

# Dynamic predict in-hospital mortality risk in intensive care unit with a new deep learning of artificial intelligence

**Yu-wen Chen**

Chengdu Institute of Computer Applications, Chinese Academy of Sciences

**Yu-jie Li**

Department of Anaesthesiology, Southwest Hospital, Third Military Medical University

**Zhi-yong Yang**

Department of Anaesthesiology, Southwest Hospital, Third Military Medical University

**Kun-hua Zhong**

Chengdu Institute of Computer Applications, Chinese Academy of Science

**Li-ge Zhang**

Chengdu Institute of Computer Applications, Chinese Academy of Science

**Yang Chen**

Department of Anaesthesiology, Southwest Hospital, Third Military Medical University

**Hong-yu Zhi**

Department of Anaesthesiology, Southwest Hospital, Third Military Medical University Southwest Hospital

**Peng Deng**

Department of Anaesthesiology, Southwest Hospital, Third Military Medical University

**Dan-dan Wang**

Department of Anaesthesiology, Southwest Hospital, Third Military Medical University

**Jian-teng Gu**

Department of Anaesthesiology, Southwest Hospital, Third Military Medical University

**Jiao-lin Ning**

Department of Anaesthesiology, Southwest Hospital, Third Military Medical University

**Kai-zhi Lu**

Department of Anaesthesiology, Southwest Hospital, Third Military Medical University

**Ju Zhang** (✉ [zhangju@cigit.ac.cn](mailto:zhangju@cigit.ac.cn))

Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Science

**Zheng-yuan Xia**

Department of Anaesthesiology, Li Ka Shing Faculty of Medicine, the University of Hongkong

**Bin Yi** (✉ [yibin1974@163.com](mailto:yibin1974@163.com))

Third Military Medical University Southwest Hospital <https://orcid.org/0000-0001-5840-2086>

Loading [MathJax]/jax/output/CommonHTML/jax.js

## Research article

**Keywords:** In-hospital mortality risk, ICU, Temporal Convolution Network, Attention Mechanism, time series, Artificial Intelligence

**Posted Date:** July 24th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-44310/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

Dynamic prediction of patients' mortality risk in ICU with time series data is limited due to the high dimensionality, uncertainty with sampling intervals, and other issues. New deep learning method, temporal convolution network (TCN), makes it possible to deal with complex clinical time series data in ICU. We aimed to develop and validate it to predict mortality risk using time series data from MIMIC III dataset.

## Methods

Finally, 21139 records of ICU stays were analyzed and in total 17 physiological variables from the MIMIC III dataset were used to predict mortality risk. Then we compared the model performances of attention-based TCN with traditional artificial intelligence (AI) method.

## Results

The Area Under Receiver Operating Characteristic (AUCROC) and Area Under Precision-Recall curve (AUC-PR) of attention-based TCN for predicting the mortality risk 48 h after ICU admission were 0.837(0.824–0.850) and 0.454. The sensitivity and specificity of attention-based TCN were 67.1% and 82.6%, compared to the traditional AI method yield low sensitivity (< 50%).

## Conclusions

Attention-based TCN model achieved better performance in prediction of mortality risk with time series data than traditional AI methods and conventional score-based models. Attention-based TCN mortality risk model has the potential for helping decision-making in critical patients.

## Background

The in-hospital mortality of patients in intense care unit (ICU) is relatively high, which ranged from 6.7–44.0% worldwide (1, 2). With the development of critical care medicine, larger amounts of data helped doctors to make decisions, however sometimes overwhelm doctors in reverse. Thus, tools which help doctors make decisions based on large amounts of monitoring and clinical data are badly needed.

In the past, score-based models, such as simplified acute physiology score (SAPS II), Acute Physiology and Chronic Health Evaluation II (APACHE II), were commonly used for patients' evaluation based on prediction of mortality risk (3, 4). When applied to large populations, the diagnostic performances of (1–5–8). Recently, methods based on artificial intelligence (AI),

including conventional machine learning (ML) methods and deep learning methods have been applied to help doctors decision-making by predicting patients' mortality risk (9–11). Comparing with statistical score-based models, methods based on AI usually had better model performance, which be related to that AI methods had the advantages of dealing with complex non-linear relationships between variables and patients' outcome over score-based models. However, there are some limitations of the referred researches. One of the biggest problems were that the worst record of a period time or statistical variables (maximum, minimal or others) replaced the repeated measured variables such as vital signs to predict the mortality risk. In ICU reality, overall trend and coupling of changes between different physiological variables may provide prognostic information, which will also help to elevate the accuracy of prediction model(12). The ideal tool to help doctors' decision-making requires optimum use of all the associated routine variables especially time series data and realizing dynamic prediction. However, due to the complex of the time series data, studies about dynamic prediction using temporal clinical data are limited.

The challenges of predicting mortality risk in ICU were summarized by Ikaro et al (12): First, measurements of time series data vary from each patient meanwhile the time interval is irregular; Secondly, the choice of measurements and the trends of time series data correlate with each other. As time series deep learning models, the Long Short-Term Memory (LSTM) (13) and its derivatives Gated-Recurrent Unit (GRU) (14), were used to predict the mortality risk of ICU patients which achieved better area under Receiver Operating Characteristic (AUCROC) and area under Precision-Recall curve (AUC-PR) than conventional score-based models. However, due to that data is processed sequentially over time, the LSTM and GRU had the shortcomings of large computing, time consuming, high hardware requirements and vanishing gradients which result in difficulties of dealing with big data and applying to clinical popularization. It is widely accepted that deep learning has the shortcomings of lower explanation and large computing. Attention mechanism simulates the data processing of human brain, and nowadays were combined with LSTM or other deep learning methods to improve computational efficiency or the interpretability (7, 15, 16). However, the limitations of inefficient particularly when processing long sequences stilled existed due to the characteristics of the method itself. A better deep learning method overcome the current limitations is badly needed. Recently, a new deep learning method, temporal convolution network (TCN) was developed, with the characteristics of parallelism, fixed gradient and smaller memory of training. Furthermore, Bai et al (17) reported that the TCN has even better performance than LSTM or GRU. Developing attention-based TCN model may not only elevate the interpretability, reduce the computation complexity, but also extend the clinical use due to its higher efficiency for long sequences. Therefore, we aimed to develop attention-based TCN model to predict the in-hospital mortality risk 48 h after admission in ICUs with time series data and compare the model performances with conventional ML methods, namely random forest (RF), logistic regression (LR), decision tree and support vector machine (SVM).

## Methods

# Ethics and data extraction

Ethical approval for this study was approved after completion of the National Institute of Health(NIH) Web-based training course named “Protecting Human Research Participants” by the author Yu-wen Chen (Certification Number: 28341490). Data used for the prediction of mortality risk were extracted from the Multi-parameter Intelligent Monitoring in Intensive Care (MIMIC) database. All the data from the database had been treated with data masking to protect patients’ privacy. So there was no requirement for written informed consent in the current study. There were 61532 records of ICU stays in Beth Israel Deaconess Medical Center ICUs, including clinical notes, physiological waveforms, laboratory measurements, and nurse-verified numerical data(18). The exclusion criteria were as following: any hospital admission with multiple ICU stays or transfers between different ICU units or wards, which would reduce the ambiguity of outcomes associated with hospital admissions rather than ICU stays; patients younger than 16; patients’ whose initial ICU stay was missing or less than 48 hours; ICU events with no events in the initial 48 hours. As shown in Fig. 1A, there were 18094 were for final analysis. We divided enrolled patients randomly into three datasets, namely training dataset (12565 patients), validation dataset (2766 patients) and testing dataset (2763 patients).

## Data preprocessing

We use 17 physiologic variables (shown in Table 1) representing a subset from the Physionet/CinC Challenge 2012(12). Up to 17 variables were recorded at least once during the first 48 hours after admission. Not all variables were available in all cases. We used all raw values for time series measurements included in the score. For Glasgow Comma Score (GCS), we included GCS (Verbal response), GCS (Motor response), GCS (Eye opening) and GCS (total) as different features. The rest of the variables included weight, height, temperature, respiratory rate (RR), heart rate (HR), diastolic blood pressure (DBP), Mean blood pressure(MBP), systolic blood pressure(SBP), fraction inspired oxygen ( $FiO_2$ ), oxygen saturation (OS), pH, glucose, and capillary refill rate (CRR). When value was more than three standard deviations away from each individual mean value, it would be removed. Twelve of them were continuous and five discrete. All time series variables were resampled into hourly rate starting from ICU admission. Forward imputation was conducted for missing values. When an entire measurement was missing during the observation time, mean value from the dataset was used. For discrete variables, we performed One-Hot encoding. For continuous variables, we performed Z-score normalization to scale the feature values. Each patient’s record was summarized into a visualization data matrix  $59 \times 48$  for 48-hour observation period as the input for deep learning.

Table 1  
Physiological variables to predict the mortality risk of patients in ICU.

Sequence number	Physiological variables	Data type
1	Capillary refill rate	Discrete value
2	Diastolic blood pressure	Continuous value
3	Fraction inspired oxygen	Continuous value
4	Glascow coma scale eye opening	Discrete value
5	Glascow coma scale motor response	Discrete value
6	Glascow coma scale total	Discrete value
7	Glascow coma scale verbal response	Discrete value
8	Glucose	Continuous value
9	Heart Rate	Continuous value
10	Height	Continuous value
11	Mean blood pressure	Continuous value
12	Oxygen saturation	Continuous value
13	Respiratory rate	Continuous value
14	Systolic blood pressure	Continuous value
15	Temperature	Continuous value
16	Weight	Continuous value
17	pH	Continuous value

## Model construction for Attention-based TCN

In this work, we developed attention-based TCN model to predict the mortality risk of ICU patients with time series data and static data. The TCN is convolutional network, which is composed of causal convolution, dilated convolution, and residual connections (shown in supplemental Fig. 1–3). The causal convolution makes the TCN a strict temporal model by using data from time  $t$  and earlier in the previous to predict the status at time  $t$  when model training. TCN allows the input of convolution to be sampled at intervals to broaden the field of perception (make the most of information) due to use of the dilated convolution. The residual connections enable the network to transmit information across layers, which

are usually used to train deep network. In addition, TCN adds Dropout to each hole in the residual module to achieve regularization. Attention mechanism was introduced into the TCN model to elevate the efficiency and the interpretability. Detailed information about TCN and attention mechanism was presented in the supplementary method.

The structure of attention-based TCN model was shown in Fig. 1B: patients' raw data were preprocessed as data flow for model in-put; The patient's vital signs data are extracted by 7-level TCN, then connected to the attention layer; Finally the mortality risk were predicted by linear layer. Implementation parameters of TCN are batch\_size = 32, dropout = 0.2, kernel size = 3, levels of TCN = 7, initial learning rate = 0.02, number of hidden units per layer = 59, optimization algorithm = Adam. The loss function used is Binary Cross Entropy:

temporal convolutional network

$$prob = \frac{1}{1 + \exp(-pred)}$$

$$L = - \sum_i label_i * \log(prob_i) + (1 - label_i) * \log(1 - prob_i)$$

pred: prediction tensor with arbitrary shape

label: target tensor with values in range [0, 1]. Must have the same size as pred. Detailed information was presented in supplemental materials.

## Non-time series model construction

We also predicted the mortality risk by non-time series ML methods such as RF(19), LR, Decision Tree (DT) and SVM. Due to the limitation of these ML methods, the in-put for these models were not time series data but results of feature extraction (statistical variables, such as the minimum, maximum, average of the variables).Then the preprocessed data were used for model construction and evaluation.

## Model evaluation

Model performance was assessed by overall performance, discrimination, and calibration. The overall performance is determined by F1 score. The F1 score is defined as the harmonic mean of accuracy and recall, which considers both the precision and the recall equally. Discrimination is the capability to distinguish between those who survival and those who do not 48 h after admission in ICU by AUCROC and the area under the Precision-Recall curve (AUC-PR). The AUC-PR is sensitive for the imbalance distribution of the negative and positive data especially for an extreme small portion of positive data. Calibration is assessed by Brier score via calculating the averaged squared deviation between the predicted probability and the actual outcome.

Loading [MathJax]/jax/output/CommonHTML/jax.js

# Statistical analysis

The statistical analyses were performed using SPSS software for Windows, V.19.0 (SPSS). Quantitative variables are presented using basic descriptive statistics: mean and SD (for normal distribution data), or median and IQR (for non-normal distribution data). Comparisons among datasets were performed using the chi-square test or Fisher's exact test, or Kruskal-Wallis test. All statistical tests were two sided, and  $P$  values less than 0.05 indicated statistical significance.

## Results

### Data distribution

Finally, there were 18094 patients for analysis. Patient demographics and characteristics of the three datasets were presented in Table 2. There were no statistically significant differences in age, gender, and ICU length of stay among them. The mortality rate of our cohort was 15.4%. Though the mortality rate of patients in the testing dataset was significantly lower than that of training and validation datasets, the mortality rate of patients in test dataset was similar to that of patients in our whole cohort.

Table 2  
The baseline of patients in training, testing and validation dataset.

Variables	Training(n = 12565)	Testing(n = 2763)	Validation(n = 2766)	$P$
Age	67.4(54.2–78.9)	67.7(53.9–79.2)	67.0(55.5–78.7)	0.49
Gender(F/M)	5637/6928	1229/1534	1224/1542	0.81
ICU admission				0.04
CCU	1696	380	375	
CSRU	2288	572*	480	
MICU	4865	1037	1054	
SICU	2145	455	509	
TSICU	1571	319	348	
survival/Death	10579/1986	2389/374*	2331/435	0.01
ICU length of stay(hours)	88.7(63.6-150.3)	86.9(62.5–147.0)	89.2(64.4-148.3)	0.38
Mean (SD) presented for normally distributed continuous variables, while median (IQR) was given to those with non-normally distributed continuous variable. Unless otherwise state n is as indicated in the column headings. The portion of admission in different ICU was statistically compared with the training dataset (* $P < 0.05$ ). F, female; M, male; CCU, Coronary Care Unit; CSRU, Cardiac Surgery Recovery Unit; MICU, Medical ICU; SICU, Surgical ICU; TSICU, Trauma Surgical intense care unit.				

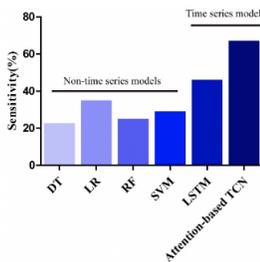
# Model performance of time series and non-time series models

As shown in Table 3 and Fig. 2A, comparing with the statistical methods, AI methods had larger AUCROC and AUC-PR which indicated a better capacity of discrimination. Though the AUCROC and AUC-PR of attention-based TCN were smaller than that of non-time series ML methods, which also had an acceptable ability of discrimination. Furthermore, comparing with non-time series ML methods, the attention-based TCN had the highest sensitivity (67.1%) and F1 score (0.46). Models with high specificity but lower sensitivity, resulting in missing patients potentially at risk, which would violate our initial purpose of helping doctors dynamic evaluating the mortality risk of patients. As for other time series methods, the sensitivity of attention-based TCN was much higher than that of model by LSTM (46.1%) based on the same database(7), with only a little difference in the AUC-PRs between them. It indicated that models developed by attention-based TCN had higher accuracy and lower omission diagnose rate than those by LSTM, which may be related to the difference between the input variables. As for model calibration, the brier score of attention-based TCN was higher than that of other conventional ML models, which may be associated with the high dimension of time series data. Taking the purpose and clinical application into consideration, due to the high sensitivity, F1 score and relative satisfied ability of discrimination, model performance of attention-based TCN was the best among the listed methods in Table 3.

Table 3

The diagnostic performances of different ML models for prediction of in-hospital mortality in the test dataset.

Methods	Sens	Spec	F1 score	Brier score	AUCROC	AUC-PR
Non-time series methods						
DT	22.7%	96.9%	0.28	0.088	0.804(0.789–0.817)	0.381
LR	35.0%	96.8%	0.43	0.081	0.838(0.824–0.850)	0.459
RF	25.1%	98.5%	0.36	0.077	0.865(0.853–0.877)	0.511
SVM	29.1%	97.9%	0.39	0.080	0.822(0.808–0.835)	0.477
SAPS-II <sup>1</sup>					0.777	0.376
APS-III <sup>1</sup>					0.750	0.357
OASIS <sup>1</sup>					0.760	0.312
Time series methods						
LSTM <sup>2</sup>	46.1%					0.451
Attention-based TCN	67.1%	82.6%	0.46	0.142	0.837(0.824–0.850)	0.454



Statistical quantifications were demonstrated with 95% CI, when applicable. ML, machine learning; attention-based TCN, attention-based Temporal Convolution Network; LR, Logistic Regression; SVM, Support Vector Machine; SAPS, Simplified Acute Physiology Score; APS, Acute Physiology Score; OASIS, Oxford Acute Severity of Illness Score; <sup>1</sup>, data referring to Hrayr et al. Scientific Data.2017; <sup>2</sup>, data referring to Ruo-xi Yu, et al. IEEE J Biomed Health Inform.2019.

## Visualization of attention weights at different time points

By visualizing the attention weight, we could clearly see that what the variable and time points were paid attention to when predicting the risk of death. The typical heatmaps for attention weight of non-survival and survival patients were shown in Fig. 2B and C. The larger portion of colored area in the

Loading [MathJax]/jax/output/CommonHTML/jax.js

that the patient was less instable. The value of the variable at

the time point represented by these colored areas contributed more than other values to the patient's death. The time point with most of the variables colored may correspond to rescue in the clinical reality. Except for the good model performance, the attention-based TCN methods may also have the potential advantage of better interpretability.

## Discussion

There are several score-based models for predicting the mortality risk, such as SAPS (3), APACHE (20), OASIS (21) and Sequential Organ Failure Assessment (SOFA) (22). These models are all non-time series and based on statistical methods, the input data are static data or statistical data, such as comorbidities and the minimum of systolic pressure in the first 24 h, which make it impossible to predict the mortality risk in the first 24 h or to update data for predicting long-term mortality risk. Despite the AUCROCs of the score-based models were satisfied, either the sensitivity or the specificity was poor (23, 24). It's not surprising that these models have been modified several times to improve their predictive performance since they first being published (25). Recently, for the complex, non-linear relationship between clinical variables and the outcome, non-time series AI methods, such as Artificial neural work (ANN), SVM, DT, RF, Naive Bayes, projective adaptive resonance theory (PART) and AutoTriage, were used to predict the mortality risk of patients in ICUs (5, 11, 24, 26, 27) with relatively satisfied model performance. However, due to the non-time series methods, all the variables are static or extracted from time series data, which makes it impossible to realize dynamic prediction. Herein, the AUCROCs and AUC-PRs of attention-based TCN model was larger than that of conventional score-based models in the same database according to Harutyunyan et al's study (8). It is a pity that Harutyunyan et al did not show the sensitivity and specificity of conventional models. Regardless of the slight difference in AUCROCs and AUC-PRs among attention-based TCN and other non-time series ML methods, the sensitivity of attention-based TCN was much higher than that of others. In clinical works, when decision-making happens, doctors should take medical history, physical examination and trend of vital signs into consideration. The ideal model for predicting mortality risk is taking both time series and static clinical data into consideration, moreover simultaneously realize dynamic prediction. Furthermore, due to the instable status of ICU patients, the sensitivity seemed more important than the specificity, for missing the potential patients at risk might be fatal. In brief, attention-based TCN method was better than non-time series methods in predicting the mortality risk of ICU patients. In addition, Hao et al (28) tried to apply attention-based TCN to language models resulting a significant elevation of model performance, which suggests attention-based TCN is a promising method for Sequence Modeling.

Recently, Yu et al (7), Harutyunyan et al (8) and Song et al (16) combined two AI methods (including one time series method) to predict the mortality risk of ICU patients with large AUCROCs and AUC-PRs but lower sensitivity (the variables and sensitivity were not presented in Harutyunyan's study). Despite the low sensitivity, there were other shortcomings of these studies. At first, Yu et al's and Harutyunyan's methods were based on LSTM, which deals with time series data sequentially from beginning to end, while TCN can do parallel processing by causal convolutions in the architecture (17). Due to the limitations of

Loading [MathJax]/jax/output/CommonHTML/jax.js be more proper for higher dimension and amounts of data

and require less in hardware, which would be more appropriate for clinical extension. Secondly, Yu et al's study included vital signs HR, SBP and temperature, while ours included RR, HR, DBP, MBP, SBP and temperature. Nowadays MBP and DBP are widely accepted as important predictors for ICU patients (29–31). So, it may be insufficient to predict the mortality risk without MBP and DBP. Moreover, some of the variables such as urinary output in Yu et al's study, which are sum or mean of clinical data in a set period time and have a longer acquisition time interval than that of vital signs. Vital signs in our study were more reasonable and easily to obtain than those in Yu et al's, meanwhile variables more frequently collected would help more for dynamic prediction. Thirdly, Harutyunyan et al's and Song et al's study focused on the algorithms, the clinical value was a bit overlooked. Fourthly, these three studies combined attention mechanism mainly aimed to elevate the efficiency of computing rather than interpretability. Comparing with time series methods combined with other AI methods for predicting mortality risk of ICU patients, our attention-based TCN method also had advantages of higher efficiency, better interpretability and easier for promotion.

As shown in Figure.3, we drew a diagram for clinical use of predicting the mortality risk of ICU patients by attention-based TCN methods: For a new critical patient, patient's baseline information and monitoring data would be put into the attention-based TCN model as data flow after automatically data preprocessing; Then the mortality risk will be predicted at different time points according to the patient's specific condition (here we predict the mortality risk 48 h after ICU admission); If the estimated mortality risk is high, the patient will receive intensive monitoring and intensive treatment; if the estimated mortality risk is low, the patient will receive intensive monitoring and routine treatment. In brief, the whole process is Warning → Intervention → Warning → Intervention → ..... → Patient outcome.

There are several limitations in this study. First of all, though the variables in our study were routine and most of them were time series, some more routine and frequently collected variables, such as lactic acid and results of arterial blood gas analysis, should be included to help elevate the prediction accuracy. Secondly, clinical data are extracted from one medical center, so the generalization ability of the model and its possibility of clinical application is not validated. Prospective multi-center studies should be carried out to investigate the clinical value of combing TCN with attention mechanism to predict patient's mortality risk using temporal clinical data.

## Conclusion

Attention-based TCN methods achieved better performance in predicting mortality risk with time series data than non-time series models, which suggested it might be be potential for decision-making in ICU by dynamic prediction of mortality risk with continuous data flow.

## Abbreviations

AI: Artificial Intelligence; ICU: Intensive Care Unit; DNN: Deep Neural Network; SVM: support vector machine; GCS: Glasgow Comma Score; ML: Machine learning; DBP: Diastolic Blood Pressure; HR: Heart

Loading [MathJax]/jax/output/CommonHTML/jax.js

Rate; TCN: Temporal Convolution Network; APS: Acute Physiology Score; SAPS II: Simplified Acute Physiology Score CRR: Capillary Refill Rate; LSTM: Long Short-Term; Memory OS: Oxygen Saturation; RNN: Recurrent Neural Network; RF: Random Forest; SBP: Systolic Blood Pressure; RR: Respiratory Rate; MBP: Mean Blood Pressure; LR: Logistic Regression; SOFA: Sequential Organ Failure Assessment; ANN: Artificial neural work; OASIS: Oxford Acute Severity of Illness Score; FiO2: Fraction Inspired Oxygen; AUCROC: Area Under The Receiver Operating Characteristic Curve; APACHE II: Acute Physiology and Chronic Health Evaluation II; MIMIC: Multi-parameter Intelligent Monitoring in Intensive Care.

## Declarations

### Acknowledgments

None.

### Authors' contributions

Conceptualization: BY, JZ and ZX; Methodology: YwC, YL, KZ and LZ; Formal analysis: ZY, YC, HZ; Writing - Original Draft: YwC and YL; Visualization: PD, DW; Project administration: JG, JN and KL.

### Funding

The project was supported by National Key R&D Program of China [No.2018YFC0116702 and No.2018YFC00116704], National Natural Science Foundation of China [No.81870422 and 81600035], Medical Innovation Capacity Improvement Program for Medical Staff of the First Affiliated Hospital of the Third Military Medical University [No. SWH2018QNKJ-27]; Technology innovation and application research and development project of Chongqing city [cstc2019jscx-msxmX0237].

### Availability of data and materials

The data that support the findings of this study are available from MIMIC III dataset but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the MIMIC III dataset with permission of Massachusetts Institute of Technology Affiliates.

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

Loading [MathJax]/jax/output/CommonHTML/jax.js eting interests.

## References

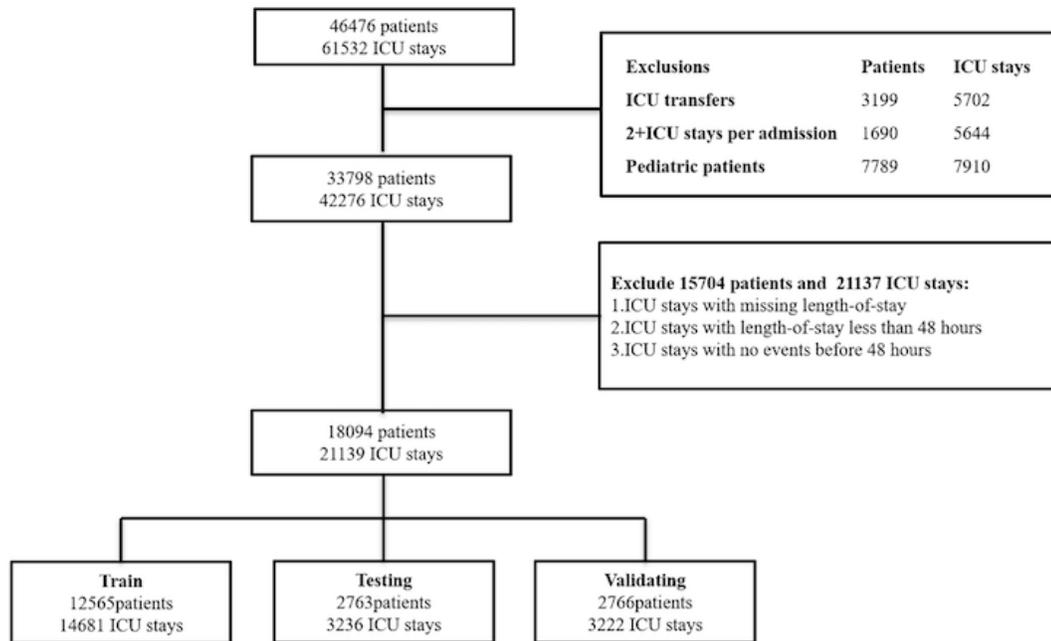
1. Weigl WA, Gorynski J, Kanski P, Hultstrom A. M. ICU mortality and variables associated with ICU survival in Poland: A nationwide database study. *Eur J Anaesthesiol.* 2018;35:949–54.
2. Weigl WAJ, Goryński P, Kański A, Hultström M. Mortality rate is higher in Polish intensive care units than in other European countries. *Intensive Care Med.* 2017;43:1430–2.
3. Le Gall J-R, Saulnier SL,F. A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study *Jama* 1993; 270:2957–63.
4. Knaus WA, Wagner EAD,DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Critical care medicine.* 1985;13:818–29.
5. Davoodi RM. M. H. Mortality prediction in intensive care units (ICUs) using a deep rule-based fuzzy classifier. *J Biomed Inform.* 2018;79:48–59.
6. Pirracchio R, Petersen, Maya L, Carone M, Rigon MR, Chevret, Sylvie van der Laan, M J. Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study. *The Lancet Respiratory Medicine* 2015;3:42–52.
7. Yu R, Zheng Y, Zhang R, Jiang Y, Poon CC. Y. Using a Multi-task Recurrent Neural Network with Attention Mechanisms to Predict Hospital Mortality of Patients. *IEEE J Biomed Health Inform* 2019.
8. Harutyunyan H, Khachatryan H, Kale DC, Ver Steeg G, Galstyan A. Multitask learning and benchmarking with clinical time series data. *Sci Data.* 2019;6:96.
9. Potes CC, Xu-Wilson B, Newth M, Inwald C, Frassica D. Joseph. A clinical prediction model to identify patients at high risk of hemodynamic instability in the pediatric intensive care unit. *Critical Care* 2017;21.
10. Christine K. Lee MS, Ira Hofer E, Gabel P, Baldi MC. Development and Validation of a Deep Neural Network Model for Prediction of Postoperative In-hospital Mortality. *Anesthesiology.* 2018;129:649–62.
11. 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.* 2011;17:232–43.
12. Silva IM, Scott G, Celi DJ, Mark LA, R. G. Predicting In-Hospital Mortality of ICU Patients: The PhysioNet/Computing in Cardiology Challenge 2012. *Comput Cardiol (2010)* 2012;39:245-8.
13. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput.* 1997;9:1735–80.
14. Dey R, Salemt FM. Gate-variants of Gated Recurrent Unit (GRU) neural networks. *IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)2017.* p. 1597 – 600.
15. Kaji DA, Zech JR, Kim JS, Cho SK, Dangayach NS, Costa AB, Oermann EK. An attention based deep learning model of clinical events in the intensive care unit. *PLoS One.* 2019;14:e0211057.
16. Huan Song DR, Jayaraman J, Thiagarajan A, Spanias. Attend and diagnose Clinical time series analysis using attention. *Association for the Advancement of Artificial Intelligence*2018.

17. Shaojie Bai JZK. Vladlen Koltun An Empirical Evaluation of Generic Convolutional and Recurrent Networks. 2018.
18. Saeed MVM, Reisner AT, Clifford G, Lehman LW, Moody G, Heldt T, Kyaw TH, Moody B, Mark RG. Multiparameter Intelligent Monitoring in Intensive Care II: a public-access intensive care unit database. *Crit Care Med*. 2011;39:952–60.
19. Xia XJT, Sohel R, Huang F. D. Random forest classification based acoustic event detection utilizing contextual-information and bottleneck features. *Pattern Recogn*. 2018;81:1–13.
20. Rivera-Fernández RV-MG, Bravo M, Aguayo-Hoyos E, Zimmerman J, Wagner D, Knaus W. The apache III prognostic system customized mortality predictions for Spanish ICU patients. *Intensive Care Med*. 1998;24:574–81.
21. Bennett CEW, Jentzer RS, Gajic J, Murphree O, Murphy DH, Mankad JG, Wiley SV, Bell BM, Barsness MR. G. W. Severity of illness assessment with application of the APACHE IV predicted mortality and outcome trends analysis in an academic cardiac intensive care unit. *J Crit Care*. 2019;50:242–6.
22. Raith EPUA, Bailey M, McGloughlin S, Maclsaac C, Bellomo R, Pilcher DV. Prognostic Accuracy of the SOFA Score, SIRS Criteria, and qSOFA Score for In-Hospital Mortality Among Adults With Suspected Infection Admitted to the Intensive Care Unit. *Jama*. 2017;317:290–300.
23. Fernando SM, Tran A, Taljaard M, Cheng W, Rochweg B, Seely AJE, Perry JJ. Prognostic Accuracy of the Quick Sequential Organ Failure Assessment for Mortality in Patients With Suspected Infection: A Systematic Review and Meta-analysis. *Ann Intern Med*. 2018;168:266–75.
24. Calvert JM, Hoffman Q, Jay JL, Desautels M, Mohamadlou T, Chettipally H, Das U. R. Using electronic health record collected clinical variables to predict medical intensive care unit mortality. *Ann Med Surg (Lond)*. 2016;11:52–7.
25. Le Gall JRNA, Hemery F, Bleriot JP, Fulgencio JP, Garrigues B, Gouzes C, Lepage E, Moine P, Villers D. Mortality prediction using SAPS II: an update for French intensive care units. *Crit Care*. 2005;9:645–52.
26. Awad AB-E-D, McNicholas M, Briggs J. Jim. Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach. *Int J Med Informatics*. 2017;108:185–95.
27. Sadeghi RB, Tanvi Romine. William. Early hospital mortality prediction using vital signals. *Smart Health* 2018;9–10:265 – 74.
28. Hong-yan Hao, Yan Wang, Yu-di Xia, Jian Zhao F-rS. Temporal Convolutional Attention-based Network For Sequence Modeling. 2020.
29. Annoni FDA, Franchi AM, Creteur F, Scolletta J, Vincent S, Taccone JL. F. S. The impact of diastolic blood pressure values on the neurological outcome of cardiac arrest patients. *Resuscitation*. 2018;130:167–73.
30. Ameloot KDDC, Ferdinande B, Dupont M, Palmers PJ, Petit T, Eertmans W, Moonen C, Belmans A, Lemmens R, Dens J, Janssens S. Mean arterial pressure of 65 mm Hg versus 85–100 mm Hg in comatose survivors after cardiac arrest: Rationale and study design of the Neuroprotect post-cardiac

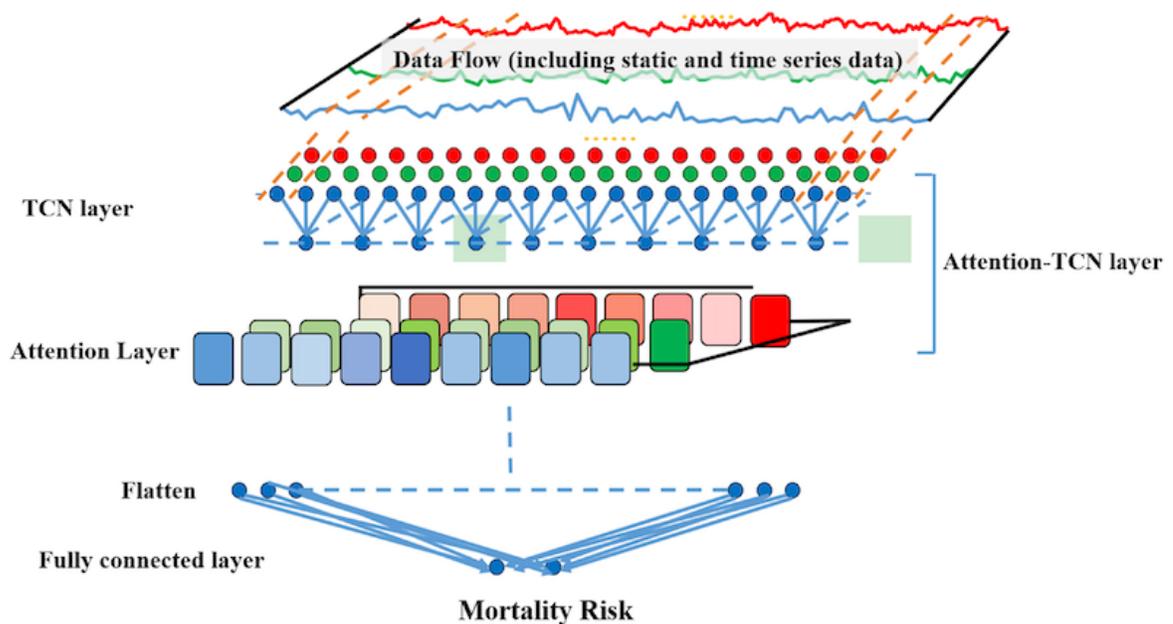
31. Houwink APRS, Bosman RJ, van der Voort PH. The association between lactate, mean arterial pressure, central venous oxygen saturation and peripheral temperature and mortality in severe sepsis: a retrospective cohort analysis. Crit Care Mar. 2016;12:20.

## Figures

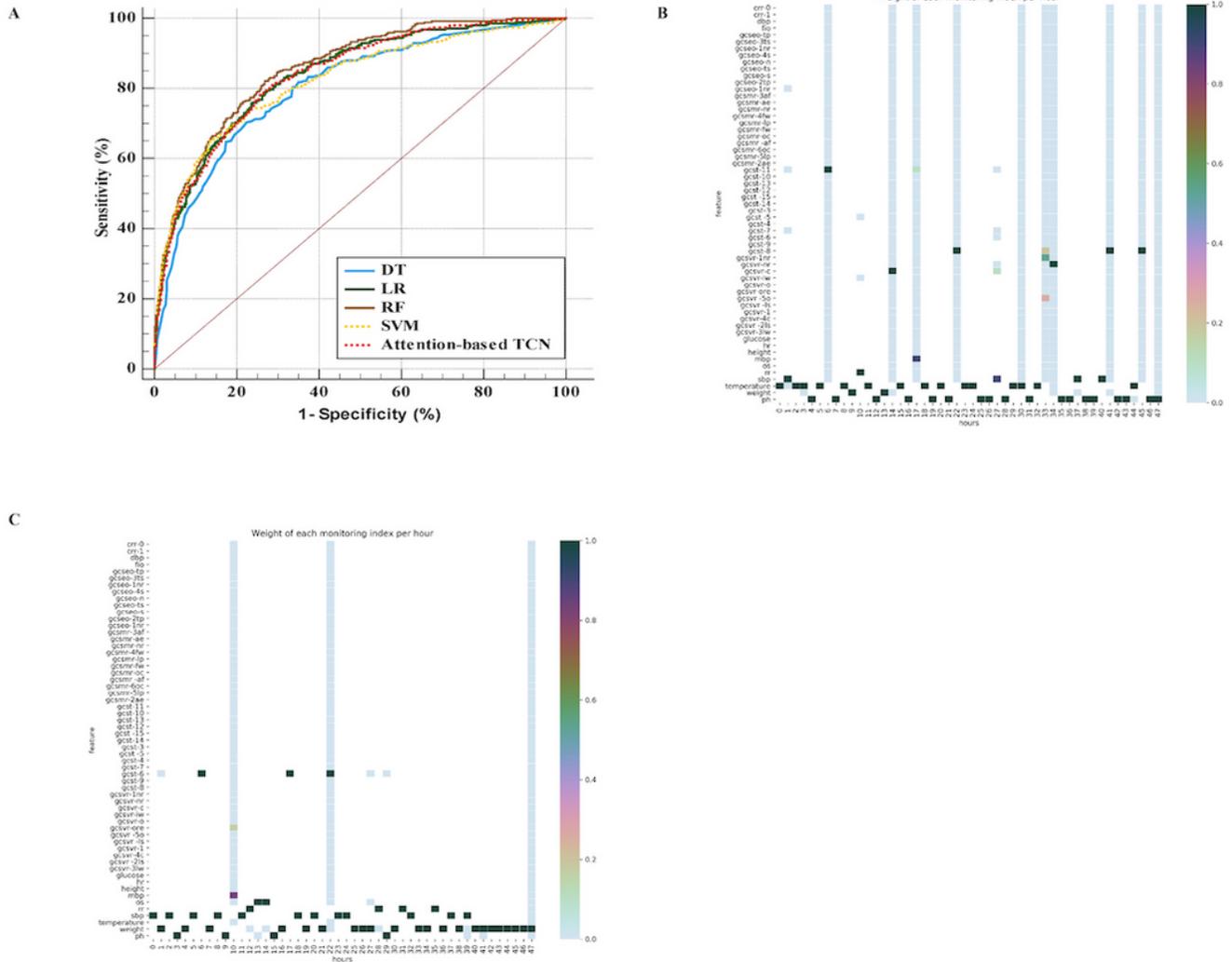
A



B

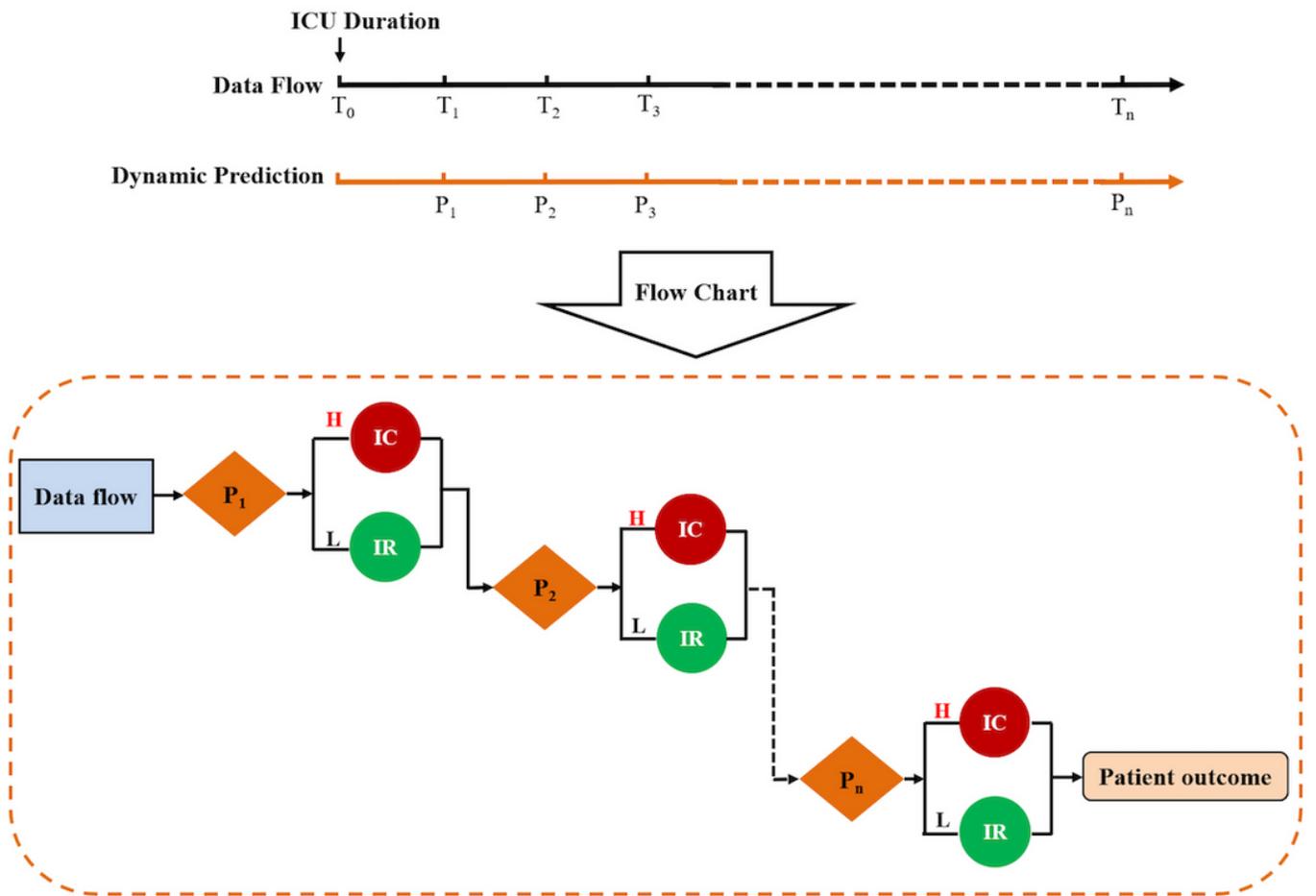


Patients enrollment from the MIMIC-III database and the structure of attention-based TCN model. (A) The flow gram of patient selection from the MIMIC-III database. (B) The structure of the attention-based TCN model for prediction of mortality risk in ICU. MIMIC, Multi-parameter Intelligent Monitoring in Intensive Care; ICU, Intensive Care Unit; TCN, Temporal Convolution Network.



**Figure 2**

The ROC curves of different AI methods and the typical visualization of attention weight. (A) The ROC curves for predicting ICU patients' in-hospital mortality 48h after admission based on different AI methods. (B) The typical heatmap for attention weight of variables and time points for the non-survival patient. (C) The typical heatmap for attention weight of variables and time points for the survival patient. AI, artificial intelligence; TCN, temporal convolution network; DT, Decision Tree; LR, Logistic Regression; RF, Random Forest; SVM, Support Vector Machine; TCN, temporal convolution network.



**Figure 3**

The diagrammatic view of dynamically prediction of mortality risk of ICU patients by attention-based TCN. (A) Data flow and dynamic prediction are briefly explained by timelines. (B) The instructions of predicting the mortality risk of a new critical patient during the treatment in ICU. T is determined by patient's main diagnosis and specific condition; P is defined as the prediction of mortality risk at different time point. H, high mortality risk; L, Low mortality risk; IC, Intensive Monitoring and Intensive Treatment; IR, Intensive Monitoring and Routine Treatment.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarmaterials.docx](#)