UNIVERSITY OF COPENHAGEN

Healthcare processes of laboratory tests for the prediction of mortality in the intensive care unit

a retrospective study based on electronic healthcare records in the USA

Zhang, Zhongheng; Goyal, Hemant; Lange, Theis; Hong, Yucai

BMJ Open

Healthcare processes of laboratory tests for the prediction of mortality in the intensive care unit: a retrospective study based on electronic healthcare records in the USA

Zhongheng Zhang, [1] Hemant Goyal,[2] Theis Lange,[3,4] Yucai Hong[1]

¹Department of Emergency Medicine, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China
²Department of Gastroenterology & Hepatology, The Wright Center for Graduate Medical Education, PA, USA
³Section of Biostatistics, University of Copenhagen, Copenhagen, Denmark
⁴Center for Statistical Science, Peking University, Beijing, China

Correspondence to
Zhongheng Zhang;
zh_zhang1984@zju.edu.cn

ABSTRACT

Objectives Healthcare process carries important prognostic information for patients, but the healthcare processes of laboratory tests have not yet been investigated for patients in the intensive care unit (ICU). The study aimed to investigate the effect of healthcare processes of laboratory tests on hospital mortality, with the hypothesis that the addition of healthcare processes could improve the discrimination for mortality outcome.
Design The study included 12 laboratory tests. There were two dimensions for each laboratory test. One was the pathophysiology value; and the other was the healthcare process variables including the clock hour, the number of measurements and the measurement time from ICU admission. Generalised additive model was employed to investigate the effect of continuous variables on mortality. Generalised linear models with and without healthcare process variables were compared for their discrimination power.
Setting ICUs in an US-based hospital.
Participants Adult patients included in the critical care big data Medical Information Mart for Intensive Care.
Primary and secondary outcome measures The hospital mortality was the primary outcome.
Results A total of 52 963 adult patients with complete ICU stay information were included for analysis. The mortality rate was 12.3%. Lower number of tests such as 1–3 times were associated with the lowest mortality for most laboratory tests. However, the hematocrit, glucose and potassium required 6–10 measurements for the first 24 hours to reach the lowest mortality rate. In n of the 12 prediction models involving laboratory tests, the addition of healthcare process variables was associated with significantly increased area under receiver operating characteristics.
Conclusions The study showed that healthcare processes of laboratory tests were independently associated with hospital mortality. The addition of healthcare processes to the pathophysiology value could increase the discrimination for mortality outcome.

Strengths and limitations of this study

► The study employed a large clinical database showing that healthcare processes of laboratory tests were independently associated with hospital mortality.
► The addition of the healthcare processes to the pathophysiology value could increase the discrimination for mortality outcome.
► Healthcare process modelling can be incorporated into the assessment of the severity of illness on the first 24 hours after intensive care unit (ICU) admission.
► Impact of healthcare processes of laboratory tests can be different for patients in ICU and those in the outpatient or general ward.
► The study is limited by a single-centre design in that different institutions may have different healthcare process that can have distinct impact on mortality outcome.

allowing the generation of a massive data via routine clinical practice. These data can help to improve healthcare management and provide large volume of data for medical researches. Laboratory tests are one of the most widely used variables in healthcare and biomedical researches because they are structuralised in most EHRs. There are generally two dimensions of a laboratory test which are the pathophysiological value and the healthcare process. The idea is that EHR laboratory test data reflect both the pathophysiology of a patient as well as the clinician's decision to order the test. Processes within the healthcare system, such as clinic hours and when hospital rounds occur, determine when most tests are ordered. Tests that are ordered at unusual times or frequency suggest that the clinician was concerned about the patient's state of health. Variables associated with healthcare processes, such as the time of day

INTRODUCTION

Electronic healthcare records (EHRs) have been widely used in modern hospitals,

of the test, can therefore provide additional insight into the patient's health, which can improve the accuracy of predictive models. Some widely investigated healthcare processes, including the timing of hospital admission, surgical procedures on weekends and handover during anaesthesia, have been shown to be associated with clinical outcomes.[1–4] More recently, healthcare processes of the laboratory tests, such as the hour of day for ordering a test, day of the week and period since the last test, have been investigated in the hospitalised patients, and the results showed that these healthcare process variables were associated with the survival outcome, independent of their pathophysiology values.[5] While pathophysiology values directly measures patients' health, the healthcare process variables such as the frequency of a test can reflect providers' knowledge of their patients' state of health.[6]

Intensive care unit (ICU) is a hospital ward where critically ill patients are closely monitored, generating a high granularity of data. Several healthcare processes have been analysed in the ICU, such as ICU admission outside daytime hours,[7] weekend admission[8 9] and the time of discharge.[10] However, the healthcare processes of laboratory tests have not been investigated in ICU. The study aimed to investigate the effect of healthcare processes of laboratory tests on hospital mortality. We hypothesised that (1) healthcare processes (eg, clock time of laboratory tests, number of measurements and the time from ICU admission to tests) of laboratory tests are independently associated with the hospital mortality and (2) the models including pathophysiology values and healthcare processes would have greater discrimination power than the models with pathophysiology values alone.

METHODS
Data source
A large US-based critical care database named Medical Information Mart for Intensive Care (MIMIC-III) was employed for this study.[11] MIMIC-III integrates deidentified, comprehensive clinical data of the patients admitted to the ICUs of Beth Israel Deaconess Medical Centre in Boston, Massachusetts, from 1 June 2001 to 31 October 2012. There were 53 423 distinct hospital admissions for adult patients (aged 16 years or above) admitted to ICUs during the study period. The study was an analysis of the third-party anonymised publicly available database with pre-existing institutional review board approval. ZZ had obtained permission to access the database after completion of the course Data or Specimens Only Research (approval no: 28776374).

Laboratory tests
The study included 12 laboratory blood tests which were bicarbonate, hematocrit, potassium, blood urea nitrogen (BUN), glucose, white cell count (WBC), creatinine, pH, sodium, albumin and bilirubin. These variables were chosen because they were included in the acute

physiological score component of the Acute Physiology and Chronic Health Evaluation (APACHE) III score, indicating that their pathophysiology values were good measures of the acute severity of illness.[12] The charting time of these tests was extracted. In the MIMIC-III database, the 'chart time' of a laboratory variable is the time at which a measurement is recorded. In almost all the cases, this is the time which best matches the time of actual measurement. The laboratory variables measured within the first 24 hours after ICU admission were obtained for the analysis. The time window was defined as the difference between the chart time of laboratory test and the ICU admission time, and it should be less than 24 hours. The reason for considering only the first 24 hours was because many severity scores like APACHE II or III were calculated during this time window, and mortality prediction was made based on variables measured in this time window.[13 14] The laboratory tests were collected for routine patient care, not for the research purpose. The quality and reliability of the tests were the same as those used for patient care. The measurements were performed by the laboratory of Beth Israel Deaconess Medical Center in Boston.

There were two dimensions for each laboratory test. One was the pathophysiology value, which was well known as a measure of health conditions. The other value was the healthcare process variables including the hour of the day when a test was ordered, the number of repeated measurements of a test for the first 24 hours after ICU admission, and the measurement time relative to ICU admission.

Statistical description
Baseline characteristics of the entire cohort were expressed as the median and IQR for continuous variables, and as the number and percentage for categorical variables. The length of stay (LOS) in the ICU and hospital were expressed as a median and IQR. Mortality outcome was expressed as the number and corresponding percentage.[15]

Generalised additive model
Generalised additive model (GAM) was employed to model the effect of continuous variables (eg, clock hour for ordering the test, number of tests within 24 hours and pathophysiology value) on mortality outcome.[16] One GAM was built for each variable of the interest. Since the mortality was a binary variable, a logistic link function was assigned to the GAMs. Other confounding factors such as the age, Sequential organ failure assessment (SOFA) score, gender and admission type were all entered into the models. Smooth terms are represented using the penalised regression splines (or similar smoothers) with smoothing parameters selected by the generalised cross-validation. The statistical output of GAMs was not directly interpretable for subject-matter audience; thus, the partial effect of the continuous variable of interest (eg, pathophysiology value, number of measurements
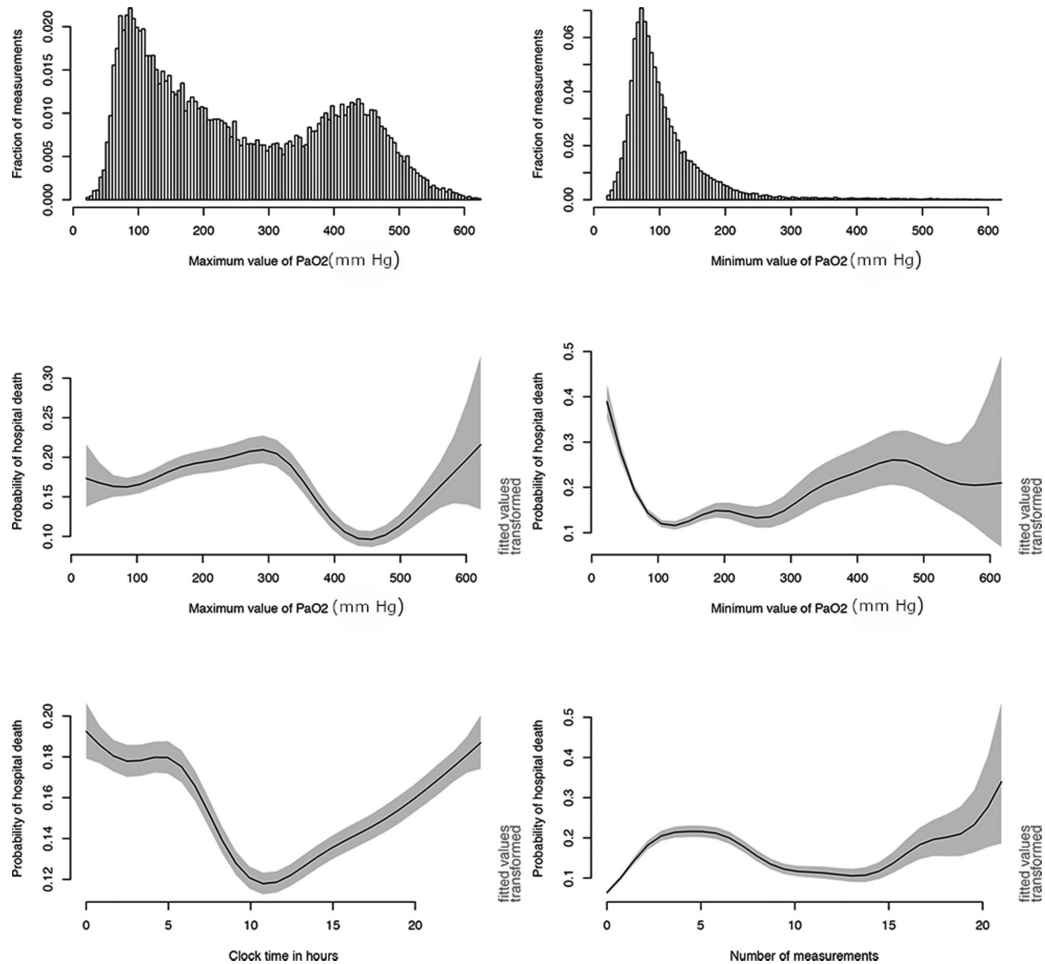
Figure 1 Distribution and impact of pathophysiology value and healthcare processes of PaO2 on mortality outcome. ICU, intensive care unit.

and hours of the day) was displayed by plotting the probability of mortality against each of the continuous variables.

Generalised linear model
While the GAM model helped to identify partial effects of the variable of interest on mortality outcome, the generalised linear model (GLM) was employed to report OR of the effects. Specifically, the continuous variables of interest were converted into the categorical variables, and then entered into the GLM for each of the variables. The response variable was the hospital mortality with logistic link function.[17 18] Pathophysiology values were categorised according to their normal ranges, as well as by the cut-off points indicated by the regression splines used in the GAM.[19] The number of measurements was categorised at 3, 6 and 10. Hours of the day were categorised by 6:00; 8:00, 14:00, 18:00 and 24:00 hours. For the investigation of the discrimination of the pathophysiology and healthcare process variables, the total sample was split into the training and testing data sets (9:1). Two logistic regression models were built for each laboratory test. Model 1 included healthcare process variables (clock hours, number of measurements) and pathophysiology value; model 2 included only the pathophysiology value. The two models were trained with the training data set,

and their discriminations for predicting mortality were calculated using the test data set. Discriminations of the models were represented by the area under receiver operating characteristics (AUROC). AUROCs of the two models were compared by using the DeLong method for each laboratory test.[20] Due to the large sample size and multiple comparisons in the analysis, we did not use p<0.05 as the cut-off for the judgement of statistical significance. Instead, we reported the point estimates along with CIs.[21]

Patient and public involvement statement
Patients and the public were not involved in the design or planning of the study.

RESULTS
Baseline characteristics of the study population
A total of 51 882 adult patients with complete ICU stay information were included for the analysis (excluding 1081 patients staying in ICU for less than 24 hours). There were a total of 29 843 male patients (56.3%). The median LOS for the entire hospitalisation was 7 days (IQR: 4–13 days). The median age of the cohort was 66 years (IQR: 53–78 years). The admission type included emergency

(44 244; 83.5%), elective surgery (7391; 14.0%) and urgent surgery (1328; 2.5%). A total of 6531 patients (12.3%) died during the study period. The median LOS in the ICU was 2 days (IQR: 1–4 days).

GAM to investigate the relationship between the pathophysiology value, healthcare process and mortality

GAM for PaO2 is shown in figure 1. The distribution of maximum and minimum values of PaO2 is shown in the top two panels. While the maximum value of PaO2 showed two peaks at 100 and 400 mm Hg, the minimum value had one peak at around 80 mm Hg. Since the maximum value of PaO2 might reflect the corrected value such as the use of mechanical ventilation, it may not reflect the true underlying lung injury. The minimum value of PaO2 was more likely to reflect the hypoxia status. As expected, the probability of the hospital death declined from over 25% at severe hypoxia (PaO2 <50 mm Hg) to 11% at PaO2 of 120 mm Hg. Thereafter, the changes in the morality outcome were unstable with wide CIs. The first panel in the third row shows the effect of hour of the day when PaO2 was measured on the mortality outcome. It appeared that PaO2 measured at 10:00 to 12:00 was associated with the lowest risk of death. The PaO2 measured at 0:00 to 6:00 was associated with high risk of death. For the number of measurements, it reached a plateau at a range of 3–6 times (table 1). Patients measured for 10–15 times for the first 24 hours showed a low mortality rate. However, as the PaO2 being obtained frequently, it is likely because another part of the blood gas was being followed closely. The PaO2 was a reflection of the blood gas measurement. Thus, we further showed the association of healthcare processes of pH on mortality (figure 2).

Generalised linear model

The results of GLM for PaO2 are shown in table 1. Table 2 displays a summary of the GLM results. As expected, pathophysiology values within the normal range were consistently associated with the lowest mortality for all laboratory tests. The number of tests associated with lowest mortality varied among these laboratory tests. Lower number of tests such as 1–3 times were associated with the lowest mortality for most laboratory tests. However, the hematocrit, glucose and potassium required 6–10 measurements for the first 24 hours to reach the lowest mortality rate. The derangements in these blood parameters are associated with acute pathophysiological deterioration or even sudden death. Therefore, close monitoring of these blood parameters is required to improve survival outcome. For example, hyperkalaemia can cause sudden cardiac arrest, acute drop in hematocrit is an indication of massive bleeding and hypoglycaemia can cause irreversible brain injury. Other tests such as albumin, WBC and bicarbonate were not a reflection of sudden deterioration, but were associated with survival outcome. However, if some urgent tests such as hematocrit, glucose and potassium were not ordered within 24 hours after ICU admission, the mortality rate could get escalated. For

Table 1 OR for the association of healthcare process variables and hospital mortality by adjusting for pathophysiological value of PaO2

Variables	OR (95% CI)	P value
Age (with every 1 year increase)	1.03 (1.02 to 1.03)	<0.001
Gender (female as reference)	0.93 (0.88 to 0.99)	0.026
Admission type (elective as reference)		
Emergency	3.61 (3.14 to 4.16)	<0.001
Urgent	3.25 (2.59 to 4.08)	<0.001
SOFA (with every 1 point increase)	1.32 (1.31 to 1.33)	<0.001
Number of measurement (none as reference)*		
1–3	2.11 (1.76 to 2.53)	<0.001
3–6	2.02 (1.59 to 2.56)	<0.001
6–10	1.08 (0.83 to 1.41)	0.558
>10	0.59 (0.44 to 0.80)	0.001
Minimum PaCO2 (<50 mm Hg as reference)		
>150 mm Hg	1.06 (0.92 to 1.23)	0.425
50–70 mm Hg	0.82 (0.72 to 0.94)	0.004
70–80 mm Hg	0.63 (0.54 to 0.73)	<0.001
80–150 mm Hg	0.64 (0.56 to 0.72)	<0.001
Clock time		
0:00–6:00	1.10 (1.00 to 1.22)	0.112
6:00–8:00	0.93 (0.85 to 1.02)	0.015
8:00–14:00	0.89 (0.80 to 0.98)	0.054
14:00–18:00	0.91 (0.84 to 1.00)	0.417
18:00–24:00	0.96 (0.88 to 1.06)	0.009
Measurement time from ICU admission (none as reference)		
Within 6 hours	1.19 (1.04 to 1.35)	0.163
From 6 to 12 hours	1.07 (0.97 to 1.19)	<0.001
Above 12 hours	1.32 (1.19 to 1.47)	<0.001

*The number of measurements was calculated as the number of measurements ordered within the first 24 hours after ICU admission. The generalised linear model was employed for binary outcome mortality.
ICU, intensive care unit; PaO2, arterial PaO2 of oxygen.

most of the tests, the measurement at 00:00–6:00 hours was associated with the lowest mortality rate.

Comparison of models with and without healthcare process variables

Comparison of models with and without healthcare process variables is shown in table 3. For albumin, hematocrit and potassium, the addition of healthcare process variables was not associated with improved prediction accuracy, as compared with the model with pathophysiology value alone. For other laboratory tests, the addition of healthcare process variables was associated with significantly increased AUROC.
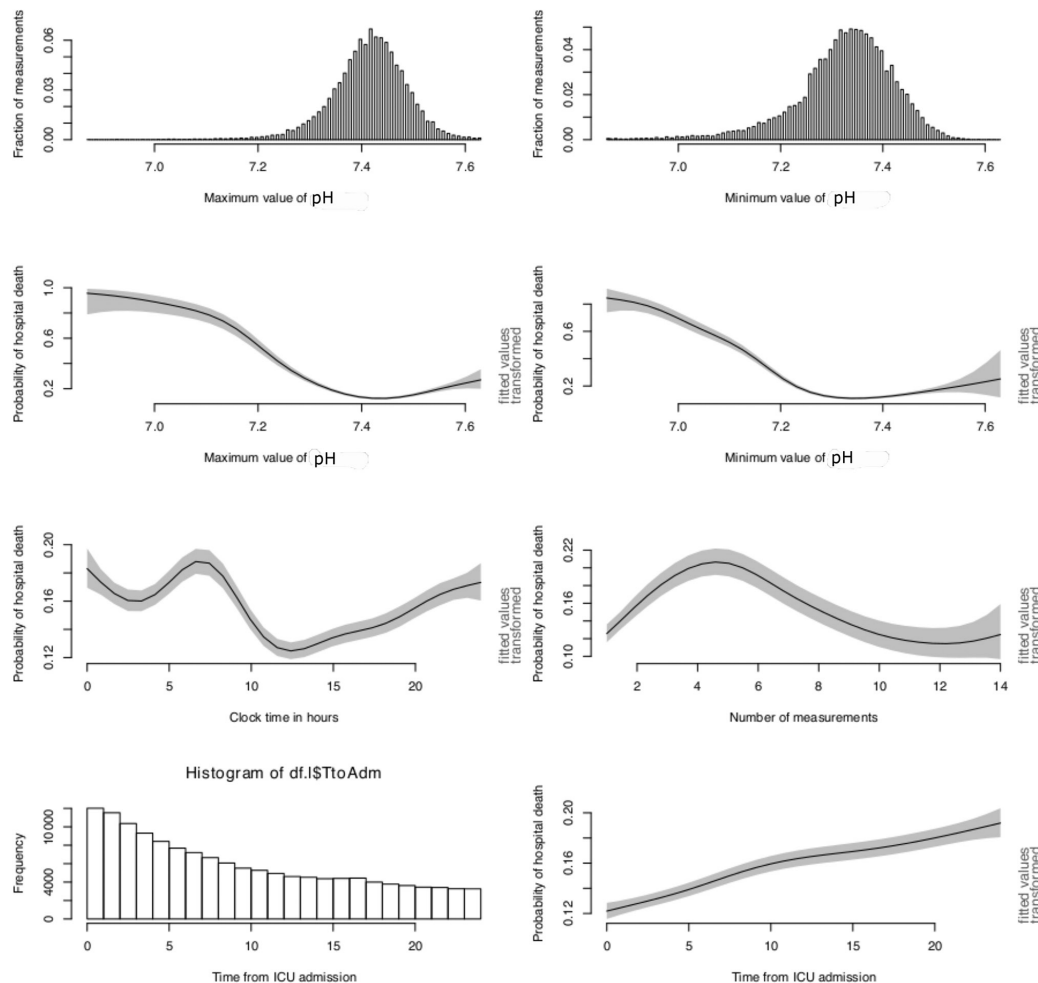
Figure 2 Distribution and impact of pathophysiology value and healthcare processes of pH on mortality outcome. ICU, intensive care unit.

DISCUSSION

Our study found that healthcare processes (eg, clock time of laboratory tests, number of measurements and the time from ICU admission to tests) of the laboratory tests were independently associated with the hospital mortality; and the addition of the healthcare processes to the pathophysiology value could increase the discrimination for mortality outcome. The healthcare process patterns of the laboratory tests were different according to the type of tests. For example, 6–10 measurements of hematocrit, glucose and potassium for the first 24 hours after the ICU admission were associated with the lowest mortality rate. For other less urgent tests such as BUN, WBC and creatinine, one to three measurements in the first 24 hours is associated with the lowest mortality.

The healthcare process of laboratory tests has been investigated in hospitalised patients. Agniel and coworkers enrolled 669 452 hospitalised patients and investigated the association of healthcare process of laboratory tests and 3 year survival.[5] They found that the WBC measured in the early morning was associated with the highest risk of death (eg, the lowest mortality was likely achieved among the people with no tests at all), which is in contrast to the findings in our study that WBC measured in the early

morning was associated with the lowest hospital mortality. This could be explained by the different settings of the two studies. There was a substantial difference in the healthcare process in ICU and general hospital wards and/ or outpatient. It can be proposed that close monitoring for ICU patients is required to circumvent catastrophic deterioration, and measurements outside daytime hours might help to identify potential lethal conditions and thus to improve outcome. Another explanation is that the physician will typically stop treatment if the chance of survival is very low. Perhaps this is why high ICU mortality was seen when tests were not ordered.[22] However, the condition can be different for general ward or outpatient, if a measurement outside of daytime hour is ordered, it is suggestive of poor conditions irrespective of the pathophysiology value of the test. Agniel's study showed that the predictive models combined with the pathophysiology values and healthcare process were better than the models with either dimension alone in 168 out of the 272 (68%) investigated laboratory tests. The result was consistent with our study that nine out of 12 (75%) of the tests had the better predictive model with both dimensions included.

Table 2 Association of healthcare process and pathophysiology aspects of laboratory tests with mortality outcome

Variables	Pathophysiological value		Number of test		Clock time		Time from ICU admission	
	Lowest mortality	Highest mortality	Lowest mortality	Highest mortality	Lowest mortality	Highest mortality	Lowest mortality	Highest mortality
Bicarbonate (mEq/L)	20–30	<10	Not tested	3–6	0:00–6:00	18:00–24:00	Over 12 hours	Within 6 hours
Hematocrit (%)	40–50	>50	>10	3–6	0:00–6:00	6:00–8:00	Over 12 hours	Within 6 hours
Potassium (mEq/L)	3.5–4.5	5.5–8.2	6–10	3–6	0:00–6:00	6:00–8:00	Over 12 hours	Within 6 hours
BUN (mg/dL)	Not tested	80–172	1–3	>10	0:00–6:00	18:00–24:00	Over 12 hours	Within 6 hours
Glucose (mg/dL)	60–200	350–861	6–10	Not tested	0:00–6:00	18:00–24:00	Over 12 hours	Within 6 hours
WBC (K/uL)	3–10	>20	1–3	6–10	0:00–6:00	18:00–24:00	Over 12 hours	Within 6 hours
Creatinine (mg/dL)	0.5–1.5	1.95–4	1–3	>10	0:00–6:00	18:00–24:00	Over 12 hours	Within 6 hours
pH	7.35–7.45	<7.1	Not tested	3–6	14:00–18:00	6:00–8:00	Within 6 hours	Over 12 hours
Sodium (mEq/L)	135–145	150–165	3–6	Not tested	0:00–6:00	18:00–24:00	Over 12 hours	6–12 hours
Albumin (g/dL)	4–5.4	<1.5	Not tested	1–3	0:00–6:00	14:00–18:00	6–12 hours	Over 12 hours
Bilirubin (mg/dL)	<2	>8	3–6	6–10	0:00–6:00	8:00–14:00	Over 12 hours	6–12 hours

BUN, blood urea nitrogen; ICU, intensive care unit; WBC, white cell count.

There are several clinical implications brought by the present study. First, although a causal statement cannot be made, this study indicates that physicians should not delay or omit ordering some critical tests such as glucose, hematocrit and potassium for ICU patients. Second, different patterns of healthcare process could be identified for different tests, and such information can be employed to identify outlier clinicians or practices. Third, healthcare process modelling can be incorporated into the assessment of the severity of illness within the first 24 hours after ICU admission. Some laboratory tests required in severity scores might be missing, and current modelling practice is to impute missing values or simply exclude patients with missing values.[23] However, our study found that the missing value per se carried important prognostic information and should be added to the assessment of the severity of illness. Similarly, a test ordered in the early morning can be different from that ordered at night. These healthcare process variables can be added to modelling predictive models. Fourth, healthcare process dimension can be incorporated into designing clinical studies. Randomised controlled trials (RCT) are often designed to minimise variations in clinical practice, which is a reason for the difference in effect sizes between the RCT and real-world setting.[24 25] Since the healthcare process carries important prognostic information, trials can be designed by a pragmatic stratification of patients according to likely disease severity according to healthcare process variables.

There are several limitations in the current study. First, the study investigated only 12 tests and it is unknown whether the results could be generalised to other tests. These 12 tests were specifically chosen because they were important prognostic tests in the first 24 hours after ICU admission (eg, the tests included in the acute physiology score for APACHE III). Other less commonly ordered tests were not included because they would produce sparsity in the data frame. Second, the study was based on a single centre and the healthcare process pattern may not be generalisable to other institutions, because the timing and frequency of tests can depend on local hospital practice patterns. However, we believe that the healthcare process variables can add more prognostic information to the pathophysiology values, irrespective of what healthcare process pattern it is. Furthermore, the healthcare processes can be influenced by the control of the practitioner, and it would be better to included variables unlikely being changed by preferences for predictive purpose. Third, underlying reasons for repeated measurements within 24 hours could not be investigated in the study, which was a source of bias known as the confounding by indication. For example, repeated measurements of PaO2 can be indicative of deteriorating respiratory function, or simply due to routines for patients on mechanical ventilation. These issues can be addressed specifically for each test in future studies.

Table 3 Discrimination performance of laboratory test for the prediction of hospital mortality in logistic regression models with and without healthcare process variables

Variables	AUC for model 1 (95% CI)	AUC for model 2 (95% CI)	P value*
PaO2	0.664 (0.634 to 0.694)	0.639 (0.609 to 0.670)	0.001
Hematocrit	0.605 (0.573 to 0.636)	0.583 (0.554 to 0.613)	0.147
Potassium	0.612 (0.580 to 0.644)	0.600 (0.570 to 0.631)	0.356
Albumin	0.594 (0.567 to 0.622)	0.590 (0.563 to 0.618)	0.333
Glucose	0.630 (0.601 to 0.660)	0.555 (0.534 to 0.575)	<0.001
Bilirubin	0.611 (0.581 to 0.640)	0.603 (0.574 to 0.632)	0.022
Bicarbonate	0.692 (0.661 to 0.723)	0.664 (0.635 to 0.694)	0.001
WBC	0.634 (0.603 to 0.666)	0.610 (0.580 to 0.641)	0.015
BUN	0.681 (0.650 to 0.712)	0.665 (0.635 to 0.695)	0.025
pH	0.665 (0.635 to 0.695)	0.636 (0.605 to 0.666)	0.001
Sodium	0.615 (0.583 to 0.647)	0.587 (0.558 to 0.616)	0.043
Creatinine	0.671 (0.640 to 0.701)	0.620 (0.590 to 0.650)	<0.001

*The AUCs of two ROC curves were compared using DeLong method, and p values for the comparisons were reported.
Model 2 included only the pathophysiology value; model 1 included healthcare process variables (number of test for the first 24 hours, clock time of the test and the measurement time relative to ICU admission) and the pathophysiology value of corresponding test.
AUC, area under curve; BUN, blood urea nitrogen; ROC, receiver operating characteristic; WBC, white cell count.

CONCLUSION

In conclusion, this study shows that the healthcare processes of laboratory tests were independently associated with the hospital mortality; and the addition of these healthcare process variables to the pathophysiology value could increase the discrimination of the model for predicting the mortality outcome. However, variations of the impacts of healthcare process on mortality exist for different tests, which requires further investigations.

REFERENCES

1. Jones PM, Cherry RA, Allen BN, *et al*. Association Between handover of anesthesia care and adverse postoperative outcomes among patients undergoing major surgery. *JAMA* 2018;319:143–53.
2. Schütte-Nütgen K, Thölking G, Dahmen M, *et al*. Is there a "weekend effect" in kidney transplantation? *PLoS One* 2017;12:e0190227.
3. Hirschy R, Sterk E, Dobersztyn R, *et al*. Time Spent in the Emergency Department and Outcomes in Patients With Severe Sepsis and Septic Shock. *Adv Emerg Nurs J* 2018;40:94–103.
4. Ramsden L, McColgan MP, Rossor T, *et al*. Paediatric outcomes and timing of admission. *Arch Dis Child* 2018;103:611–7.
5. Agniel D, Kohane IS, Weber GM. Biases in electronic health record data due to processes within the healthcare system: retrospective observational study. *BMJ* 2018;361:k1479.
6. Weber GM, Kohane IS. Extracting physician group intelligence from electronic health records to support evidence based medicine. *PLoS One* 2013;8:e64933.
7. Buck DL, Christiansen CF, Christensen S, *et al*. Out-of-hours intensive care unit admission and 90-day mortality: a Danish nationwide cohort study. *Acta Anaesthesiol Scand* 2018;62:77.
8. Brunot V, Landreau L, Corne P, *et al*. Mortality Associated with Night and Weekend Admissions to ICU with On-Site Intensivist Coverage: Results of a Nine-Year Cohort Study (2006-2014). *PLoS One* 2016;11:e0168548.
9. Zampieri FG, Lisboa TC, Correa TD, *et al*. Role of organisational factors on the 'weekend effect' in critically ill patients in Brazil: a retrospective cohort analysis. *BMJ Open* 2018;8:e018541.
10. Yang S, Wang Z, Liu Z, *et al*. Association between time of discharge from ICU and hospital mortality: a systematic review and meta-analysis. *Crit Care* 2016;20:390.
11. Johnson AE, Pollard TJ, Shen L, *et al*. MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:160035.
12. Knaus WA, Wagner DP, Draper EA, *et al*. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991;100:1619–36.
13. Knaus WA, Draper EA, Wagner DP, *et al*. APACHE II: a severity of disease classification system. *Crit Care Med* 1985;13:818–29.
14. Chan T, Bleszynski MS, Buczkowski AK. Evaluation of APACHE-IV predictive scoring in surgical abdominal sepsis: a retrospective cohort study. *J Clin Diagn Res* 2016;10:PC16–18.
15. Zhang Z. Univariate description and bivariate statistical inference: the first step delving into data. *Ann Transl Med* 2016;4:91.
16. Wood SN. *Generalized additive models: an introduction with R*. 2nd ed: Chapman and Hall/CRC, 2017.
17. Zhang Z. Model building strategy for logistic regression: purposeful selection. *Ann Transl Med* 2016;4:111.
18. Tolles J, Meurer WJ. Logistic regression: relating patient characteristics to outcomes. *JAMA* 2016;316:533–4.
19. Zhang Z, Zhang H, Khanal MK. Development of scoring system for risk stratification in clinical medicine: a step-by-step tutorial. *Ann Transl Med* 2017;5:436.

20. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* 1988;44:837.

21. Lederer DJ, Bell SC, Branson RD, *et al*. Control of confounding and reporting of results in causal inference studies. guidance for authors from editors of respiratory, sleep, and critical care journals. *Ann Am Thorac Soc* 2019;16:22–8.

22. Wilkinson DJC, Savulescu J. Knowing when to stop: futility in the ICU. *Curr Opin Anaesthesiol* 2011;24:160–5.

23. Zhang Z. Missing data imputation: focusing on single imputation. *Ann Transl Med* 2016;4:9.

24. Oude Rengerink K, Kalkman S, Collier S, *et al*. Series: Pragmatic trials and real world evidence: Paper 3. Patient selection challenges and consequences. *J Clin Epidemiol* 2017;89:173–80.

25. Zhang Z. Big data and clinical research: focusing on the area of critical care medicine in mainland China. *Quant Imaging Med Surg* 2014;4:426–9.