Supplementary Table S1: Summary of 61 studies qualified for quantitative descriptive analysis.

Study	Health Application in ICU	Clinical Variables	ML-Models	Dataset Size	ML Model Performance	Imputation	Feature Importance	Validation
Lee et al, (2010) [59]	Early prediction of hypotensive episodes.	Age, SBP, DBP, mean arterial blood pressure, heart rate, and medication	ANN	1,311 ICU records.	AUROC ANN: 0.918	Not clear	Yes	5-fold cross validation.
Lee et al, (2011) [60]	To predict hypotensive episode based on heart rate and blood pressure extracted from time series data	Time series data for Health Rate and Arterial Blood Pressure	Three layer Artificial Neural Network	130,325 control and 3,953 hypotensive examples	AUROC ANN: 0.934 (Best Mean)	No	Yes	5-fold cross validation
Mikhno et al, (2012) [71]	Prediction of extubation failure in Neonates with respiratory distress syndrome.	Age, demographic, birth weight, gestational age, BP, LOS, time from birth to intubation and extubation, respiratory rate, heart rate, Monocytes, Fio2, pao2, chart data, labs, and ventilation settings, and values	LOG-R	179 neonates that were intubated within 24 hours of birth.	AUC LOG-R: 0.871	No	Yes	Yes.
Celi et al, (2012) [21]	To build customized mortality prediction model on three subsets: patients with acute kidney injury , patients with subarachnoid hemorrhage (SAH) , and elderly patients	For AKI: 118 variables were used For SAH: 53 variables were used For Open-heart surgery: 41 variables were used. Refer to paper for specific details.	Logistic regression (LR), Bayesian network (BN) and artificial neural network (ANN)	1400 AKI patients 223 individual hospitalizations of subarachnoid hemorrhage (SAH) 3,261	AUROC (Best Performance) ANN: 0.875 for AKI Bayesian Network: 0.958 for subarachnoid hemorrhage ANN: 0.941	No	Yes	5-fold cross validation

	undergoing open heart surgery			patients who undergone heart surgery.	for open heart surgery			
Behar et al, (2013) [72]	False alarm detection in cardiology.	Waveform data	SVM	4050 life threatening heart rate related arrhythmia alarms.	Overall Sensitivity SVM: 0.744 Overall Specificity SVM: 0.935	Not mentioned	Not mentioned	5-fold cross validation.
Marafino et al, (2014) [42]	ML classifier to identifying a range of procedures and diagnoses from ICU clinical notes.	Discharge summaries, nursing notes, radiology reports, and if available, physician progress notes	SVM	4,191 NICU and 2,198 adult ICU patients.	Accuracy for identifying mechanical ventilation SVM: 0.982-0.987 Accuracy for identifying phototherapy use SVM: 0.924-0.940 Accuracy for identifying Jaundice SVM: 0.865-0.898 Accuracy for identifying	Not mentioned	Yes	10-fold cross validation.

					ICH SVM: 0.927-0.938			
Marafino et al, (2015) [22]	Predict mortality from ICU nursing notes.	All nursing notes only dated within 24 h of the first recorded ICU admission time.	SGD classifier	25,826 ICU patients.	Mortality Prediction Accuracy SGD classifier: 0.85-0.90	No	Yes	10-fold cross validation.

Pirracchio et al, (2015) [23]	Predict ICU mortality.	Age, gender, type of admission, GCS, SBP, heart rate, body TEMP, PaO 2/FiO2, urinary output, serum urea nitrogen concentration, WBC, serum bicarbonate concentration, sodium concentration, potassium concentration, and bilirubin concentration, immunodeficiency syndrome, metastatic cancer, and hematological cancer	Super learner	24,508 patients.	AUROC Super learner: 0.85	No	Yes	10-fold cross validation.
Wang et al, (2015) [73]	Identify false alarms generated by ICU bedside monitors.	216 relevant features to capture the characteristics of all alarms, from ABP and ECG signals	SVM,DT,BDT, and KNN	5,569 alarms.	AUC (best performer) SVM: 90.61% Please refer to paper for full model performances	Yes	Yes	5-fold cross validation.
Hoogendoorn et al, (2016) [24]	Predict ICU mortality.	Demographics, Medications, categorical measurements, observations, and continuous and ordinal measurements	LOG-R, Cox, and KNN	13,923 patients.	AUC LOG-R and Cox (predictive modelling): 0.84 KNN (patient similarity): 0.68	Yes	Yes	5-fold cross validation.

Huddar et al, (2016) [17]	Predict complications from clinical notes.	Demographic, Vitals, Lab Tests, Medication, Procedures, Diagnoses, Nursing Notes, Radiology, and comorbidities	LOG-R, SVM, DT, AB, and RF	About 700 patients.	AUC (with all combined features) LR:0.881 SVM:0.582 DT:0.619 AB:0.831 RF:0.848	Yes	Yes	5-fold cross validation.
Du et al, (2016) [83]	Predict mortality in ICU. (28 day mortality)	73 clinical variables, demographics, vitals, severity, comorbidities, lab, and interventions.	DBN, SVM, and GB	15,647 patients.	Mortality Prediction Accuracy DBN: 86.0 SVM: 84.0 GB: 85.5	No	Yes	No
Desautels et al, (2016) [51]	To study and validate a sepsis prediction method, InSight, for the new Sepsis-3 definitions, make predictions using a minimal set of variables from within the electronic health record data.	Age, gender, vital signs, LOS, ICU type, death during hospital, and other clinical variables.	Insight	22853 ICU stays	Insight AUROC of 0.781 at sepsis onset time	Yes	No	4-fold cross-validation

Awad et al, (2017) [26]	Early hospital mortality prediction. (In general Mortality, within 6 hours of admission)	Age, Type of admission, Heart Rate, SBP, TEMP, Respiratory Rate, GCS, Arterial Blood Oxygen, Fractional inspired Oxygen, Serum urea nitrogen, serum creatinine, INR, Sodium, Potassium, WBC, Bilirubin, Platelets count, Hematocrit, AIDS, and Metastatic cancer	RF, NB, DT, and PART (Rule based)	11,722 patients.	AUROC (with all attributes) RF: 0.85± 0.01 For detailed accuracy refer to paper.	Yes	Yes	Cross validation and SMOTE.
Dervishi et al, (2017) [43]	ICU risk assessment and stratification.	Heart rate, peripheral ASBP, peripheral ADBP, peripheral arterial mean blood pressure, and SpO2	A combination of fuzzy c-means clustering (FCM), and RF	127 ICU patients.	AUROC combination of FCM and RF: 93.2	Yes	Yes	Yes
Paradkar et al, (2017) [61]	Coronary artery disease detection using photo plethysmography.	Features extracted by photo plethysmography signal and its second derivative. Analyzing temporal position of systolic and diastolic phases and characteristic points	SVM	55, 35 coronary artery disease patients.	Sensitivity SVM: 0.85 Specificity SVM: 0.85	Not mentioned	Yes	Fifty iterations of 5-fold cross validation.

Ghosh et al, (2017) [52]	Septic shock.	Three waveforms: mean arterial pressure, heart rate, and respiratory rate	CHMM	1,310 samples.	Prediction of septic shock Accuracy CHMM: 83.7-85.3 for four combinations of gap interval and observation windows	No	No	Multiple 5-fold cross validation.
Kam et al, (2017) [53]	Early prediction of sepsis.	Age, SBP, pulse pressure, heart rate, body TEMP, respiration rate, WBC, pH and blood oxygen saturation	DFN, Insight, and LSTM	5443 episodes	AUC DFN (109 and 209 features): 0.887-0.915 Insight: 0.83 LSTM: 0.929	Yes	Yes	Yes
Desautels et al, (2017) [76]	Identifying patients who are likely to suffer unplanned ICU readmission	Age, vital signs, bilirubin, creatinine, international normalized ratio (INR), lactate, white cell count, platelet count and pH, FiO2 and total Glasgow Coma Score (GCS)	AdaBoost	2018 ICU episodes	ICU Readmission AUROC (Best Performance) Ensemble model: 0.7095	Yes	No	10-fold cross-validation
Morid et al, (2017) [66]	Predict Adverse event in ICU using temporal patterns AKI as case study	Labs, Vital Signs	Deep Learning and Traditional ML algorithms.	22542 patients	AUC (Best performed) Random Forest 0.809	Yes	Yes	20-fold cross validation

Meyer et al, (2018) [27]	Predict complications, bleeding, mortality, and renal failure.	Age, sex, height, weight systolic mean and diastolic arterial pressure, systolic mean and diastolic pulmonary artery pressure, central venous pressure, ventilator FiO2 setting, heart and respiratory frequency, body TEMP, Bicarbonate, glucose, hemoglobin, oxygen saturation, partial pressure of carbon dioxide and oxygen, pH level, potassium, sodium Albumin, bilirubin, urea, creatinine kinase, hemoglobin , hematocrit, international normalized ratio, creatinine, white blood cell count, lactate dehydrogenase, magnesium, partial thromboplastin time, platelets, prothrombin time Bleeding rate, and urine flow rate	RNN	5,898 patients.	AUC RNN-Bleeding: 0.75 RNN-Mortality: 0.81 RNN-Renal Failure: 0.91	No	No	10-fold cross validation with weight decay of 0.003.

Purushotham et al, (2018) [28]	Benchmarking various ML algorithms by predicting outcomes: mortality, LOS, and ICD9 group.	Three sets of features used. Set A: Total 17 data elements. Age, admission type ,TEMP, SBP, Heart rate, Body pao2/fio2 ratio, urine output, serum urea nitrogen level, WBC count, serum bicarbonate level, Sodium level, potassium level, bilirubin level, Glasgow coma scale, AIDS, hematologic malignancy, and metastatic cancer, Set B: elements from Set A plus three additional variables for the GCS, PaO2, and FiO2 Set C: 136 elements that includes 20 elements from set B	Super Learner, FFN, RNN, and Multimodal deep learning	38,425 admissions.	Exhaustive benchmarking evaluation of models was done, please refer to paper for detailed accuracy.	Yes	No	Multiple rounds of cross validation.

Ding et al, (2018) [74]	ICU patients health status monitoring.	Age, gender and other demographics, waveform data of vital signs: heart rate, mean artery blood pressure, ASBP and ADBP, and respiratory rate, and other parameters	LWPR-PCA, L-PCA, LWPR-JPCA, LWPR-MPCA, LWPR-PLS, KPCA, PLS, and PCA	18 patients.	Mean fault detection rate (95% confidence interval) LWPR-PCA:94.4 L-PCA:90.1 LWPR-JPCA:94.7 LWPR-MPCA:97.6 LWPR-PLS:95.2 KPCA:25.7 PLS:61.7 PCA:66.6	No	Yes	Stochastic leave-one-out cross validation.
Davoodi et al, (2018) [29]	Predict mortality in ICU. (In general Mortality, 48hours from ICU admission)	Age, gender, TEMP, SBP, DBP, albumin, anion gap, Bicarbonate, In-hospital mortality index, bilirubin, BUN, chloride, creatinine, glucose lab test, hematocrit, hemoglobin, INR, lactate, mechanical ventilation, platelet, potassium, PT, PTT, sodium, WBC, glucose, heart rate, mean blood pressure, respiratory rate, and SPO2	DRBFS, NB, DT,GB, DBN, and D-TSK-FC	10,972 patients.	AUROC (%) NB: 73.51 DT: 61.81 GB: 72.98 DBN: 70.07 D-TSK-FC: 66.74 DRBFS: 73.90	Yes	Yes	Hold out method of validation.

Li et al, (2018) [86]	Predicting AKI using clinical notes.	Age, gender, race, ethnicity, clinical notes during the first 24 hours of ICU admission, and 72-hour serum creatinine after admission	SVM, LOG- R, RF, GB, NB, and CNN	77,160 clinical notes of 14,470 patients.	AUCRidge LOG-R (best performed) : 0.779Please refer to paper for other models performance	No	No	5-fold cross validation.
Anand et al, (2018) [34]	Predict mortality in diabetic patients.	Gender, ethnicity, type of admission, location of admission, insurance, diagnoses, admissions, lab values (HbA1c, blood glucose, and serum creatinine) and medications	RF and combined LOG-R	4,111 patients.	AUC RF: 0.787 Binomial LOG-R: 0.785	No	Yes	5-fold cross validation.
Zahid et al, (2018) [30]	Mortality prediction. (In hospital mortality and 30 days after discharge)	Age, gender, type of admission, vitals, labs, ICU service type, primary ICD-9 code, mechanical ventilation or continuous Positive Airway Pressure during the first 24 hours in the ICU, and the receipt of vasopressor therapy during the first 24 hours, GCS, and urine output	SNN	17,150 patient with 29,149 ICU admissions.	AUROC for 30 days mortality SNN: 0.8445 (± 0.08) AUROC for in hospital mortality SNN: 0.86 (± 0.12)	No	No	10-fold cross validation.
Ren et al, (2018) [75]	Predict unexpected respiratory decompensation requiring intubation in ICU patients.	Age, gender, heart rate, BP, respiratory rate, Pao2,GCS, PaCo2, pH, hematocrit, hemoglobin, chloride,	GB, LOG-R, and FFN	12,470 patients.	AUROC (at window sizes 8 and lead time 3) GB: 89%	Yes	Yes	10-fold cross validation.

		calcium, creatinine, bilirubin, platelet, PT, PTT, INR, blood urea nitrogen, WBC, urine output, Presence of congestive failure, presence of pulmonary circulation disorders, Albumin, and Methemoglobin			LOG-R: 81% FFN: 80%			
McWilliams et al, (2018) [77]	Detecting patients that are ready for discharge from intensive care.	Age, gender, vitals, BMI, LOS, discharge delay, in-hospital mortality, readmission, hours since admission, potassium, sodium and creatinine, PAO2, SPO2, PACO2, and hemoglobin.	RF and LOG-R	7592 patients.	Accuracy in predicting patients that are ready for discharge. RF: 0.8387-0.8531 LOG-R: 0.8366-0.8494	Yes	Yes	Multiple-source cross validation.

Weissman et al, (2018) [31]	Early prediction of mortality or prolonged ICU stay using unstructured data.	Age, type of admission, vital signs, lab data, clinical notes, and modified Elixhauser score	LOG-R ,GBM, RF, and ENR applied with and without unstructur ed clinical text data	25,947 admissions.	AUC (RF and GBM performed the best) GBM on both structured and unstructured data: 0.88-0.90 GBM on unstructured only: 0.81-0.84 RF on both structured and unstructured data: 0.87-0.89 RF on unstructured only: 0.81-0.83	Yes	Yes	10-fold cross validation.
Rojas et al, (2018) [78]	Predict ICU readmissions.	demographics, vital signs, labs, medications administered during the ICU admission, ICU interventions, nursing , diagnostic tests, and ICD codes from prior admissions	GBM, SWIFT, and MEWS	42,303 ICU transfers.	AUC (with 95% confidence interval with early, late, and ever time to re-admission) GBM: 0.71–0.78 SWIFT: 0.60–0.68 MEWS: 0.52–0.62	No	No	10-fold cross validation.
Chen et al, (2018) [44]	To develop a personalized diagnostic model	Demographics, comorbidities, vital signs, laboratory data,	Logistic Regression, C4.5, LogitBoost,	38,597 adult patients	AUROC logistic regression: 0.86	Yes	Yes	5-fold cross-validation

	for kidney stone disease	and other clinical variables	Random Forest, Super Learner					
Jain et al, (2018) [32]	To Develop a Predictive Model for ICU Mortality in Patients with Acute Exacerbation COPD	Demographics, comorbidities, number of readmissions, and social economic status. Refer paper for more details.	univariate GLM-derived logistic, (2) Mean Gini-derived logistic (MGDL), and (3) random forest	1198 admissions	AUROC (Best performer) MGDL 0.778 for predicting mortality.	No	Yes	5-fold cross validation
Tang et al, (2018) [33]	To conduct a systematic comparative study of different ML algorithms for several predictive modeling problems in urgent care. (mortality and prediction, differential diagnostics, and disease marker discovery)	Demographics, vital signs, insurance, comorbidities, and labs	Deep Learning and Traditional ML algorithms.	37787 patients	For detailed accuracy by problem refer to paper	Yes	Yes	5-fold cross validation
Lin et al, (2019) [35]	Predict in hospital mortality for AKI patients.	Age, type of admission, AIDS, metastatic cancer, hematologic malignancy, and12 physiological variables.	RF, ANN, SVM, and customize d SAPS II	19,044 patients with AKI.	AUROC RF: 0.866 (0.862–0.870) SVM: 0.861 (0.855-0.868) ANN: 0.833 (0.818–0.848) SAPSII 0.795 (0.781–0.809)	Yes	Yes	5-fold cross validation.

Liu et al, (2019) [54]	Septic shock prediction in sepsis patients.	Heart rate, respiratory rate, TEMP, SBP, DBP, mean, BP, CVP, PaO2, FiO2, GCS, bilirubin, platelets, creatinine, lactate, BUN, arterial, pH, WBC, PaCO2, hemoglobin, hematocrit, potassium, epinephrine, dopamine, dobutamine, norepinephrine, phenylephrine, vasopressin, and urine output.	GLM, XGBoost, RNN, and Cox	38,418 patients.	AUC GLM: 0.87 XGBoost: 0.85 RNN: 0.93 Cox: 0.82	No	Yes	10-fold cross validation.
Zhang et al, (2019) [67]	Predict AKI in patients with Oliguric.	Age, gender, ethnicity, type of admission, heart rate, BP, respiratory rate, TEMP, elective surgery, ICU type, vasopressor, infection, mechanical ventilation, serum creatinine, glucose, bicarbonate, bilirubin, chloride, hematocrit, lactate, platelet, potassium, aPTT, INR, Sodium, BUN, WBC, albumin, urinary pH, and Urinary creatinine	LOG-R and XGBoost	6,682 patients.	AUC XGBoost: 0.842-0.878 LOG-R: 0.703-753	Yes	Yes	Cross validation done using 300 iterations.

Zimmerman et al, (2019) [68]	Predict AKI after day 1 of ICU admission.	Age, Gender, Ethnicity, Creatinine, Heart Rate, BP(SBP and DBP), TEMP, SpO2, Glucose, Bicarbonate, Hemoglobin, Platelet count, Potassium, Partial Thromboplastin, INR, Prothrombin	Multivariate LOG-R, RF, and MLP	23,950 patients.	AUC LOG-R: 0.783 with all-features.	Yes	Yes	5-fold cross validation, 10 runs of cross validation.
Kaji et al, (2019) [62]	To predict clinical events, myocardial infarction (MI), and vancomycin antibiotic administration over two week patient ICU courses	MI model contained 221 features, the sepsis model contained 225 features, and the vancomycin model contained 224 features that includes demographics, vital signs, lab results and other clinical variables	LSTM	56,841 patients	AUC of 0.876 for sepsis, 0.823 for MI, and 0.833 for vancomycin administration using LSTM	Yes	Yes	No
Barrett et al, (2019) [63]	To predict one-year mortality in patients diagnosed with acute myocardial infarction or post myocardial infarction syndrome	Demographics, Admissions, Diagnostic Information, Labs	Deep FNN and multiple traditional ML algorithm including Logistic regression, Random forest	5436 admissions	AUC of 2 Best performed models. Logistic Model Tree and Simple Logistic: 0.901	No	Yes	10-fold cross validation
Lin et al, (2019) [79]	To predict the ICU readmission of patients within 30	Demographics, vital signs, comorbidities, and GCS.	Recurrent Neural Networks (RNN)	35,334 patients with 48,393 ICU stays.	AUROC LSTM: 0.791	Yes	Yes	5-fold cross validation

	days of their discharge		with Long Short-Term Memory (LSTM)					
Caicedo-Torres et al, (2019) [36]	A multi-scale deep convolutional architecture to predict ICU mortality	22 features including demographics, Labs, Vitals, comorbidities		22.413 distinct patients	AUROC CovNet: 0.8735	Yes	Yes	5-fold cross validation
Payrovnaziri et al, (2019) [64]	To predict one-year mortality in ICU patients with Acute Myocardial Infarction(AMI) and Post Myocardial Infarction(PMI)	Demographics, Labs, Vital Signs, comorbidities, admissions, discharge summary	Deep Learning	5,436 admissions	Deep learning model achieved 82.02% accuracy	No	Yes	10-fold cross validation
Sun et al, (2019) [69]	To identify early AKI onset Using Clinical Notes and Structured Multivariate Physiological Measurements	Demographics, Labs, Vital Signs, comorbidities, Ventilations, ICU clinical notes for first 24 hours of ICU admissions	CNN and traditional ML models including SVM	A total of 16,558 ICU stays of 14,469 patients	AUC above 0.83 with SVM best performer	Yes	Yes	5-fold cross validation
Scherpf et al, (2019) [55]	Predicting sepsis onset with a recurrent neural network	Demographics, vital signs, and labs.	RNN and Insight	Patients selected based on exclusion from previous studies	Prediction 3 h prior to sepsis onset, network achieves an AUROC of 0.81	Yes	Yes	4-fold-stratified-cross-validation
Cramer et al, (2019) [45]	prognostic tools for determining a patient's risk of hospital-acquired pressure ulcers (PUs) in intensive care units	Demographic parameters, diagnosis codes, laboratory values and vitals available	Multiple Traditional ML and DL models	50,851 admissions	A weighted linear regression model showed precision 0.09 and recall 0.71 for future pressure	Yes	Yes	5-fold cross validation

					ulcers development			
Fagerström et al, (2019) [56]	Develop an improved algorithm for early detection of septic shock	Demographics, Vital signs, and labs, GCS, diagnosis.	LiSep LSTM	59,000 ICU patients	AUROC LSTM: 0.8306	Yes	Yes	6-fold cross validation
Xia et al, (2019) [46]	Build an Ensemble Approach for Improving the Outcome Prediction in Intensive Care Unit	Demographics, vital signs, and labs	eLSTM	18415 cases	AUROC Ensemble LSTM: 0.8451	Yes	Yes	No
Garcia-Gallo et al, (2020) [37]	Develop a model for predicting 1-year mortality in critical patients diagnosed with sepsis.	Demographics, comorbidities, vital signs, laboratory data, and other clinical variables	Stochastic gradient boosting (SGB)	5,650 admissions	AUROC SGB: 0.8039	Yes	Yes	No
Rongali et al, (2020) [47]	Learning Latent Space Representations to Predict Patient Outcomes	Demographics, Labs, Diagnosis, Medications	long short-term memory (LSTM) outcome prediction using comprehen sive feature relations or in short, CLOUT	7,537 patients	AUROC (Best performed) CLOUT Model: 0.89	No	Yes	No
Cherifa et al, (2020) [65]	Prediction of an Acute Hypotensive Episode During an ICU Hospitalization	Demographics, Medications, Vital signs, cause of admission	Super Learner	1,151 MIMIC patients 55 external cohort	AUROC (Best) Super Learner: 0.929	No	Yes	10-fold cross validation

Lee et al, (2020) [48]	Develop risk prediction models using the gradient boosted tree method to derive risk estimates for acute onset diseases in the near future.	Demographics, time-series clinical observations, labs, medications, diagnosis.	Decision tree based models.	21,981 MIMIC hospital admissions 14,506University of Washington Clinical Data Repository hospital admissions	Refer paper for detailed accuracy	Yes	Not clear	No
Su et al, (2020) [49]	To predict the effects of heparin treatment using machine learning methods	Demographics, labs, and medication	Shallow Neural Network and Traditional ML models	MIMIC 2789 patients eICU 575 patients	Best F1 score by shallow neural network of 87.26%	Yes	Yes	5-fold cross validation
Sha et al, (2020) [38]	To develop a gated recurrent unit-based recurrent neural network with hierarchical attention for mortality prediction	Diagnosis, comorbidities, admissions, demographics, vital signs and labs	Traditional ML and DL methods	7,537 patients who had at least two hospital admissions in MIMIC-III	AUROC best performer GRNN 0.8650±0.01	No	Yes	4-fold cross validation
Song et al, (2020) [57]	Early Detection of Late-Onset Neonatal Sepsis using ML	Demographics, vital sign data, blood gas estimations, blood cell counts, and pH levels.	Traditional ML and DL methods	7,870 patients	Best AUROC of the 48-hour prediction model was 0.861 with Logistic Regression and onset detection model was 0.868 with Gradient	Yes	Yes	10-fold cross validation

					Boosting Classifier.			
Ahmed et al, (2020) [39]	Develop a machine learning-based model to predict mortality in trauma patients admitted to ICU	Demographics, admissions, LOS, Labs, diagnosis, SOFA score	Traditional ML and DL methods	3,041 trauma patients	Best performed AUROC of Deep-FLAIM Model: 0.912	No	No	No
Eickelberg et al, (2020) [50]	To develop a novel framework to identify ICU patients with a low risk of BI as candidates for earlier EAT discontinuation	Demographics, comorbidities, diagnosis, Labs, medication, and vital signs	Traditional ML and DL methods	12,232 ICU encounters (10,290 unique patients)	Best models identified patients at low risk of BI with AUROCs up to 0.8	Yes	Yes	10-fold cross validation
Yao et al, (2020) [58]	To develop ML model to predict sepsis post-surgical procedures.	Demographics, Comorbidities, Labs, Vital Signs, SOFA scores	XGBoost and Linear Regression	3,713 patients	Best model XGBoost c-statistics 0.835	Yes	Yes	4-fold cross validation
Wang et al, (2020) [70]	predict acute kidney injury by ensemble learning and time series model	Labs, Medication, Vital Signs	Ensemble Time Series Model (ETSM).	ICUC patients 13053 and MIMIC patients 52152	Refer paper for detailed performances.	Yes	Yes	Yes
Kong et al, (2020) [40]	to predict in-hospital mortality of sepsis patients in the ICU	Demographics, vital signs, laboratory tests and comorbidities	Traditional ML models	16,688 sepsis patients	AUROC Best performed GBM model: 0.845	Yes	Not clear	Not Clear
Zhang et al, (2020) [41]	Predict hospital mortality, readmissions and LOS using both structured and unstructured data	Demographics, vital signs, lab test results, medications, diagnosis codes, as well as clinical notes	Fusion-CNN and Fusion-LSTM	39,429 unique admissions	For Mortality Prediction : Fusion LSTM - AUROC 0.871 For LOS: Fusion CNN:	Yes	Not clear	Not Clear

					AUROC 0.784 For 30-day readmissions: Fusion LSTM- AUROC 0.674			

List of Abbreviation used in Multimedia Appendix 1

Abbreviation	**Description**
AB	AdaBoost
ADBP	Arterial Diastolic Blood Pressure
AIDS	Acquired Immuno Deficiency Syndrome
AKI	Acute Kidney Injury
ANN	Artificial Neural Networks
aPTT	activated Partial Thromboplastin Time
ASBP	Arterial Systolic Blood Pressure
AUC	Area Under the Curve
AUROC	Area Under the Receiver Operating Characteristic curve
BDT	Bagged Decision Tree
BI	Bacterial infections
BMI	Body Mass Index
BP	Blood Pressure
BUN	Blood Urea Nitrogen
CHMM	Coupled Hidden Markov Models
CNN	Convolutional Neural Network
CVP	Central Venous Pressure
DBN	Deep Belief Networks
DBP	Diastolic Blood Pressure
DFN	Deep Feedforward Network
DRBFS	Deep Rule-Based Fuzzy System
DT	Decision Trees

D-TSK-FC	deep Takagi-Sugeno-Kang (TSK) fuzzy classifier
EAT	Empiric Antibiotic Therapy
ECG	Electrocardiogram
eGFR	Estimated Glomerular Filtration Rate
ENR	Elastic Net Regression
FCM	Fuzzy c-means Clustering
FFN	Feedforward Neural Networks
GB	Gradient Boosting
GBM	Gradient Boosting Machines
GCS	Glasgow Coma Score
GLM	Generalized Linear Models
ICD	International Classification of Diseases
ICD9	International Classification of Diseases-9
ICH	Intracranial Hemorrhage
ICU	Intensive Care Units
ICUC	ICU data in China
INR	International Normalized Ratio
KNN	K Nearest Neighbor
KPCA	Kernel Principal Component Analysis
LOG-R	Logistic Regression
LOS	Length Of Stay
L-PCA	Learning-type Principal Component Analysis
LSTM	Long Short-Term Memory
LWPR-JPCA	Locally Weighted Projection Regression - Joint Principal Component Analysis
LWPR-MPCA	Locally Weighted Projection Regression - Modified Principal Component Analysis
LWPR-PCA	Locally Weighted Projection Regression - Principal Component Analysis
LWPR-PLS	Locally Weighted Projection Regression - Partial Least Squares
MEWS	Modified Early Warning Score
MLP	Multi Layer Perceptron
NB	Naïve Bayes

NICU	Neonatal Intensive Care Unit
PART	Projective Adaptive Resonance Theory
PCA	Principal Component Analysis
PLS	Partial Least Squares
PT	Prothrombin Time
PTT	Partial Thromboplastin Time
RF	Random Forest
RNN	Recurrent Neural Network
SAH	Subarachnoid hemorrhage
SAPS II	Simplified Acute Physiology Score II
SBP	Systolic Blood Pressure
SGD	Stochastic Gradient Descent
SMOTE	Synthetic Minority Oversampling Technique
SNN	Self normalizing Neural Network
SVM	Support Vector Machine
SWIFT	Stability and Workload Index for Transfer
TEMP	Temperature
WBC	White Blood Cell