to about 80% of high-risk patients, while creating one out of nine false positives (high specificity of about 90%). It should be noted that there are scores that predict mortality among the patients with obesity. Examples include Kings Obesity Staging Criteria and Edmonton Obesity Staging System.[41] However, these tools are not used for acute hospitalized patients.

The complexity of predicting mortality among hospitalized patients with severe obesity was evident in this study. Intuitively, it could be assumed that patients with severe obesity who did not survive would have more comorbidities. Yet, this was not evident in this study. Additionally, most single variables did not have a distinguishing capability. Thus, it is clear that the physician's ability to predict mortality of patients with severe obesity is limited. The medical staff must be aware of this population's unique characteristics rather than rely on familiar patterns. Therefore, a decision support system that can alert to high-risk patients with severe obesity can potentially enhance the care delivered to this population.

This study has several limitations. This is a retrospective study. Missing data was recorded in some patients. Nevertheless, the GB model can integrate missing values. Additionally, although this was a multi-center study, it was limited to an urban New York setting. Moreover, in this study, BMI was used to define severe obesity. Alternative indices can be used, such as the waist circumference and the waist-to-hip ratio. Also, different cutoff values of BMI could be used.

In conclusion, machine learning can predict in-hospital mortality among patients with severe obesity. The utilization of a decision support system may assist with the management of this complex population.

ACKNOWLEDGMENTS

The authors received no financial support for the research, authorship, and/or publication of this article.

CONFLICT OF INTEREST

The authors declare that they have no relevant conflict of interest.

AUTHOR CONTRIBUTIONS

Eyal Klang had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data

analysis. Concept and design: Shelly Soffer, Eyal Zimlichman, Matthew A. Levin, Alexis M. Zebrowski, Benjamin S. Glicksberg, Robert Freeman, David L. Reich, and Eyal Klang. Acquisition, analysis, or interpretation of data: Eyal Klang and Shelly Soffer. Drafting of the manuscript: Eyal Klang and Shelly Soffer. Critical revision of the manuscript for important intellectual content: Shelly Soffer, Eyal Zimlichman, Matthew A. Levin, Alexis M. Zebrowski, Benjamin S. Glicksberg, Robert Freeman, David L. Reich, and Eyal Klang. Statistical analysis: Eyal Klang, Alexis M. Zebrowski, and Benjamin S. Glicksberg. Administrative, technical, or material support: Shelly Soffer, Eyal Zimlichman, Matthew A. Levin, Alexis M. Zebrowski, Benjamin S. Glicksberg, Robert Freeman, David L. Reich, and Eyal Klang.

ORCID

Shelly Soffer https://orcid.org/0000-0002-7853-2029
Robert Freeman https://orcid.org/0000-0003-4946-6533

REFERENCES

1. Hossain MA, Amin A, Paul A, et al. Recognizing obesity in adult hospitalized patients: a retrospective cohort study assessing rates of documentation and prevalence of obesity. *J Clin Med*. 2018;7:203.
2. Fusco K, Robertson H, Galindo H, Hakendorf P, Thompson C. Clinical outcomes for the obese hospital inpatient: an observational study. *SAGE Open Med*. 2017;5.2050312117700065
3. Coppack SW. Pro-inflammatory cytokines and adipose tissue. *Proc Nutr Soc*. 2001;60:349-356.
4. Parameswaran K, Todd DC, Soth M. Altered respiratory physiology in obesity. *Can Respir J*. 2006;13:203-210.
5. Schetz M, De Jong A, Deane AM, et al. Obesity in the critically ill: a narrative review. *Intensive Care Med*. 2019;45:757-769.
6. Persaud SR, Lieber AC, Donath E, et al. Obesity paradox in intracerebral hemorrhage: national inpatient sample analysis. *Stroke*. 2019;50:999-1002.
7. Camprubi M, Cabrera S, Sans J, Vidal G, Salvado T, Bardaji A. Body mass index and hospital mortality in patients with acute coronary syndrome receiving care in a university hospital. *J Obes*. 2012;2012:287939.
8. Ghanta RK, LaPar DJ, Zhank Q, et al. Obesity increases risk-adjusted morbidity, mortality, and cost following cardiac surgery. *J Am Heart Assoc*. 2017;6:e003831.
9. Akinyemiju T, Sakhuja S, Vin-Raviv N. In-hospital mortality and post-surgical complications among cancer patients with metabolic syndrome. *Obes Surg*. 2018;28:683-692.
10. Schiffl H. Obesity and the survival of critically ill patients with acute kidney injury: a paradox within the paradox? *Kidney Dis*. 2020;6:13-21.
11. Antonopoulos A, Oikonomou E, Antoniades C, Tousoulis D. From the BMI paradox to the obesity paradox: the obesity–mortality association in coronary heart disease. *Obes Rev*. 2016;17:989-1000.
12. Klang E, Kassim G, Soffer S, Freeman R, Levin MA, Reich DL. Severe obesity as an independent risk factor for COVID-19 mortality in hospitalized patients younger than 50. *Obesity*. 2020;28:1595-1599.
13. Barash Y, Guralnik G, Tau N, et al. Comparison of deep learning models for natural language processing-based classification of non-English head CT reports. *Neuroradiology*. 2020;62:1247-1256.
14. Klug M, Barash Y, Bechler S, et al. A gradient boosting machine learning model for predicting early mortality in the emergency department triage: devising a nine-point triage score. *J Gen Intern Med*. 2020;35:220-227.
15. Segar MW, Vaduganathan M, Patel KV, et al. Machine learning to predict the risk of incident heart failure hospitalization among patients with diabetes: the WATCH-DM risk score. *Diabetes Care*. 2019;42:2298-2306.
16. Kwon JM, Kim KH, Jeon KH, Park J. Deep learning for predicting in-hospital mortality among heart disease patients based on echocardiography. *Echocardiography*. 2019;36:213-218.
17. Kwon J-M, Kim K-H, Jeon K-H, et al. Artificial intelligence algorithm for predicting mortality of patients with acute heart failure. *PloS One*. 2019;14:e0219302.
18. Kwon J-M, Jeon K-H, Kim HM, et al. Deep-learning-based risk stratification for mortality of patients with acute myocardial infarction. *PloS One*. 2019;14:e0224502.
19. Lin K, Hu Y, Kong G. Predicting in-hospital mortality of patients with acute kidney injury in the ICU using random forest model. *Int J Med Inf*. 2019;125:55-61.
20. Klang E, Kummer BR, Dangayach NS, et al. Predicting adult neuroscience intensive care unit admission from emergency department triage using a retrospective, tabular-free text machine learning approach. *Sci Rep*. 2021;11:1381.
21. Subudhi S, Verma A, Patel AB, et al. Comparing machine learning algorithms for predicting ICU admission and mortality in COVID-19. *NPJ Digit Med*. 2021;4:1-7.
22. Chen T, Guestrin C. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, 13–17 August. Association for Computing Machinery.
23. Liaw A, Wiener M. Classification and regression by Random Forest. *R News*. 2002;2:18-22.
24. Oshiro TM, Perez PS, Baranauskas JA. *International Workshop on Machine Learning and Data Mining in Pattern Recognition*. New York: Springer; 2012.
25. Ogutu JO, Piepho H-P, Schulz-Streeck T. Pre-selection of markers for genomic selection. *BMC Proc*. 2011;5:S12.
26. Ugon A, Karlsson D, Klein G, Moen A. *Building Continents of Knowledge in Oceans of Data: The Future of Co-created eHealth*. IOS Press; 2018.
27. Taylor RA, Pare JR, Venkatesh AK, et al. Prediction of in-hospital mortality in emergency department patients with sepsis: a local big data–driven, machine learning approach. *Acad Emerg Med*. 2016;23:269-278.
28. Weir CB, Jan A. BMI classification percentile and cut off points [Internet]. StatPearls; 2021.
29. Soffer S, Klang E, Barash Y, Grossman E, Zimlichman E. Predicting in-hospital mortality at admission to the medical ward: a big-data machine learning model. *Am J Med*. 2021;134:227-234.
30. Zhang G, Xu J, Yu M, Yuan J, Chen F. A machine learning approach for mortality prediction only using non-invasive parameters. *Med Biol Eng Comput*. 2020:1-44.
31. Thorsen-Meyer H-C, Nielsen AB, Nielsen AP, et al. Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records. *Lancet Dig Health*. 2020;2:e179-e191.
32. Wang B, Chen G, Cao Y, Xue J, Li J, Wu Y. Correlation of lactate/albumin ratio level to organ failure and mortality in severe sepsis and septic shock. *J Crit Care*. 2015;30:271-275.
33. Torabi M, Mirafzal A, Rastegari A, Sadeghkhani N. Association of triage time Shock Index, Modified Shock Index, and age Shock Index with mortality in Emergency Severity Index level 2 patients. *Am J Emerg Med*. 2016;34:63-68. doi:10.1016/j.ajem.2015.09.014
34. Jellinge ME, Henriksen DP, Hallas P, Brabrand M. Hypoalbuminemia is a strong predictor of 30-day all-cause mortality in acutely admitted medical patients: a prospective, observational, cohort study. *PloS One*. 2014;9:e105983.

35. Fieber JH, Sharoky CE, Wirtalla C, Williams NN, Dempsey DT, Kelz RR. The malnourished patient with obesity: a unique paradox in bariatric surgery. *J Surg Res*. 2018;232:456-463.
36. Lee BY, Ferguson MC, Cox SN, Phan PH. Big data and systems methods: the next frontier to tackling the global obesity epidemic. *Obesity*. 2021;29:263-264.
37. Timmins KA, Green MA, Radley D, Morris MA, Pearce J. How has big data contributed to obesity research? A review of the literature. *Int J Obes*. 2018;42:1951-1962.
38. Morris MA, Wilkins E, Timmins KA, Bryant M, Birkin M, Griffiths C. Can big data solve a big problem? Reporting the obesity data landscape in line with the Foresight obesity system map. *Int J Obes*. 2018;42:1963-1976.
39. Hamad R, Pomeranz JL, Siddiqi A, Basu S. Large-scale automated analysis of news media: a novel computational method for obesity policy research. *Obesity*. 2015;23:296-300.
40. Chiang J, Arons A, Pomeranz JL, Siddiqi A, Hamad R. Geographic and longitudinal trends in media framing of obesity in the United States. *Obesity*. 2020;28:1351-1357.
41. Abdelaal M, le Roux CW, Docherty NG. Morbidity and mortality associated with obesity. *Ann Transl Med*. 2017;5:161.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Soffer S, Zimlichman E, Levin MA, et al. Machine learning to predict in-hospital mortality among patients with severe obesity: proof of concept study. *Obes Sci Pract*. 2022;8(4):474-482. https://doi.org/10.1002/osp4.571