Predicting Mortality in Diabetic ICU Patients Using Machine Learning and Severity Indices

Rajsavi S. Anand[1,2], Paul Stey, PhD[1,2], Sukrit Jain[1,2], Dustin R. Biron[1,2], Harikrishna Bhatt, MD[1], Kristina Monteiro, PhD[1], Edward Feller, MD[1], Megan L. Ranney, MD, MPH[1,3], Indra Neil Sarkar, PhD, MLIS[1,2], Elizabeth S. Chen, PhD[1,2]
Alpert Medical School[1], Center for Biomedical Informatics[2], and Emergency Digital Health Innovation Program, Department of Emergency Medicine[3], Brown University, Providence, RI, USA

Abstract

Diabetes constitutes a significant health problem that leads to many long term health issues including renal, cardiovascular, and neuropathic complications. Many of these problems can result in increased health care costs, as well risk of ICU stay and mortality. To date, no published study has used predictive modeling to examine the relative influence of diabetes, diabetic health maintenance, and comorbidities on outcomes in ICU patients. Using the MIMIC-III database, machine learning and binomial logistic regression modeling were applied to predict risk of mortality. The final models achieved good fit with AUC values of 0.787 and 0.785 respectively. Additionally, this study demonstrated that robust classification can be done as a combination of five variables (HbA1c, mean glucose during stay, diagnoses upon admission, age, and type of admission) to predict risk as compared with other machine learning models that require nearly 35 variables for similar risk assessment and prediction.

Introduction

Diabetic patients constitute 7% of the United States population, with at least 22.3 million diabetics in the United States today.[1] These patients use significantly more healthcare resources than patients with other chronic diseases, accounting for more than 45% of intensive care unit (ICU) patient stays above the age of 65.[2] Diabetes itself leads to a higher incidence of nearly all comorbidities including renal, cardiovascular, and neuropathic disease.[3] Furthermore, diabetic complications can directly impact how patients will fair in the ICU, with a strong association between diabetes and ICU bloodstream infections established.[4] However, the effect of stand-alone diabetes on a patient's risk of ICU mortality has been debated.[5] A recent meta-analysis of 141 studies by Siegelaar *et al.* suggests that having a diagnosis of diabetes in the ICU does not itself directly lead to increased mortality in most ICU settings, only specifically in the cardiac surgical unit.

Management of diabetes in the ICU itself is a controversial topic as tight glucose control in hyperglycemic patients has been seen to have both positive and negative results on mortality outcomes depending on the trial setting.[6] The current clinical practice is to have moderate control of blood glucose in the ICU (140 mg/dl), rather than tight control (<110 mg/dl). To date, no study has looked at the combination of glucose control, hemoglobin A1c (HbA1c) values, and comorbidities to predict mortality outcomes. A recent study by Mahmoodpoor *et al.* demonstrated that higher HbA1c values and admission glucose values were predictive of mortality for ICU patients, showing the value of these variables in predictive modeling for diabetics.[7] By better understanding how to treat diabetic patients in the ICU, mortality outcomes could potentially be improved, thus leading to better patient health with less economic burden.

Algorithms exist to predict mortality risk in general patient populations, with the Charlson Comorbidity Index (CCI) and the Elixhauser Comorbidity Index well agreed upon in the literature.[8] These two measures calculate risk scores based on the ICD-9-CM diagnosis codes for each patient. The CCI scores are based on the severity of symptoms, weighing more serious conditions with more points; its scoring of general hazard risks have been validated for survival at one and two years.[9] The Elixhauser score is based solely on category of comorbidity, and gives one point for each. Furthermore, both of these indices have been applied to ICU patients. However, these indexes have not been validated for diabetics in the ICU. The Diabetes Complications Severity Index (DCSI) has recently been established as an alternative to these measures with a specific focus on diabetic patients using ICD-9-CM diagnosis codes and renal labs.[10] The DCSI has yet to be validated in an ICU population of diabetic patients, but it has been shown to be effective in predicting hospitalizations in general diabetic populations. Additionally, this metric has proved successful in predicting which diabetics are most likely to utilize increased healthcare expenditure.[11]

Machine learning algorithms have been applied to ICU settings, but never specifically to diabetic patient populations in the ICU. Studies have shown the ability to predict risk of mortality in the ICU based on many variables including vitals, labs, surgical history, diagnoses, ventilator settings, and past medical history.[12] Additionally, diabetic specific machine learning has been used in a variety of contexts: detecting hypoglycemia, predicting diagnosis, and further complications, as well as blood glucose classification.[13] Recently, there has been success in predicting which diabetic patients are most at risk for specific complications (retinopathy, neuropathy, and nephropathy) with area under the curve statistics of 0.838.[14] While useful in predicting complications, neither study investigated diabetic patients in the ICU.

One of the most widely used mortality prediction algorithms in the ICU, the Acute Physiology and Chronic Health Evaluation (APACHE) II algorithm, requires 15 variables (of which 12 are physiological measurements) and does not take into account specific patient populations or diagnoses.[15] Subsequent iterations of APACHE (III and IV) have become more and more time consuming, with the need to spend 37.3 minutes on average to enter in relevant information.[16] While more accurate at predicting mortality, these algorithms may prove to be unrealistic in their implementation due to the sheer amount of information and time required to perform them. This had led in part to the continued use of the APACHE-II algorithm first described in 1985. Additionally, other scoring algorithms exist for predicting mortality in the ICU such as the Simplified Acute Physiology Score (SAPS) III, which uses 17 variables, including physiological and disease manifestation variables.[17] The Sequential Organ Failure Assessment (SOFA) score is also well researched and gives a risk calculation for organ failure.[18] Algorithms that require the least possible information upon admission have high practical clinical utility as they allows for rapid implementation and changes in patient treatment strategies. Thus combining diabetic specific metrics and using the fewest possible variables may be of high clinical use in this cohort of patients.

This study aimed to test which of the established comorbidity indices best predicts mortality in diabetic patients in the ICU. By combining other variables of diabetic health including HbA1c, mean glucose during stay, insulin status, as well as these various indices, predictions are made about which diabetics are more likely to have adverse outcomes using both logistic modeling techniques as well as machine learning algorithms.

Methods

Data were extracted from a local MySQL database containing Medical Information Mart for Intensive Care III (MIMIC-III) data using SQL queries. Pre-processing (e.g., data recoding) and analyses were done using the Julia programing language (v.0.5). Data were then transformed into various indices and calculations were conducted to create the other variables of interest. The Charlson Comorbidity Index, Elixhauser Comorbidity Index, and the Diabetic Severity Index were calculated as described in the literature, with points assigned for the presence of specific ICD-9-CM codes. Finally, predictive models of mortality were generated on training sets using available Julia packages for statistics and machine learning. Validation occurred on the remaining 30% of available data with results presented from this set of data.

Data and Variable Selection

The MIMIC-III database, which includes data from Beth Israel Deaconess Medical Center's ICU and Hospital from 2001 to 2012, was used for this study.[19] This database has comprehensive information regarding ICU admissions and all the data needed to address this question. Data were selected from MIMIC-III detailing admissions, medications, ICU stays, ICD-9-CM diagnoses, and lab values. First, all patients with an ICD-9-CM diagnosis of 250.0 – 259.0 (diabetes inclusive) were selected. A dataset was created that contained the diagnoses, admissions, lab values, and ICU medications for this sample upon which analysis was conducted. Labs of interest were extracted using the following LOINC codes: '4548-4' for HbA1c, '2345-7' for Blood Glucose, and '2160-0' for Serum Creatinine. Generic medication names as well as text encoding ethnicity, insurance, admission type, gender, and admission location were also extracted from the medical record. Patients who did not have HbA1cs or Blood Glucose values listed were excluded.

Data Processing

All data processing was conducted using the Julia programming language. The *DataFrames.jl* package was used to aggregate data in tabular formats. The average blood glucose per stay and comorbidity index scores were calculated for each patient. Demographic data were merged with the lab values and index scores into a single data frame for analysis. Descriptive statistics for the sample were calculated.

Admission type was recoded from the original options (elective, emergent, newborn, or urgent) to give numerical outputs. "Elective" was defined as a previously planned hospital admission, while "emergent" or "urgent" were defined as unplanned medical care. Ethnicity was recoded from 41 categories into numerical outputs. Insurance was also recoded from the original options (government, Medicaid, Medicare, private, or self pay) to give numerical outputs.

Insulin status was determined by cross-referencing patient drug lists with generics that included "insulin" in the name. Blood glucose was averaged for each patient's stay to create a mean glucose variable. HbA1c was averaged for each patient's history to gauge overall diabetic health. The former predicts how the patient's diabetes was handled during the stay while the later was a measure of overall diabetic health over their history.

A DCSI score (0-13) was calculated for each patient as described in the literature using diagnosis codes and serum creatinine levels.[20] Diagnosis codes were categorized into seven different comorbidity categories: (1) ophthalmic, (2) nephropathy, (3) neuropathy, (4) cerebrovascular, (5) cardiovascular, (6) peripheral vascular disease, and (7) metabolic. Depending on the presence or absence of a diagnosis code in a patient's record, a value was assigned: 0 for no comorbidity, 1 for minor, and 2 for major. Scores were assigned using regular expressions to search for the presence of the diagnosis codes in a patient's list of diagnosis codes. Patients with serum creatinine values > 1.5 mg/dL were assigned one point, while those with serum creatinine values > 2.0 mg/dL were assigned two points. The nephropathy category was the maximum point assignment of either the diagnosis code or the serum creatinine value. All of these categories were summed and calculated as the maximum for each patient per visit.

An Elixhauser score (0-36) was calculated for each patient using diagnosis codes and assigning a single point for each comorbidity within the 36 categories established in the literature.[8] The CCI score (0-34) was calculated by giving values to comorbidities in one of 17 categories.[9] Patients who had more serious conditions were given more points for these conditions as per CCI guidelines in the literature: 2 (hemiplegia, moderate or severe renal disease, diabetes with end stage-organ damage, tumor without metastasis, leukemia, lymphoma); 3 (moderate or severe liver disease); 6 (metastatic solid tumor, AIDS). Age brackets (<50, 50-60, 60-70, 70-80, >80) were determined by calculating the difference between ICU stay and date of birth and assigned point values of 0, 1, 2, 3, 4, and 5 respectively. The diagnosis score and age bracket score values were then summed and calculated as described above for the DCSI.

Prediction Modeling

The outcome of interest was all cause mortality in the index visit. All model fitting was conducted using packages from the Julia programming language. The *GLM.jl* package was used to fit binomial logistic regression models. Model creation utilized data from 70% of the sample, while the remaining 30% of the sample was used for model validation. Using binomial logistic regression, all of the variables were used in bivariate analyses to correlate with the outcome of mortality. The significant variables with a *p*-value less than 0.05 were combined for multivariate analysis.

Receiver operating characteristic (ROC) curves were generated by varying the thresholds of each model and the models were compared based on their area under the curve (AUC) values. Each index was first run by itself and compared to one another. Then all of the metrics that had proved significant on preliminary logistic regression analysis were compiled into a final binomial logistic regression model (three indices, demographic variables, and diabetic health metrics).

The *DecisionTrees.jl* package was used to fit random forest models. All of the variables extracted from the MIMIC-III database were used in this analysis. The following meta-parameters were used: three features randomly selected at each node, 5000 total trees, maximum of five observations in leaf nodes, and a maximum depth of ten. All meta-parameters were validated and selected using five-fold cross validation.

The *SciKitLearn.jl* package was used for fitting L1 penalized regression (i.e., least absolute shrinkage and selection operator [Lasso]) models on a combined model that contained all of the significant variables from the logistic regression models to determine which variables had the greatest impact on mortality. The following meta-parameters were used: 6 models with varying alpha values from 10^{-4} to 10^{2}.

Results

Of the 10,318 patients in the MIMIC-III database with diabetes as a diagnosis, 4111 diabetic patients had values for HbA1c and Blood Glucose; all patients in this subset also had values for Serum Creatinine. Of those patients, 3729 (90.7%) survived, while 382 (9.3%) died during their ICU stay. Summary statistics for this patient population are

shown in Table 1**.** The six health related variables that were used in the binomial logistic regression are shown at the bottom of Table 2. Differences between the patients with the outcome of death and those who were alive are shown. Of note, there was an increase in all of the health metrics in the subset of patients with a death outcome, other than that of mean A1c value and insulin status. An increase in these categories mirrors an increase in a negative outcome. Patients who died in the ICU on average had 1.71 HbA1c values recorded as opposed to those who survived who had 2.39 HbA1c values. Patients who died in the ICU furthermore had on average 16.23 blood glucose measurements as opposed to those who survived who had 11.54 blood glucose values recorded.

Table 1: Summary statistics on the cohort of patients who died and those who survived

Category		Alive		Dead	
Age	<50	490	13.1%	20	5.2%
	50-60	725	19.4%	48	12.6%
	60-70	1023	27.4%	103	27.0%
	70-80	926	24.8%	112	29.3%
	>80	565	13.7%	99	25.9%
Gender	Female	1527	40.9%	171	44.8%
	Male	2202	53.6%	211	55.2%
Ethnicity	Asian	94	2.5%	7	1.8%
	Black	511	12.4%	69	18.1%
	Hispanic/Latino	183	4.9%	10	2.6%
	Other	447	12.0%	35	9.2%
	White	2494	60.7%	261	68.3%
Insurance	Government	112	3.0%	1	0.3%
	Medicaid	291	7.8%	23	6.0%
	Medicare	2238	60.0%	294	77.0%
	Private	1061	28.5%	60	15.7%
	Self Pay	27	0.7%	4	1.0%
Admission Type	Elective	649	17.4%	15	3.9%
	Emergency	3015	73.3%	362	94.8%
	Urgent	65	1.6%	5	1.3%

Table 2: Average values for variables of interest in cohort of patients who died and those who survived

Variable	Alive Mean	Death Mean
HbA1c	7.472 ± 1.86	6.965 ± 1.45
CCI Score	5.454 ± 2.31	7.113 ± 2.49
DCSI Score	2.974 ± 1.81	4.166 ± 1.72
Elixhauser Score	4.449 ± 1.86	5.835 ± 2.00
Insulin Status	0.968 ± 0.13	0.966 ± 0.18
Mean Glucose	156.4 ± 48.38	169.07 ± 57.7

Figure 1 shows the ROC curve for the models explored. The AUC for the DCSI only model was 0.694, while the AUC for the Elixhauser model and CCI model alone were 0.682 and 0.656 respectively. When all of the metrics were combined into one binomial logistic regression, the AUC was 0.785 suggesting that combining the three

metrics was the most accurate in predicting the outcome of mortality. Additionally, the AUC for the random forest model was 0.787.

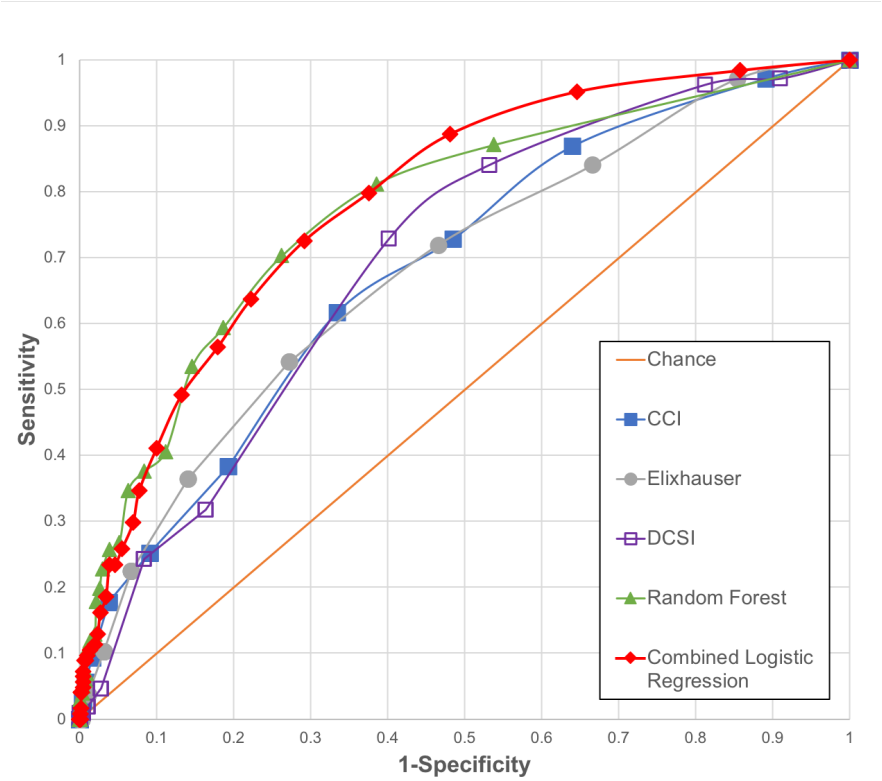

Figure 1: Receiver operating curve of various modeling techniques

The best performing logistic regression model was the Combined Model that showed the best AUC. Table 3 shows the significance and coefficients for each of the major metrics that contributed to the Combined Model. Each of the metrics showed significance, with the largest coefficient related to Elixhauser Sum and admission type. Ethnicity, insurance, age bracket, gender, and insulin status were not shown to be significant predictors of mortality as they had *p*-values greater than 0.05.

Table 3: Binomial logistic regression modeling results with coefficients, standard errors, z values, and statistical significance

Variable	Coefficient	Std. Error	Z-value	P-value
DCSI Sum	0.173	0.034	5.011	< 0.001
Elixhauser Sum	0.203	0.033	6.146	< 0.001
CCI Sum	0.152	0.027	5.612	< 0.001
Mean Glucose	0.010	0.001	8.337	< 0.001
Mean A1c	-0.285	0.046	-6.21	< 0.001
Admission Type	0.455	0.129	3.533	< 0.001

Each model showed high sensitivity and specificity at various thresholds, with the Combined logistic regression model and random forest models performing the best. Table 4 shows the sensitivity and specificities after varying the threshold for classifying the linear models and the decision tree algorithm. The combined model showed the highest sensitivity and specificity. The bolded values show the optimal thresholds for sensitivity/specificity combinations. At optimal thresholds of 0.10 for DCSI and Combined; 0.08 for Elixhauser and CCI; and 0.04 for Random Forest, there were 76-78 classified correctly as positive (true positives), 580-777 classified correctly as negative (true negatives), 349-545 classified incorrectly as positive (false positives), and 29-31 classified incorrectly as negative (false negatives). For example, at a threshold of 0.10, the DCSI model classified 78/107 cases correctly as positive, and 674/1126 of the cases correctly as negative. This is in comparison to the CCI model at a threshold of 0.08 that classified 78 correctly as positive and 580 correctly as negative.

Table 4: Sensitivities and specificities of various models

Model	DCSI		ELIX		CCI		**Combined**		**Random Forest**	
Threshold	sens.	spec.	sens.	spec.	sens.	spec.	**sens.**	**spec.**	sens.	spec.
0.04	97%	9%	97%	15%	97%	11%	95%	35%	**81%**	**61%**
0.06	96%	19%	84%	33%	87%	36%	89%	52%	**70%**	**74%**
0.08	**84%**	**47%**	**72%**	**53%**	**73%**	**52%**	**80%**	**62%**	59%	81%
0.1	**73%**	**60%**	**54%**	**73%**	**62%**	**67%**	**73%**	**71%**	53%	85%
0.12	32%	84%	54%	73%	38%	81%	64%	78%	41%	89%
0.14	32%	84%	36%	86%	38%	81%	56%	82%	38%	92%
0.16	24%	92%	36%	86%	25%	91%	49%	87%	35%	94%

A series of Lasso models were used to estimate which variables had the largest impact on the model. Significant variables from the logistic regression models were applied to determine which metrics carried the greatest significance. The results appear in Figure 2. Mean glucose, Elixhauser score, and CCI score were the last three metrics whose coefficients shrank to 0 as alpha increased.

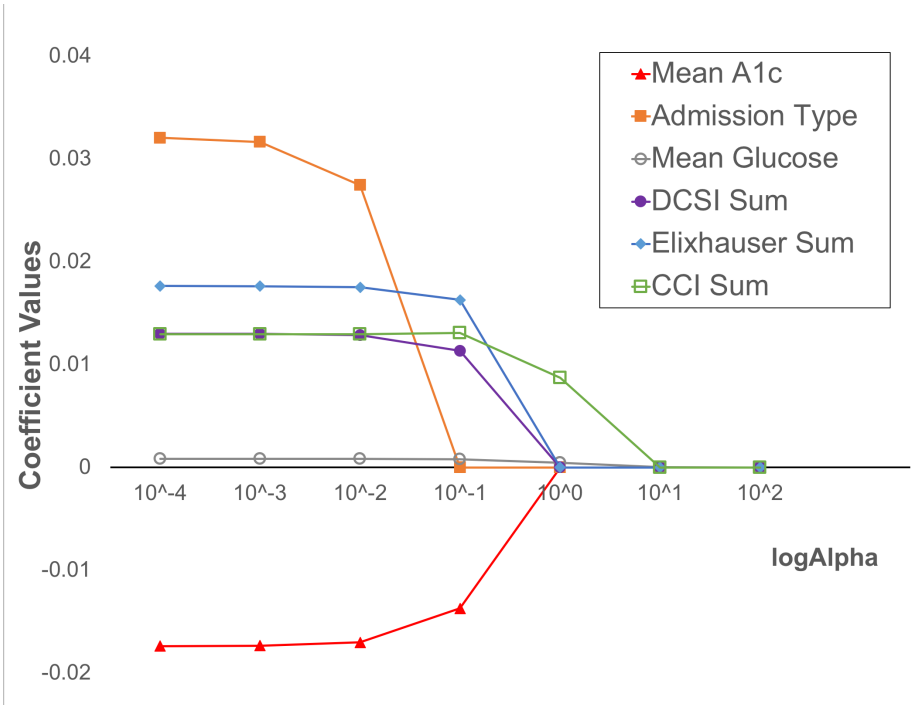

Figure 2: Lasso logistic regression models with increase penalty terms (alpha)

Discussion

Mortality in the ICU is a multifactorial issue that depends upon severity of the admitting disease, quality of care, and infections, amongst other issues. Diabetic patients in the ICU present a unique problem of management of a chronic condition while simultaneously managing acute exacerbation of a comorbidity. It is important to predict mortality of diabetic patients on ICU admission because this will allow for better management of care in the ICU. The summary data shown in Table 2 suggest that no single existing comorbidity score or single variable is able to accurately predict mortality, as the means of the individual predictors demonstrate overlap between the subsets of patients who survived and those who died. Thus, the use of multivariate linear modeling and machine learning algorithms are useful to better predict the complex interplay between the variables in the prediction of mortality.

As a first study, mortality was considered a binomial outcome variable. Future work might include additional outcomes (e.g., bins consisting of range of time to death values). Additionally, this study focused on patients who had HbA1cs and Blood Glucose values. The lack of HbA1c data for certain diabetic patients does not seem to be

based on clinical differences between the two populations, but rather as a function of missing data from a patient's prior history at the hospital. Future work could investigate the differences between diabetic patients in the MIMIC-III database who had HbA1cs and those who did not. Additional characterization of these two groups will give further insight into the problem at hand.

Based on the results of the logistic regression analysis, the variables that significantly correlated with mortality were mean glucose of each ICU visit, mean HbA1c for each patient, index admission type for each patient, and the various indices' scores (Table 3). Patient ethnicity, admission location, insurance, and insulin status were not significantly associated with mortality. Lack of correlation with insulin status is likely due to lack of stratification of this variable (with 96.6% of the population using insulin). Insulin use in type 2 diabetics begins in later stages and with more advanced disease, suggesting that this population has had a long diabetic history. A recent study by Vincent, *et al.*, described this phenomenon by showing that insulin treated diabetes did not increase the risk of mortality in the ICU.[21] However this study did not stratify diabetic patients with other variables as this present study has. Perhaps surprisingly, neither insurance nor patient ethnicity were an indicator of future mortality. Previous studies have shown that insurance type has a significant effect on diabetic treatment with 36% of privately insured patients getting all recommended preventative treatment vs. 16% for publicly insured.[22] Despite this result, however, insurance did not prove to affect the model.

In this particular data set, the mean blood glucose was the variable most strongly associated with diabetic mortality in the ICU based on the series of Lasso models (Figure 2). Based on these results, every increase in mean blood glucose by 1 mg/dl led to an 0.010 increase risk in death suggesting that 25 mg/dl changes in blood glucose can increase log odds risk by 25%. This result is of significant interest because the question of how to manage diabetic blood sugar in the ICU is highly debated.

The first randomized control trial of blood glucose in the ICU, the Leuven trial showed this trend, suggesting that tight control to 80-110 mg/dl was more successful in treating patients than the control group which had glucose values of 180–215 mg/dL.[6] Compared to the control group, absolute mortality was decreased by 3%. In contrast, the NICE-SUGAR trial showed that tight control of diabetic blood sugars led to an absolute mortality increase of 3%. Instead, they argued that physicians should target 140 mg/dl for their patients as they had targeted in their control group. The latter article has been taken to be the standard of care leading many ICUs to practice maintaining higher levels of glucose despite the contradicting evidence shown in the earlier studies. Both research studies that were performed used different methods of blood glucose measurement (blood gas analyzers vs. hand glucose monitors). The result from this current study suggests that this issue may need to be revisited as the greatest correlative factor with mortality was that of mean hospital blood glucose level.

The negative effect of HbA1c on the risk of death is problematic suggesting that as a variable there is not a normative distribution. HbA1c is clinically used as a measure of diabetic health over the past three months indicating how well a patient's blood sugar has been maintained. While mean glucose during the patient's stay is most likely to affect the immediate situation, a higher HbA1c value is correlated with higher risk of complications by itself.[3] Perhaps the number of HbA1c tests in the patients who died (1.78) versus those who survived (2.24) may have had an impact on this variable's effect on the model. A future analysis that warrants investigation is whether or not the most recent A1c has as strong of an impact as well as the number of a patient's HbA1cs or Blood Glucoses. Additionally, further investigation needs to be conducted into whether these patients were being managed on stronger medications. It is possible that stronger medication regimens are more strongly correlated to a low HbA1c value than the actual diabetic health of the patient. This type of analysis is not possible with the MIMIC-III database as it only contains data from previously documented visits within the same health system and many patients do not have records other than the current ICU stay. Future studies validating this algorithm should look into the direct effect of HbA1c on ICU mortality.

Amongst the various indices, all three had similar AUC values with the DCSI having a slight advantage. In comparing the three linear models (Table 4 and Figure 1), the AUC for the DCSI model was 0.694 as compared to 0.682 and 0.656 for the Elixhauser model and CCI model respectively. Despite the slightly better performance of the DCSI, none of these three models performed at a level of clinical efficacy. The fact that the DCSI, however, did perform better in detecting diabetic mortality than the other metrics, suggests that its use may be expanded past its previous uses: outpatient, inpatient non-ICU hospitalization, and healthcare cost analysis. Of note, previous studies that validated the DCSI did not compare it to these more accepted comorbidity indices of the CCI and Elixhauser.[10,11,23,24] There is significant overlap between the three indices, with variance in (a) the categorization of the various comorbidities, (b) the inclusion of age as a metric, (c) the inclusion of serum creatinine, and (d) the

granularity of the metric. When the three indices were combined, the AUC of the model became the highest at 0.785 suggesting that this model was the best at predicting mortality in these patients. At a threshold of 0.08, the sensitivity of the model was 80%, and the specificity was 62%, while at 0.10, 73% and 71% respectively. These values suggest that this model is a good predictor of outcomes for diabetic ICU patients.

These AUC values are higher than that reported in the literature of CCI predictability alone in 30-day mortality using retrospective data (0.755).[25] When the model from Stavem *et al.* was performed upon the dataset, we achieved an AUC of 0.704. The binomial regression and random forest models outperformed this model with an AUC of 0.785 and 0.787 respectively. Additionally, unlike the study performed by Stavem *et al.*, gender was not seen to be predictive in the logistic regression model, as it had a probability > 0.05. One reason for this discrepancy may be due to their choice of outcomes of 30 day and one year mortality, whereas this current study predicted outcomes based on the length of the index ICU visit, which varied. Another potential confounder is that their study was conducted in a nine patient ICU unit at a hospital in Norway and may not be generalizable to data collected in a large tertiary care center in the U.S.

In comparison, the random forest modeling was comparable to the other models in predicting mortality suggesting that it has use in predicting which of these patients will have the greatest risk of mortality. The AUC value of this model was the highest of any of the models at 0.787. At a threshold of 0.06, the model had a sensitivity of 70% and specificity of 73%. This combination suggests that it is a useful method for predicting mortality. Further optimization of the model may allow for an even better prediction value. However, the random forest model still performed slightly better than the logistic regression suggesting it may be better used in this case.

When mean glucose is removed from the combined logistic regression model, the AUC decreases slightly to 0.757, but still outperforms all of the individual indexes alone. This suggests that based on four variables (diagnoses at admission, type of admission, patient's HbA1c history, and age), one can predict with a sensitivity of 76% and specificity 60% which patients will be likely to die during their stay. Prospective studies are needed to assess whether this relationship is merely correlation or causation. This is in stark contrast to previous studies leveraging machine learning algorithms to predict mortality in ICU patients, including the Super ICU Learner Algorithm as well as the Artificial Neural Networks (ANN).[25,26] While both of these studies showed improved AUCs than the present study (0.85), the number of variables needed for prediction rise from five to between 15-35 including various labs, vital signs, and surgical history. Complex models with multiple variables take longer to calculate and require more comprehensive labs and vitals upon admission.[16] Additionally some of the metrics may be heavier on computing power. The model presented in this study, however, requires only a few aspects of a patient's prior medical history. A case can be made for a simpler, faster to calculate modeling approach for a number of clinical settings. For example, the approach developed here could be used as a pragmatic triage mechanism excluding mean glucose.

A potential limitation of this study is the use of a random 70/30 class sampling for analysis as there was a class imbalance in which less than 10% of the total were positive cases. Future work may consider the use of class balancing techniques such as Synthetic Minority Over-Sampling Technique (SMOTE) to account for this imbalance. It was felt that the imbalance would not significantly invalidate the robustness of the model. Other potential future work may look into whether patients who are admitted directly for diabetes care (e.g., with either hypo- or hyperglycemia) have different mortality outcomes compared to those who were admitted primarily with non-diabetes related issues. It should be noted that such an analysis using the MIMIC-III database can be difficult as patients are admitted with a list of ordered diagnosis codes as well as a free-text, non-coded chief complaint. Deciphering the meaning of chief complaints mapped to codes (e.g., ICD-9-CM) through use of natural language processing could therefore involve assumptions. For example, a patient may be admitted for a certain condition, but experience hypo- or hyperglycemia while in the hospital – this would be impossible to distinguish from patients who were admitted primarily for failure to control their diabetes.

Further studies need to be conducted to test other machine learning models against the logistic regression and random forest models explored in this study to see if better prediction values may be obtained. Potential implementation of deep learning techniques, such as ANNs, may also allow for significant improvement on the model. Additionally, this model needs to be tested directly against established algorithms (e.g., APACHE-II, SAPS III, and SOFA) to gauge its performance against the clinical gold standards to see both efficacy and the amount of time it would take to complete both.[15,17,18] Finally these models need to be validated in prospective cohorts with analysis on length of stay as a variable. Future work could also examine whether modifications of variables in the predictive model during the ICU stay, alter mortality risk.

Conclusion

Managing diabetic patients in the ICU presents a unique problem of managing an acute emergency while maintaining care for a severe chronic disease that may have an immediate impact on survival. This study shows the utility of machine learning modeling to predict which patients are most likely to die in the ICU and gives the opportunity to better treat critical patients, by utilizing only five readily available variables. The promising results set the foundation for future work in developing rapid and robust classification algorithms that leverage the minimal amount of available data.

Acknowledgements

This work was funded in part by the Brown Center for Biomedical Informatics and National Institutes of Health grant U54GM115677. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. American Diabetes Association. Economic costs of diabetes in the U.S. in 2012. Diabetes Care. 2013Jun;36(4):1033–46.
2. Fuchs L, Chronaki CE, Park S, Novack V, Baumfeld Y, Scott D, et al. ICU admission characteristics and mortality rates among elderly and very elderly patients. Intensive Care Med. 2012;38(10):1654–61.
3. Forbes JM, Cooper ME. Mechanisms of Diabetic Complications. Physiological Reviews. 2013Jan;93(1):137–88.
4. Michalia M, Kompoti M, Koutsikou A, Paridou A, Giannopoulou P, Trikka-Graphakos E, et al. Diabetes mellitus is an independent risk factor for ICU-acquired bloodstream infections. Intensive Care Med. 2008;35(3):448–54.
5. Siegelaar SE, Hickmann M, Hoekstra JB, Holleman F, Devries J. The effect of diabetes on mortality in critically ill patients: a systematic review and meta-analysis. Crit Care. 2011;15(5).
6. Mesotten D, Berghe GVD. Glycemic targets and approaches to management of the patient with critical illness. Curr Diab Rep. 2011;12(1):101–7.
7. Mahmoodpoor A, Hamishehkar H, Shadvar K, Beigmohammadi M, Iranpour A, Sanaie S. Relationship between glycated hemoglobin, intensive care unit admission blood sugar and glucose control with ICU mortality in critically ill patients. Indian J Crit Care Med. 2016;20(2):67-71.
8. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi J-C, et al. Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data. Med Care. 2005;43(11):1130–9.
9. Charlson ME, Pompei P, Ales KL, Mackenzie C. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis. 1987;40(5):373–83.
10. Young BA, Lin E, Kroff MV, Simon G, Ciechanowski C, Ludman E. Diabetes complications severity index and risk of mortality, hospitalization, and healthcare utilization. Am Jnl of Manag Care. 2008;14(1):15-24.
11. Chang HY, Weiner JP, Richards TM, Bleich SN, Segal JB. Validating the adapted diabetes complications severity index in claims data. Am Jnl of Manag Care. 2012;18(11):721-726.
12. Kim S, Kim W, Park RW. A comparison of intensive care unit mortality prediction models through the use of data mining techniques. Healthc Inform Res. 2011Dec;17(4):232.
13. Rigla M, García-Sáez G, Pons B, Hernando ME. Artificial intelligence methodologies and their application to diabetes. J Diabetes Sci Technol. 2017May; Epub
14. Dagliati A, Marini S, Sacchi L, Cogni G, Teliti M, Tibollo V, et al. Machine learning methods to predict diabetes complications. J Diabetes Sci Technol. 2017May; Epub
15. Knaus, WA, Draper, EA, Wagner, DP, Zimmerman, JE. APACHE II: a severity of disease classification system. Crit Care Med. 1985Oct; 13(10):818-29.
16. Niewiński, G, Starczewska, M, Kański, A. Prognostic scoring systems for mortality in intensive care units--the APACHE model. Anaesthesiol Intensive Ther. 2014Mar;46(1):46-9.
17. Moreno RP, Metnitz PGH, Almeida E, Jordan B, Bauer P, Campos RA, et al. SAPS 3—From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. Intensive Care Med. 2006;32(5):1345-55.
18. Vincent J-L, Moreno R, Takala J, Willatts S, Mendonça AD, Bruining H, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. Intensive Care Med. 1996;22(7):707–10.

19. Johnson AE, Pollard TJ, Shen L, Lehman L-WH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. Sci Data. 2016;3:160035.
20. Glasheen WP, Renda A, Dong Y. Diabetes Complications Severity Index (DCSI)—Update and ICD-10 translation. J Diabetes Complications. 2017;31(6):1007–13.
21. Vincent J-L, Preiser J-C, Sprung CL, Moreno R, Sakr Y. Insulin-treated diabetes is not associated with increased mortality in critically ill patients. Crit Care. 2010;14(1).
22. Saaddine JB, Engelgau MM, Beckles GL, Gregg EW, Thompson TJ, Narayan KV. A Diabetes report card for the United States: quality of care in the 1990s. Ann Intern Med. 2002;136(8):565.
23. Hazel-Fernandez L, Li Y, Nero D, Moretz C, Slabaugh L, Meah Y, et. al. Relationship of diabetes complications severity to healthcare utilization and costs among Medicare Advantage beneficiaries. Am J Manag Care. 2015Jan;21(1):e62-70.
24. Chang HY, Weiner JP, Richards TM, Bleich SN, Segal JB. Predicting costs with diabetes complications severity index in claims data. Am J Manag Care. 2012Apr;18(4):213-9.
25. Stavem K, Hoel H, Skjaker SA, Haagensen R. Charlson comorbidity index derived from chart review or administrative data: agreement and prediction of mortality in intensive care patients. Clin Epi. 2017;(9):311–20.
26. Pirracchio R, Petersen ML, Carone M, Rigon MR, Chevret S, Laan MJVD. Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study. Lancet Respir Med. 2015Jan;3(1):42–52.