

RESEARCH

Open Access



Machine learning for the prediction of volume responsiveness in patients with oliguric acute kidney injury in critical care

Zhongheng Zhang^{1*} , Kwok M. Ho² and Yucai Hong¹

Abstract

Background and objectives: Excess fluid balance in acute kidney injury (AKI) may be harmful, and conversely, some patients may respond to fluid challenges. This study aimed to develop a prediction model that can be used to differentiate between volume-responsive (VR) and volume-unresponsive (VU) AKI.

Methods: AKI patients with urine output < 0.5 ml/kg/h for the first 6 h after ICU admission and fluid intake > 5 l in the following 6 h in the US-based critical care database (Medical Information Mart for Intensive Care (MIMIC-III)) were considered. Patients who received diuretics and renal replacement on day 1 were excluded. Two predictive models, using either machine learning extreme gradient boosting (XGBoost) or logistic regression, were developed to predict urine output > 0.65 ml/kg/h during 18 h succeeding the initial 6 h for assessing oliguria. Established models were assessed by using out-of-sample validation. The whole sample was split into training and testing samples by the ratio of 3:1.

Main results: Of the 6682 patients included in the analysis, 2456 (36.8%) patients were volume responsive with an increase in urine output after receiving > 5 l fluid. Urinary creatinine, blood urea nitrogen (BUN), age, and albumin were the important predictors of VR. The machine learning XGBoost model outperformed the traditional logistic regression model in differentiating between the VR and VU groups (AU-ROC, 0.860; 95% CI, 0.842 to 0.878 vs. 0.728; 95% CI 0.703 to 0.753, respectively).

Conclusions: The XGBoost model was able to differentiate between patients who would and would not respond to fluid intake in urine output better than a traditional logistic regression model. This result suggests that machine learning techniques have the potential to improve the development and validation of predictive modeling in critical care research.

Keywords: Acute kidney injury, Critical care, Extreme gradient boosting, Urine output, Predictive modeling

Background

Acute kidney injury (AKI) is common in the intensive care unit (ICU), and there is evidence that even a small increase in serum creatinine may be associated with increased risk of mortality [1–3]. AKI can be defined by either an elevation in serum creatinine or a reduced urine output according to the Kidney Disease: Improving Global Outcomes (KDIGO) guidelines [4]. Oliguric AKI constitutes a substantial proportion of the overall AKI population, and it imposes a

great challenge for fluid management. Pathophysiologically, oliguria may represent an adaptive response in AKI, and once an effective circulatory volume is restored by positive fluid balance, urine output would improve. Under this circumstance, fluid administration or positive fluid balance can be considered beneficial and those who improve with more fluid can be considered as having volume-responsive (VR) AKI.

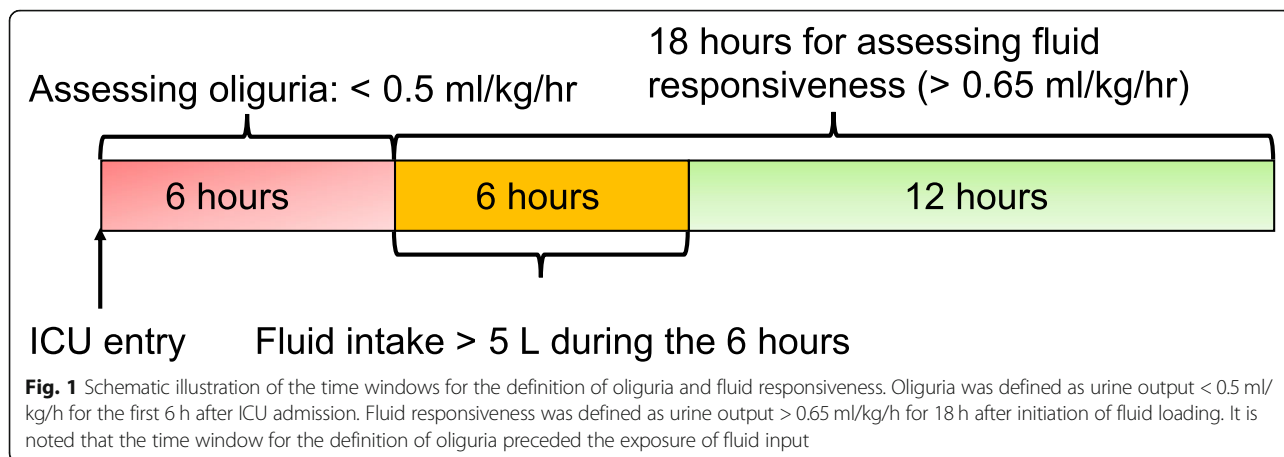
Intravenous fluid challenges are often used in critical care to restore blood pressure to improve urine output in patients with hypotension and oliguria, respectively [5, 6]. Recent epidemiological evidence, however, suggested that large positive fluid balance may not be useful

* Correspondence: zh_zhang1984@zju.edu.cn

¹Department of Emergency Medicine, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, No. 3, East Qingchun Road, Hangzhou 310016, Zhejiang Province, China

Full list of author information is available at the end of the article





to improve urine output in many patients with AKI and can even be harmful in worsening renal function through a number of possible mechanisms including excessive kidney edema [7]. This latter type of AKI can be considered as volume-unresponsive (VU) AKI [8]. Because early improvement in organ dysfunction is associated with an improved survival [9], it would be best if we can adjust the fluid treatment strategy by identifying which patients are having VR or VU AKI.

In animal models, fractional excretion of electrolytes was found to perform well in early differentiation between VR- and VU-AKI [10]. However, these promising results could not be replicated in human studies. Legrand et al.

investigated the discrimination of urinary sodium to distinguish volume responsiveness, and it was found to have limited predictive value [11]. Currently, there is little clinical information on how we can identify patients with AKI who are VR and VU in terms of urinary output response. We hypothesized that advanced machine learning techniques may be useful to identify the most important clinical factors that can differentiate between patients with VR and VU AKI. In this study, we aimed to use machine learning techniques to develop and validate an AKI fluid-responsiveness model, called extreme gradient boosting (XGBoost), and compared the performance of this model to a conventional logistic regression model.

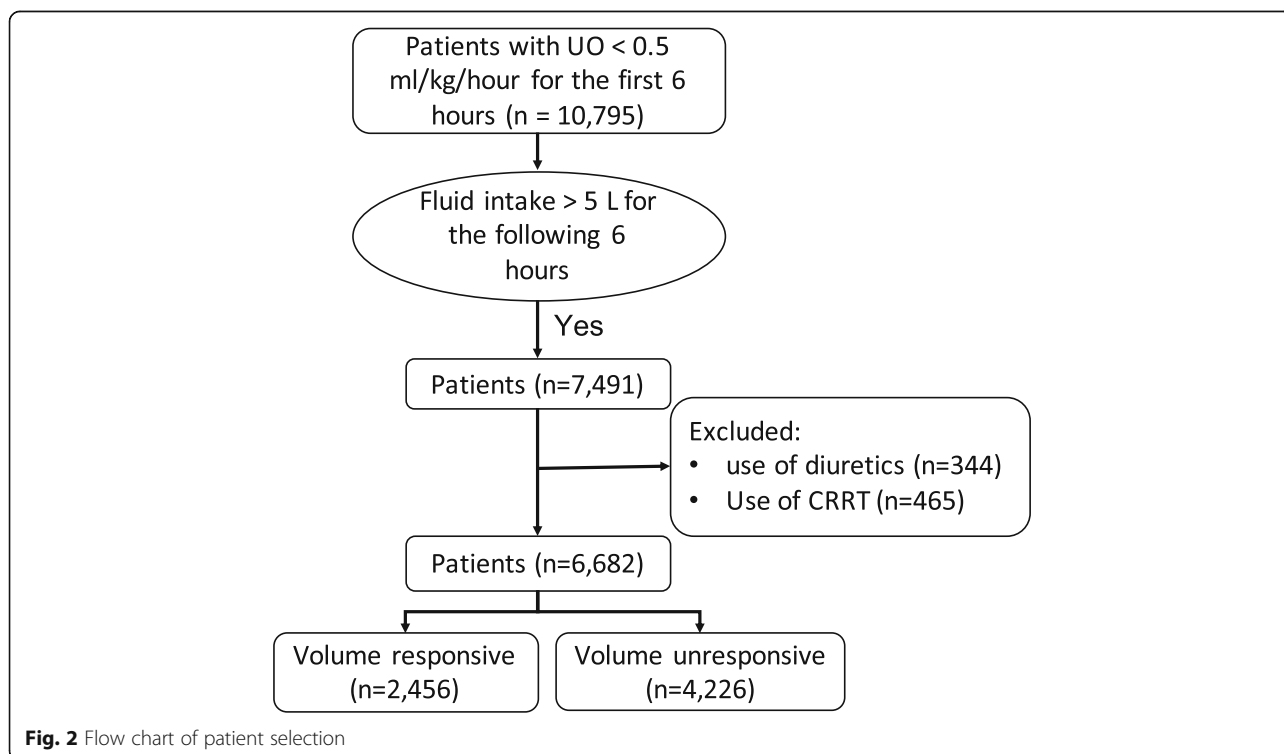


Table 1 Characteristics between fluid responsive and non-responsive groups

Variables	Volume unresponsive (n = 4226)	Volume responsive (n = 2456)	p value
Demographic variables			
Gender male, n (%)	2203 (52.1)	1495 (60.9)	< 0.001
Age (mean (SD))	68.81 (15.14)	63.82 (16.67)	< 0.001
Ethnicity, n (%)			0.013
Asian	58 (1.4)	45 (1.8)	
Black	351 (8.3)	197 (8.0)	
Hispanic	94 (2.2)	83 (3.4)	
Unknown	517 (12.2)	326 (13.3)	
White	3206 (75.9)	1805 (73.5)	
Admission type (%)			< 0.001
Elective	656 (15.5)	485 (19.7)	
Emergency	3454 (81.7)	1908 (77.7)	
Urgent	116 (2.7)	63 (2.6)	
Elective surgery, n (%)	574 (13.6)	448 (18.2)	< 0.001
ICU type (%)			< 0.001
CCU	391 (9.3)	180 (7.3)	
CSRU	619 (14.6)	646 (26.3)	
MICU	1939 (45.9)	932 (37.9)	
SICU	771 (18.2)	385 (15.7)	
TSICU	506 (12.0)	313 (12.7)	
Vasopressor, n (%)	1266 (30.0)	561 (22.8)	< 0.001
Infection, n (%)	2639 (62.4)	1168 (47.6)	< 0.001
Mechanical ventilation, n (%)	1804 (42.7)	768 (31.3)	< 0.001
Serum laboratory variables, mean (SD) if not otherwise specified			
Serum creatinine ($\mu\text{mol/l}$)	1.74 (1.50)	1.26 (0.98)	< 0.001
Maximum glucose (mg/dl)	186.32 (99.16)	180.38 (85.71)	0.013
Minimum bicarbonate (mmol/l)	20.82 (5.41)	22.24 (4.32)	< 0.001
Maximum bilirubin (mg/dl, median [IQR])	0.80 [0.50, 1.90]	0.70 [0.40, 1.30]	< 0.001
Maximum bicarbonate (mmol/l)	23.80 (4.94)	24.88 (3.97)	< 0.001
Minimum chloride (mmol/l)	102.93 (6.65)	102.63 (6.16)	0.062
Maximum chloride (mmol/l)	107.94 (6.69)	108.28 (5.85)	0.036
Minimum hematocrit (%)	28.76 (5.98)	28.46 (6.21)	0.052
Maximum hematocrit (%)	35.14 (5.69)	35.93 (5.79)	< 0.001
Maximum lactate (mmol/l)	3.30 (2.85)	2.75 (1.90)	< 0.001
Minimum platelet ($\times 10^9/l$, median [IQR])	181.00 [120.25, 253.75]	178.00 [128.00, 240.00]	0.479
Maximum potassium (mmol/l)	4.80 (0.94)	4.78 (0.96)	0.444
Maximum aPTT (median [IQR])	34.20 [28.42, 47.80]	32.50 [27.50, 40.60]	< 0.001
Maximum INR	1.89 (1.82)	1.59 (1.02)	< 0.001
Minimum sodium (mmol/l)	136.42 (5.92)	136.23 (5.34)	0.210
Maximum sodium (mmol/l)	140.20 (5.57)	140.23 (4.42)	0.850
Maximum BUN (median [IQR])	28.00 [19.00, 45.00]	20.00 [14.00, 30.00]	< 0.001
Minimum WBC ($\times 10^9/l$)	11.67 (7.68)	10.61 (6.48)	< 0.001
Maximum WBC ($\times 10^9/l$)	15.77 (10.08)	14.62 (9.69)	< 0.001
Minimum albumin (g/dl)	2.83 (0.66)	3.03 (0.63)	< 0.001

Table 1 Characteristics between fluid responsive and non-responsive groups (*Continued*)

Variables	Volume unresponsive (n = 4226)	Volume responsive (n = 2456)	p value
Vital signs, mean (SD)			
Maximum heart rate (/min)	106.76 (22.48)	105.97 (19.98)	0.147
Minimum systolic BP (mmHg)	85.09 (18.03)	89.16 (16.46)	< 0.001
Minimum diastolic BP (mmHg)	40.38 (11.69)	43.60 (10.97)	< 0.001
Maximum respiratory rate (/min)	28.17 (6.75)	27.52 (6.57)	< 0.001
Maximum temperature (°C)	37.48 (0.88)	37.64 (0.76)	< 0.001
Urinary biomarkers, mean (SD)			
Urinary pH	5.66 (0.79)	5.78 (0.86)	< 0.001
Urinary creatinine (mg/dl)	132.51 (79.89)	111.61 (73.01)	< 0.001

Abbreviations: ICU intensive care unit, BP blood pressure, CCU coronary artery unit, CSRU cardiac surgery recovery unit, MICU medical ICU, SICU surgical ICU, TSICU trauma-neuro surgical ICU, SD standard deviation, IQR interquartile range, pH potential hydrogen, aPTT activated partial thromboplastin time, WBC white blood cell count

Methods

Source of data

A large US-based critical care database named Medical Information Mart for Intensive Care (MIMIC-III) was analyzed [12]. The MIMIC-III database is an integrated, de-identified, comprehensive clinical dataset containing all the patients admitted to the ICUs of Beth Israel Deaconess Medical Center in Boston, MA, from June 1, 2001, to October 31, 2012. There were 53,423 distinct hospital admissions for adult patients (aged 16 years or above) admitted to the ICUs during the study period. Since the study was an analysis of a third-party anonymized publicly available database with pre-existing institutional review board (IRB) approval, IRB approval from our institution was exempted. The study was reported according to the recommendations of the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement [13].

Participants

Patient eligibility was considered when urine output was less than 0.5 ml/kg/h for the first 6 h after ICU admission. This definition was consistent with the urine output component of the KDIGO criteria [4]. To examine the impact of fluid resuscitation on subsequent urine output response, only patients with substantial fluid intake (> 5 l) within 6 h following the initial 6 h for assessing oliguria were eligible (Fig. 1). Fluid intake and urine output were extracted from the nursing chart system, which more accurately reflected the actual volume intake or output than the medical order system. Patients who left ICU during the observation period were excluded. Furthermore, patients receiving any diuretics and/or renal replacement therapy (RRT) on day 1 were excluded.

Outcome (volume responsiveness)

The urine output within an 18-h period following the initial 6 h for defining oliguria was used as the outcome.

Patients were considered as VR-AKI if he/she had urine output greater than 0.65 ml/kg/h, corresponding to a 30% increase as compared with the baseline value. Otherwise, they were defined as VU-AKI.

Predictors of VR-AKI

Routinely collected clinical and laboratory variables obtained within the first 6 h of ICU admission were assessed for their ability to predict volume responsiveness. For some variables with multiple measurements, both the maximum and minimum values were assessed. Age, gender, ethnicity, admission type, elective surgery, type of ICU and presence of infection, and vital signs including respiratory rate, blood pressure, heart rate, and temperature were analyzed. In addition, laboratory data including glucose, white blood cell count (WBC), hematocrit, chloride, potassium, sodium, lactate, creatinine, blood urea nitrogen (BUN), coagulation profile, PaO₂, PaCO₂, and pH were included.

Because this was a hypothesis-generating epidemiological study, no attempt was made to estimate the sample size of the study; instead, all eligible patients in the database were included to maximize the statistical power of the predictive model. Because missing data may create bias, variables with > 70% missing values were excluded from further analysis. Other variables with a lesser degree of missing values were analyzed using multiple imputation method [14].

Statistical analysis

Clinical characteristics between VR-AKI and VU-AKI groups were compared using either Student *t* test or rank-sum test as appropriate. Chi-square test or Fisher's exact test was employed to compare the differences of the categorical variables [15]. A stepwise logistic regression model was used to select variables which were predictive of volume responsiveness. Both forward selection and backward elimination were used, testing at each step

for variables to be included or excluded. Akaike Information Criterion (AIC) was used as the selection criteria to eliminate the predictors [16].

Extreme gradient boosting (XGBoost) combined with decision trees was employed to predict VR versus VU. A classification tree was used as the weak learner, and the learning objective function was binary logistic. The boosting method works by iteratively refitting a weak classifier (decision tree) to residuals of previous models. Each successive classifier focused more on misclassified observations during the previous round of fitting [17]. In this study, we employed 300 rounds of iterations for cross-validation process, which were expected to result in a powerful ensemble classifier with superior predictive accuracy. Overfitting can be a major problem in using machine learning techniques. The ability to understand the complex relationship in data while avoiding overfitting requires fine-tuned hyperparameters. XGBoost hyperparameters included learning rate, minimum loss reduction required to make a further partition on a leaf node of the tree, maximum depth of a tree, subsample ratio of the training instance. The original dataset was randomly partitioned into 5 equal-sized subsamples for bootstrap validation (BV). Specifically, 4 subsamples were used to train the model, which was then validated in the remaining 1 subsample. Hyperparameters were considered to be sufficiently tuned if (1) the BV training log-loss decreased as the number of trees increased and (2) BV testing log-loss was less than 0.693 and only slightly greater than the training log-loss (e.g., a log-loss of 0.693 is the performance of a binary classifier that performs no better than chance: $-\log 0.5 \approx 0.693$). We used a loop function (grid search) to select the hyperparameters that the minimum training log-loss should be greater than the 85th percentile, and the minimum testing log-loss should be less than the 8th percentile. After choosing the hyperparameters, the BV process was run for 100 times to determine the number of trees required for the final model. The number of trees in the final XGBoost model was determined by the minimum BV testing log-loss in each of the last 100 BV iterations and computed the 5th percentile of that distribution. This was a conservative approach to the selection of the number of trees, which was also described in the literature [18].

Results

Participants

Of the 10,795 patients with urine output < 0.5 ml/kg/h for the first 6 h after ICU admission, 7491 patients (69.4%) received fluid intake > 5 l within the following 6 h. A number of 809 patients were excluded because they received diuretics and/or RRT on the first day. A total of 6682 patients were included in our analysis; 2456 patients had VR-AKI, and 4226 patients had VU-AKI on day 1 in ICU (Fig. 2).

Table 2 Multivariable logistic regression model with stepwise variable selection

Variables	OR (95% CI)	<i>p</i> value
Gender (female as reference)	1.56 [1.36, 1.78]	< 0.001
Ethnicity (Asian as reference)		
Black	0.85 [0.49, 1.50]	0.574
Hispanic	1.09 [0.57, 2.10]	0.785
Unknown	0.88 [0.51, 1.52]	0.632
White	0.74 [0.44, 1.26]	0.261
ICU type (CCU as reference)		
CSRU	1.39 [1.04, 1.85]	0.027
MICU	0.94 [0.72, 1.22]	0.630
SICU	0.88 [0.66, 1.18]	0.391
TSICU	0.82 [0.60, 1.11]	0.200
Infection		
Mechanical ventilation	0.71 [0.61, 0.81]	< 0.001
Bilirubin (with every unit increment)	0.97 [0.95, 0.98]	< 0.001
Lactate (with every unit increment)	0.91 [0.88, 0.94]	< 0.001
Albumin (with every unit increment)	1.53 [1.38, 1.70]	< 0.001
Temperature (with every unit increment)	1.19 [1.10, 1.30]	< 0.001
Urinary pH (with every unit increment)	1.05 [0.97, 1.13]	0.252
Age (with each 20 years increment)	0.69 [0.63, 0.75]	< 0.001
Serum creatinine (with every 0.1 mg/dl increment)	0.98 [0.97, 0.99]	< 0.001
Maximum chloride (with every 20 mmol/l increment)	2.94 [2.08, 4.16]	< 0.001
Minimum chloride (with every 20 mmol/l increment)	0.44 [0.31, 0.60]	< 0.001
Maximum glucose (with every 20 mmol/l increment)	0.99 [0.97, 1.00]	0.164
Minimum bicarbonate (with every 5 mmol/l increment)	1.12 [1.02, 1.22]	0.014
Minimum hematocrit (with every 5% increment)	0.91 [0.86, 0.97]	0.002
BUN (with every 10 mmol/l increment)	0.94 [0.90, 0.99]	0.017
Maximum heart rate (with every 10 beats/min increment)	1.03 [1.00, 1.06]	0.095
Minimum systolic BP (with every 20 mmHg increment)	1.13 [1.04, 1.23]	0.003
Urinary creatinine (with every 50 mg/dl increment)	0.71 [0.68, 0.75]	< 0.001
Maximum aPTT (for every 10 s increment)	0.96 [0.93, 0.98]	< 0.001

An OR value greater than 1 indicates that the presence of a variable or increase in a continuous variable is associated with higher probability of volume responsiveness

Abbreviations: OR odds ratio, BP blood pressure, aPTT activated partial thromboplastin time, BUN blood urea nitrogen, ICU intensive care unit, BP blood pressure, CCU coronary artery unit, CSRU cardiac surgery recovery unit, MICU medical ICU, SICU surgical ICU, TSICU trauma-neuro surgical ICU

The differences in characteristics between VR and VU groups are described in Table 1. VR group had more patients of elective surgery prior to ICU admission than the VU

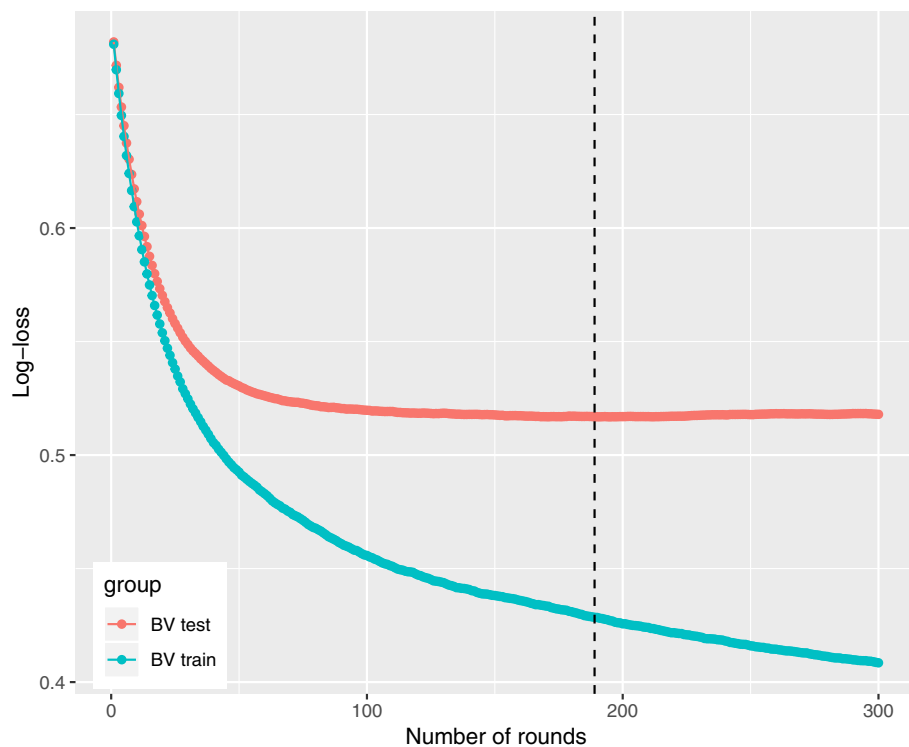


Fig. 3 Training process of the extreme gradient boosting machine. Sample output of bootstrap validation (BV) during XGBoost hyperparameter tuning, using the values specified in the final XGBoost model (learning rate = 0.04, minimum loss reduction = 10, maximum tree depth = 9, subsample = 0.6, and number of trees = 300). Log-loss value for the training and testing datasets is shown in the vertical axis. The dashed vertical line indicates the number of rounds with the minimum log-loss in the test sample. The conditions of well-tuned model were satisfied: BV training log-loss decreases as the number of trees in an ensemble increases, and BV testing log-loss is less than 0.693 (e.g., a log-loss of 0.693 is the performance of a binary classifier that performs no better than chance: $-\log 0.5 \approx 0.693$) and only slightly more than BV training log-loss as the tree grows.

group (18.2% vs. 13.6%; $p < 0.001$). The maximum serum creatinine concentration (1.74 ± 1.50 vs. 1.26 ± 0.98 $\mu\text{mol/l}$; $p < 0.001$) was higher, and the minimum bicarbonate concentration (20.82 ± 5.41 vs. 22.24 ± 4.32 mmol/l ; $p < 0.001$) and maximum hematocrit (35.14 ± 5.69 vs. $35.93 \pm 5.79\%$; $p < 0.001$) were lower in the VU group. The VR group had higher urinary pH (5.78 ± 0.86 vs. 5.66 ± 0.79 ; $p < 0.001$), lower urinary creatinine (111.61 ± 73.01 vs. 132.51 ± 79.89 mg/dl ; $p < 0.001$), higher systolic blood pressure (89.16 ± 16.46 vs. 85.09 ± 18.03 mmHg ; $p < 0.001$), higher albumin (3.03 ± 0.63 vs. 2.83 ± 0.66 ; $p < 0.001$), lower rate of mechanical ventilation (31.3% vs. 42.7%; $p < 0.001$), vasopressor use (22.8% vs. 30.0%; $p < 0.001$), and infection (47.6% vs. 62.4%; $p < 0.001$) than the VU group (Table 1).

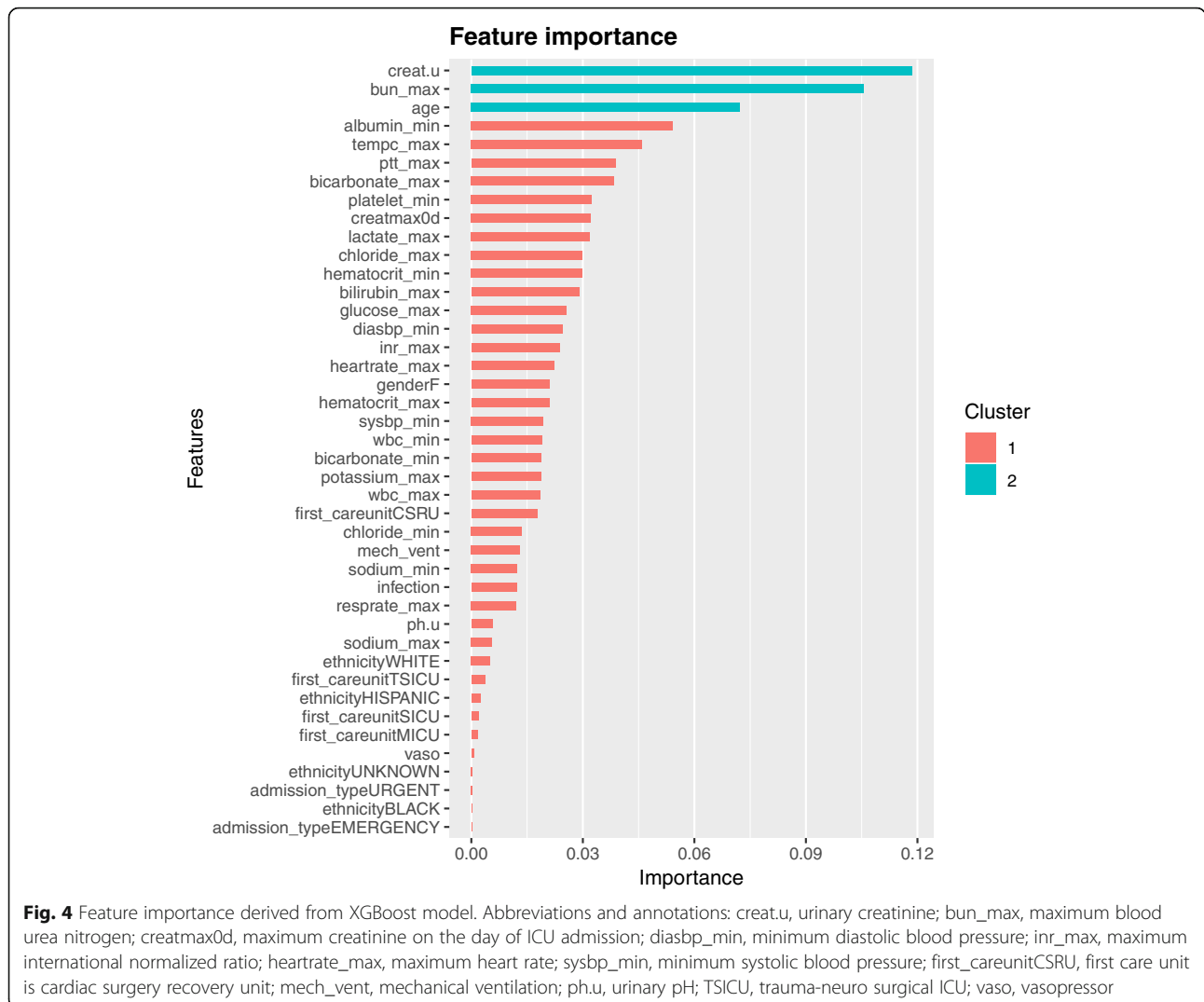
The stepwise logistic regression model

The results of stepwise logistic regression model are shown in Table 2. As expected, advanced age (odds ratio [OR] for a 20-year increase, 0.69; 95% confidence interval [CI], 0.63 to 0.75), a higher serum creatinine (OR for each 0.1 mg/dl increase, 0.98; 95% CI, 0.97 to 0.99), and lactate (OR for each 1- mmol/l increase, 0.91; 95% CI,

0.88 to 0.94) were associated with decreased probability of volume responsiveness. On the contrary, a greater value of albumin (OR for each 1-g/dl increase, 1.53; 95% CI, 1.38 to 1.71), systolic BP (OR for each 20- mmHg increase, 1.13; 95% CI, 1.04 to 1.23), and minimum bicarbonate (OR for each 5- mmol/l increase, 1.12; 95% CI, 1.02 to 1.22) were associated with increased probability of volume responsiveness (Table 2).

The XGBoost model

The hyperparameters used in our analysis were as follows (determined by grid search): learning rate = 0.04, minimum loss reduction = 10, maximum tree depth = 9, subsample = 0.6, and number of trees = 300. With these hyperparameters, bootstrap validation (BV) training log-loss decreases as the number of trees in an ensemble increases, and the BV testing log-loss was less than 0.693 and only slightly more than BV training log-loss as the tree grows (Fig. 3). Feature importance was calculated by the sum of the decrease in error when split by a variable, which reflects the contribution each variable makes in classifying VR versus VU. The urinary



creatinine was the most important variable to distinguish VR and VU group, followed by maximum BUN, age, albumin, and maximum temperature (Fig. 4).

Model performance

Model discrimination was assessed using the area under receiver operating characteristic curve (AU-ROC). The XGBoost had a significantly greater AU-ROC than the logistic regression model (AU-ROC, 0.860; 95% CI, 0.842 to 0.878 vs. 0.728; 95% CI, 0.703 to 0.753, respectively; Fig. 5). Table 3 describes the classification or confusion matrix for the two models in identifying the VR and VU status.

Discussion

In this hypothesis-generating study, we showed that some clinical factors were more likely to be associated with VR-AKI than VU-AKI. Using advanced machine learning techniques, we could identify some important

clinical factors associated with VR-AKI such as age, urinary creatinine concentration, maximum BUN concentration, and albumin. These results have some implications and require further consideration.

First, an ability to accurately identify volume responsiveness in critically ill patients with AKI is clinically important to avoid both hypervolemia and hypovolemia. Currently, there is a lack of a reliable tool to distinguish between VR and VU AKI at an early stage. In this study, we showed that sophisticated machine learning techniques such as the XGBoost modeling can enrich the amount of information we can obtain from analyzing a database and allow us to develop and validate a better-performing predictive model compared to the conventional logistic regression technique. The potential usefulness of the model is that it can help to stratify oliguria patients immediately after ICU admission. As a result, large volume fluid can be more accurately given to patients who are very likely to respond to fluid challenge. There is evidence that fluid

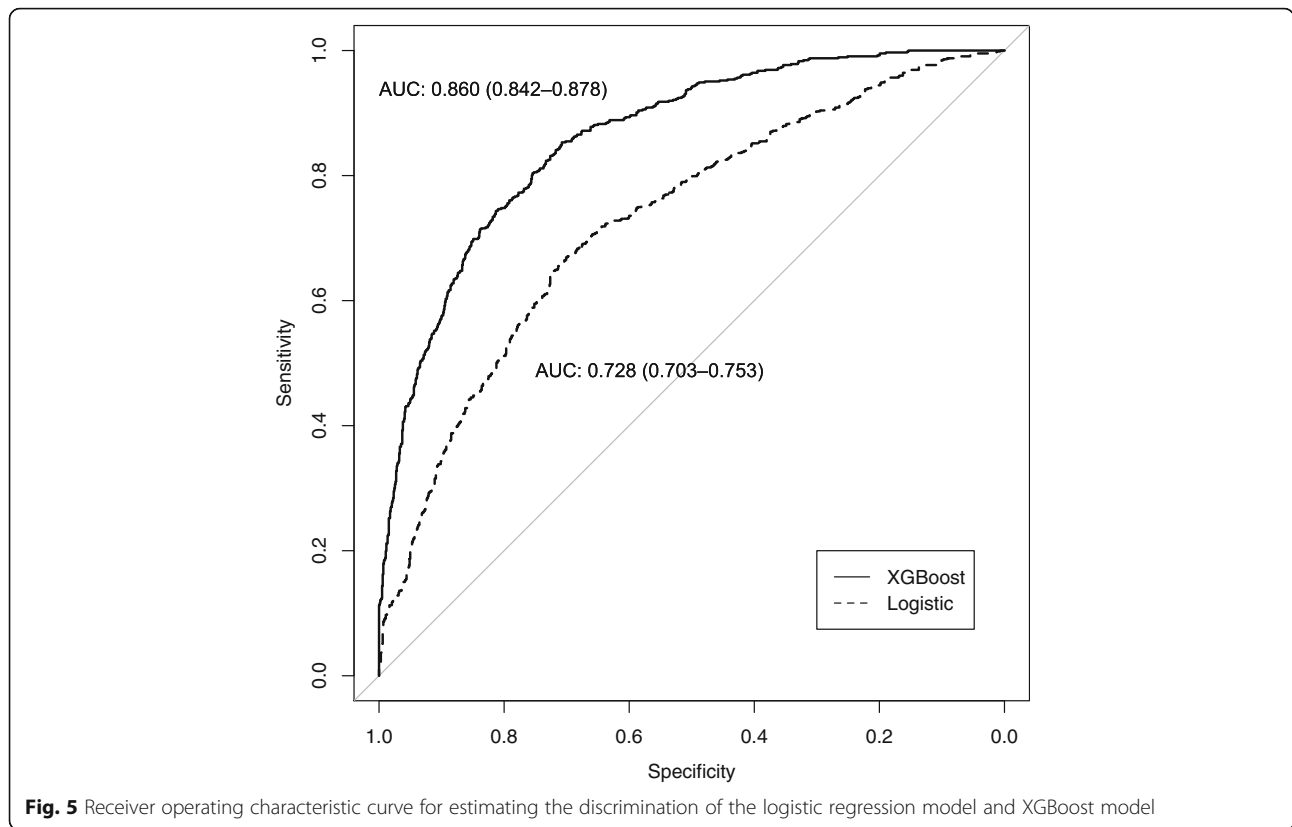


Fig. 5 Receiver operating characteristic curve for estimating the discrimination of the logistic regression model and XGBoost model

overload can result in organ dysfunctions, prolonged mechanical ventilation, and even death [19–21]. Thus, it is of vital importance to identify patients who will benefit from fluid resuscitation. However, the present study cannot provide a higher level of evidence on the effectiveness of the XGBoosting model. Future randomized controlled trials comparing the treatments with and without the prediction model are warranted to explore the effectiveness.

Second, our results showed that urinary creatinine was potentially useful to differentiate between patients in AKI who were VR and VU. Probably, patients with higher serum creatinine may also have higher excretion of creatinine to the urine. Since the former is a biomarker of kidney injury (e.g., higher serum creatinine was associated with higher risk of intrinsic injury), the latter is also associated with increased risk of VU-AKI. The utility of urinary biochemistry to predict AKI

outcome has been controversial in the literature. Although urinary biomarkers such as creatinine and fractional excretion of electrolytes were significantly different between VR and VU groups in some animal and human studies [10], there are also studies showing that urinary biochemistry may not be useful in differentiating between VR and VU AKI [22–25]. In our study, we could not analyze the ability of urinary sodium and potassium to differentiate between VR and VU AKI because a large proportion (> 70%) of patients did not have this data.

Third, we found that patients with AKI and oliguria after elective surgery were more likely to respond to fluid challenges in univariate analysis. Patients who underwent elective surgery are generally in better clinical condition than patients requiring urgent surgery or with a medical emergency. Postoperative oliguria can be

Table 3 Classification matrix for the XGBoost and logistic regression models in the out-of-sample validation cohort

	XGBoost Observed		Stepwise Logistic regression Observed	
	Non-responsive	Responsive	Non-responsive	Responsive
Predicted				
Non-responsive	846	180	737	228
Responsive	177	467	286	419

Correct classification (accuracy) of volume responsiveness for the XGBoost and the logistic models were 0.79 (95% CI, 0.77–0.81) and 0.69 (95% CI, 0.67–0.71), respectively

explained by hypovolemia due to intraoperative and postoperative insensible fluid loss. As such, they will be more likely to benefit from a larger amount of fluid after major surgery. However, the association of elective surgery with volume responsiveness disappeared after adjusting for other physiological variables, indicating that hypovolemia can be represented by these variables such as systolic blood pressure, heart rate, and hematocrit. We found that an increased hematocrit within the first 24 h of ICU admission was also an independent predictor of VR (in both the logistic and XGBoost models). This result could be explained by the fact that hematocrit has a direct relationship with the intravascular plasma volume and a higher hematocrit may suggest a relative hypovolemic state [26]. Conversely, a higher serum creatinine on ICU admission might indicate established renal intrinsic damage, which is more likely to be unresponsive to fluid challenges.

This study has some strengths and weaknesses. The XGBoost modeling is a novel technique that has not been widely adopted in critical care research. The XGBoost algorithm has been successfully used in some complex scenarios such as the prediction of the failure of the treatment for parapneumonic empyema [18], in which the predictive accuracy of the XGBoost model was significantly better than a generalized linear model. This is not surprising because the XGBoost model is an ensemble of weak prediction trees, which is able to capture complex relationships in data without the need for high-order interactions and non-linear functions to be explicitly specified [27]. Furthermore, this technique is well designed to prevent overfitting by cross-validation and regularization [17]. Our results suggest that this technique has the potential to improve the power of critical care epidemiological studies in the future. Nonetheless, this was a hypothesis-generating study, and external validation of our model is essential to confirm its utility. A limitation of this study is that we did not have data on the indications for large volume resuscitation. The study was not a designed clinical trial that the indications for large volume loading could be prespecified. However, we have randomly selected 30 cases and found that patients receiving > 5 l fluid during a 6-h period were those with indications for fluid loading in order to increase the urine output. Furthermore, the overall population had low blood pressure (mean SBP, 85 mmHg), elevated heart rate (mean, > 100/min), and lactate (mean, 3 mmol/l) in our study, which were consistent with the indications for fluid loading (Table 1). Clinicians might have given their patients fluid for a variety of different reasons such as hypotension or elevated heart rate in addition to any attempt to improve oliguria. The study only explored the short-term effect of large volume intake, other long-term outcomes such as persistent AKI,

organ-failure free days, and mortality were not investigated.

Conclusion

In conclusion, this hypothesis-generating study showed that some clinical factors were more likely to be associated with VR-AKI than VU-AKI. The XGBoost modeling technique could identify the predictors of VR-AKI that were not apparent using logistic regression, resulting in a better-performing predictive model to identify patients with VR-AKI. Further epidemiological studies using advanced machine learning techniques to validate our results will help us to identify the most suitable patients to be included in clinical trials assessing the benefits of fluid therapy in AKI.

Abbreviations

AKI: Acute kidney injury; BUN: Blood urea nitrogen; KDIGO: Kidney Disease: Improving Global Outcomes; MIMIC: Medical Information Mart for Intensive Care; TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis; VR: Volume responsive; VU: Volume unresponsive; WBC: White blood cell count; XGBoost: Extreme gradient boosting

Acknowledgements

None

Funding

There is no funding.

Availability of data and materials

The data were available on the MIMIC-III website at <https://mimic.physionet.org/>

Authors' contributions

ZZ conceived the idea, performed the analysis, and drafted the manuscript. KH interpreted the results and helped to revise the manuscript. YH helped to frame the idea of the study and helped to analyze the data. All authors read and approved the final manuscript.

Ethics approval and consent to participate

The study was an analysis of a third-party anonymized publicly available database with pre-existing institutional review board (IRB) approval.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Emergency Medicine, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, No. 3, East Qingchun Road, Hangzhou 310016, Zhejiang Province, China. ²School of Population and Global Health, University of Western Australia, Perth, Australia.

Received: 31 January 2019 Accepted: 26 March 2019

Published online: 08 April 2019

References

1. Saxena A, Meshram SV. Predictors of mortality in acute kidney injury patients admitted to medicine intensive care unit in a rural tertiary care hospital. *Indian J Crit Care Med.* 2018;22:231–7.

2. Anseeuw K, Mowry JB, Burdman EA, Ghannoum M, Hoffman RS, Gosselin S, et al. Extracorporeal treatment in phenytoin poisoning: systematic review and recommendations from the EXTRIP (Extracorporeal Treatments in Poisoning) Workgroup. *Am J Kidney Dis.* 2016;67:187–97.
3. Ostermann M, Chang RWS. Acute kidney injury in the intensive care unit according to RIFLE. *Crit Care Med.* 2007;35:1837–43–quiz1852.
4. Kidney Disease Improving Global Outcomes KDIGO Acute Kidney Injury Work Group. KDIGO clinical practice guideline for acute kidney injury. *Kidney Int.* 2012;suppl 2:1–138.
5. Miller TE, Bunke M, Nisbet P, Brudney CS. Fluid resuscitation practice patterns in intensive care units of the USA: a cross-sectional survey of critical care physicians. *Perioper Med (Lond) BioMed Central.* 2016;5:15.
6. Cecconi M, Hofer C, Teboul J-L, Pettilä V, Wilkman E, Molnar Z, et al. Fluid challenges in intensive care: the FENICE study: a global inception cohort study. *Intensive Care Med.* 2015;41:1529–37.
7. Raimundo M, Crichton S, Martin JR, Syed Y, Varrier M, Wyncoll D, et al. Increased fluid administration after early acute kidney injury is associated with less renal recovery. *Shock.* 2015;44:431–7.
8. Himmelfarb J, Joannidis M, Molitoris B, Schietz M, Okusa MD, Warnock D, et al. Evaluation and initial management of acute kidney injury. *Clin J Am Soc Nephrol.* 2008;3:962–7.
9. Levy MM, Macias WL, Vincent J-L, Russell JA, Silva E, Trzaskoma B, et al. Early changes in organ function predict eventual survival in severe sepsis. *Crit Care Med.* 2005;33:2194–202.
10. Troia R, Guarini M, Grisetti C, Serafini F, Magna L, Monari E, et al. Fractional excretion of electrolytes in volume-responsive and intrinsic acute kidney injury in dogs: diagnostic and prognostic implications. *J Vet Intern Med.* 2018;32:1372–82.
11. Legrand M, Le Cam B, Perbet S, Roger C, Darmon M, Guerci P, et al. Urine sodium concentration to predict fluid responsiveness in oliguric ICU patients: a prospective multicenter observational study. *Crit Care.* 2016;20:165.
12. Johnson AEW, Pollard TJ, Shen L, Lehman L-WH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. *Sci Data.* 2016;3:160035.
13. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMC Med.* 2015;13:1.
14. Zhang Z. Multiple imputation with multivariate imputation by chained equation (MICE) package. *Ann Transl Med.* 2016;4:30.
15. Zhang Z, Gayle AA, Wang J, Zhang H, Cardinal-Fernández P. Comparing baseline characteristics between groups: an introduction to the CBCgrps package. *Ann Transl Med.* 2017;5:484.
16. Zhang Z. Variable selection with stepwise and best subset approaches. *Ann Transl Med.* 2016;4:136.
17. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. *arXiv.* New York: ACM Press; 2016. p. 785–94.
18. Khemasuwan D, Sorensen J, Griffin DC. Predictive variables for failure in administration of intrapleural tissue plasminogen activator/deoxyribonuclease in patients with complicated parapneumonic effusions/empyema. *Chest.* 2018;154:550–6.
19. Chao W-C, Tseng C-H, Chien Y-C, Sheu C-C, Tsai M-J, Fang W-F, et al. Association of day 4 cumulative fluid balance with mortality in critically ill patients with influenza: a multicenter retrospective cohort study in Taiwan. *Lazzeri C, editor. PLoS One* 2018;13:e0190952.
20. Silversides JA, Major E, Ferguson AJ, Mann EE, McAuley DF, Marshall JC, et al. Conservative fluid management or deresuscitation for patients with sepsis or acute respiratory distress syndrome following the resuscitation phase of critical illness: a systematic review and meta-analysis. *Intensive Care Med.* 2017;43:155–70.
21. National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network, Wiedemann HP, Wheeler AP, Bernard GR, Thompson BT, Hayden D, et al. Comparison of two fluid-management strategies in acute lung injury. *N Engl J Med.* 2006;354:2564–75.
22. Vanmassenhove J, Glorieux G, Hoste E, Dhondt A, Vanholder R, Van Biesen W. Urinary output and fractional excretion of sodium and urea as indicators of transient versus intrinsic acute kidney injury during early sepsis. *Crit Care.* 2013;17:R234.
23. Darmon M, Vincent F, Dellamonica J, Schortgen F, Gonzalez F, Das V, et al. Diagnostic performance of fractional excretion of urea in the evaluation of critically ill patients with acute kidney injury: a multicenter cohort study. *Crit Care.* 2011;15:R178.
24. Wlodzimirow KA, Abu-Hanna A, Royackers AANM, Spronk PE, Hofstra LS, Kuiper MA, et al. Transient versus persistent acute kidney injury and the diagnostic performance of fractional excretion of urea in critically ill patients. *Nephron Clin Pract.* 2014;126:8–13.
25. Dewitte A, Biais M, Petit L, Cochard JF, Hilbert G, Combe C, et al. Fractional excretion of urea as a diagnostic index in acute kidney injury in intensive care patients. *J Crit Care.* 2012;27:505–10.
26. Schiller AM, Howard JT, Lye KR, Magby CG, Convertino VA. Comparisons of traditional metabolic markers and compensatory reserve as early predictors of tolerance to central hypovolemia in humans. *Shock.* 2018;50:71–7.
27. Friedman JH. Stochastic gradient boosting. *Comput Stat Data Anal.* 2002;38:367–78.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

