

# Development and validation of a deep learning model to screen hypokalemia from electrocardiogram in emergency patients

Chen-Xi Wang<sup>1</sup>, Yi-Chu Zhang<sup>1</sup>, Qi-Lin Kong<sup>1</sup>, Zu-Xiang Wu<sup>1</sup>, Ping-Ping Yang<sup>1</sup>, Cai-Hua Zhu<sup>1</sup>, Shou-Lin Chen<sup>2</sup>, Tao Wu<sup>1</sup>, Qing-Hua Wu<sup>1</sup>, Qi Chen<sup>1</sup>

<sup>1</sup>Department of Cardiovascular Medicine, The Second Affiliated Hospital of Nanchang University, Nanchang University School of Medicine, Nanchang, Jiangxi 330006, China;

<sup>2</sup>Department of Anesthesiology, The Second Affiliated Hospital of Nanchang University, Nanchang University School of Medicine, Nanchang, Jiangxi 330006, China.

## Abstract

**Background:** A deep learning model (DLM) that enables non-invasive hypokalemia screening from an electrocardiogram (ECG) may improve the detection of this life-threatening condition. This study aimed to develop and evaluate the performance of a DLM for the detection of hypokalemia from the ECGs of emergency patients.

**Methods:** We used a total of 9908 ECG data from emergency patients who were admitted at the Second Affiliated Hospital of Nanchang University, Jiangxi, China, from September 2017 to October 2020. The DLM was trained using 12 ECG leads (lead I, II, III, aVR, aVL, aVF, and V<sub>1-6</sub>) to detect patients with serum potassium concentrations <3.5 mmol/L and was validated using retrospective data from the Jiangling branch of the Second Affiliated Hospital of Nanchang University. The blood draw was completed within 10 min before and after the ECG examination, and there was no new or ongoing infusion during this period.

**Results:** We used 6904 ECGs and 1726 ECGs as development and internal validation data sets, respectively. In addition, 1278 ECGs from the Jiangling branch of the Second Affiliated Hospital of Nanchang University were used as external validation data sets. Using 12 ECG leads (leads I, II, III, aVR, aVL, aVF, and V<sub>1-6</sub>), the area under the receiver operating characteristic curve (AUC) of the DLM was 0.80 (95% confidence interval [CI]: 0.77–0.82) for the internal validation data set. Using an optimal operating point yielded a sensitivity of 71.4% and a specificity of 77.1%. Using the same 12 ECG leads, the external validation data set resulted in an AUC for the DLM of 0.77 (95% CI: 0.75–0.79). Using an optimal operating point yielded a sensitivity of 70.0% and a specificity of 69.1%.

**Conclusions:** In this study, using 12 ECG leads, a DLM detected hypokalemia in emergency patients with an AUC of 0.77 to 0.80. Artificial intelligence could be used to analyze an ECG to quickly screen for hypokalemia.

**Keywords:** Deep learning; Hypokalemia; Electrocardiogram; Artificial intelligence

## Introduction

Hypokalemia is one of the most common electrolyte disturbances encountered in clinical practice.<sup>[1]</sup> Detection of the serum potassium concentration is the main diagnostic method for hypokalemia, but a long detection time and poor repeatability may delay clinical intervention and allow a patient's condition to deteriorate.<sup>[2,3]</sup> This situation is undesirable for emergency patients who require a rapid diagnosis.

Hypokalemia can increase the excitability and self-regulation of cardiomyocytes and slow conductivity, which manifests as a series of well-defined ECG abnormalities, such as T-wave changes, ST-segment decline, QT-interval prolongation, and U wave values  $\geq 0.1$  mV.<sup>[2,4]</sup> However, physicians in clinical practice are not particularly attentive to changes in ECGs when diagnosing

electrolyte disturbances.<sup>[5]</sup> Traditional artificial intelligence applications have gradually evolved into those for specialized medicine.<sup>[6,7]</sup> We hypothesize that a deep learning model (DLM) based on convolutional neural networks (CNNs) can be used to effectively screen emergency patients for hypokalemia. Therefore, we trained and validated a DLM to screen for hypokalemia based on the ECGs of emergency patients.

The objective of this study was to improve the detection efficiency of hypokalemia in emergency patients by using an electrocardiogram (ECG) to develop and verify non-invasive screening tests.

## Methods

This study was approved by the Ethics Review Committee of the Second Affiliated Hospital of Nanchang University (No.

### Access this article online

Quick Response Code:



Website:  
www.cmj.org

DOI:  
10.1097/CM9.0000000000001650

**Correspondence to:** Qi Chen, Department of Cardiovascular Medicine, The Second Affiliated Hospital of Nanchang University, Nanchang University School of Medicine, No. 1, Minde Road, Nanchang, Jiangxi 330006, China  
E-Mail: efyccq@189.cn

Copyright © 2021 The Chinese Medical Association, produced by Wolters Kluwer, Inc. under the CC-BY-NC-ND license. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Chinese Medical Journal 2021;134(19)

Received: 27-04-2021 Edited by: Peng Lyu

2019-086). The study was conducted in accordance with the *Declaration of Helsinki* (as revised in 2013). Clinical data, including ECGs stored as electronic data, serum potassium and magnesium ion concentrations, B-type natriuretic peptide (BNP) levels, free thyroxine levels, sex, and age, were obtained from the Second Affiliated Hospital of Nanchang University. As the acquired data were anonymously processed by the hospital scientific research platform and the study was retrospective, the Ethics Review Committee exempted the patients' informed consent.

**Deep learning model**

Deep learning is a new research direction in machine learning that is moving the field toward its original goal, namely, artificial intelligence (AI). Deep learning learns the internal laws and representation levels of sample data and uses many hidden neuron layers to generate increasingly abstracted, non-linear representations of the underlying data. The goal of deep learning is to endow machines with the analytical and learning abilities of humans and to recognize data such as images and sounds. Image recognition was the first application of deep learning to the clinical field.<sup>[7]</sup>

In this study, a DLM was built on an Anaconda platform using Python (version 3.5.2; Python Software Foundation, Beaverton, OR, USA) and the TensorFlow neural network framework (Google LLC, Mountain View, CA, USA). The framework had 11 CNN layers, of which the first 10 were convolutional layers, and the last layer was a fully connected SoftMax layer. The network output result was between 0 and 1, indicating the probability of detecting hypokalemia from an ECG [Figure 1]. We trained the DLM on 12 ECG leads (leads I, II, III, aVR, aVL, aVF, and V1–6). We also trained the DLM using a single lead (lead II) that can be easily detected by wearable devices.<sup>[8,9]</sup> The electrocardiograph used to record the ECG data comes from Nippon Optoelectronics Tomioka Co., Ltd.(model ECG-1150).

**Development and validation data sets**

A total of 310,256 ECGs were obtained from September 2017 to October 2020 from the emergency department of the Second Affiliated Hospital of Nanchang University, including 4615 ECGs of patients with hypokalemia. By

excluding patients who have taken >10 min between draw blood from vein and ECG examination, or who have had potassium supplementation treatment or any other medical orders during this period, we ensured that the ECG data could reflect the truest serum potassium level during the ECG examination as much as possible. In addition, the deletion of ECGs without associated demographic data and indication of death yielded a total of 4315 hypokalemia ECGs, as confirmed by serum potassium test results. Considering the limited computing power of existing machines, non-hypokalemic ECGs were randomly selected and aggregated with the same number of selected hypokalemic ECGs for the same period to serve as a DLM development data set in this study. The final development data set included 8630 ECGs.

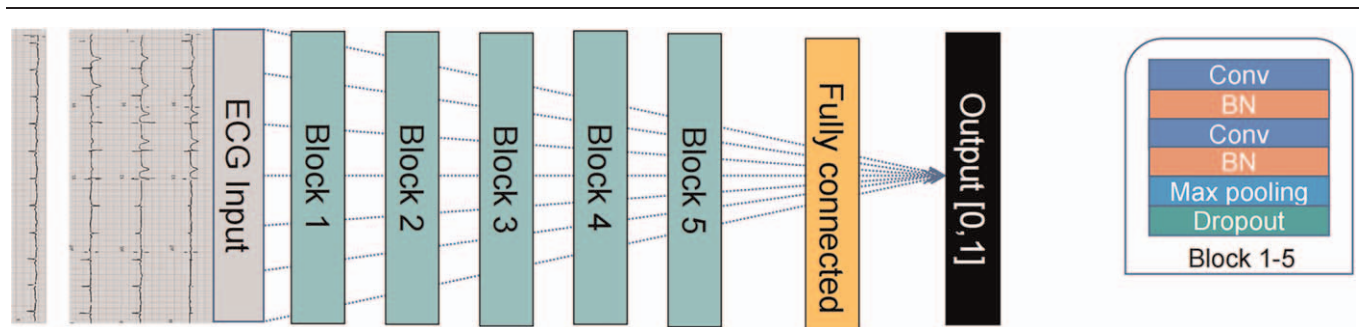
All patients underwent at least one standard 10-s 12-lead ECG in the resting supine position at the time of emergency treatment. Each digitally stored ECG lead was recorded at 500 data points per second (500 Hz) for 10 s. We compiled a data set consisting of the data from all 12 ECG leads. As shown in Figure 2, the ECGs were randomly divided into a training data set (80%) and an internal validation data set (20%). A total of 1278 ECG data from 1278 patients from the Jiangling branch of the Second Affiliated Hospital of Nanchang University were used only for external verification, showing the powerful capability of the DLM for various data sets. The Jiangling branch (Hospital B) is located in the suburbs and has a distinctly different environment from the hospital headquarters (Hospital A). After review, Hospital A and Hospital B do not have duplicate emergency visit records.

**Hypokalemia and non-hypokalemia**

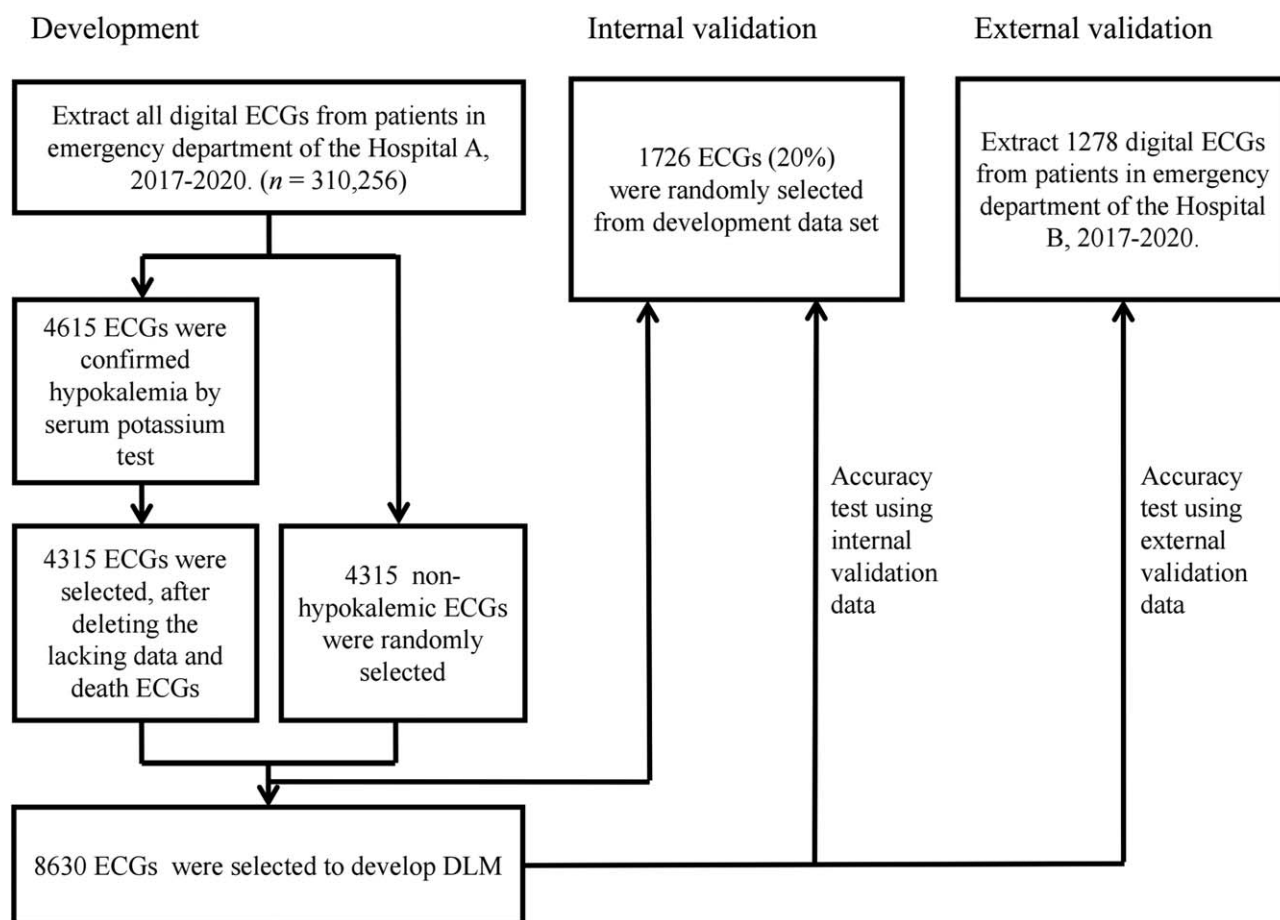
A normal serum potassium concentration is 3.5 to 5.5 mmol/L, with an average of 4.2 mmol/L. Hypokalemia usually occurs at <3.5 mmol/L serum potassium.<sup>[10]</sup> In this study, hypokalemia and non-hypokalemia were defined as corresponding to serum potassium concentrations <3.5 and ≥3.5 mmol/L, respectively.

**Statistical analysis**

We used the size of the area under the receiver operating characteristic curve (AUC) to evaluate the DLM perfor-



**Figure 1:** DLM for predicting hypokalemia. The DLM consisted of a convolutional CNN with 11 layers, with the first ten layers being convolutional and the last fully connected SoftMax layer. A RELU activation function was used. Skip connections, dropout, BN, and max pooling were all utilized to improve generalization and convergence properties. The DNN was designed as a binary classifier that output a number from 0 to 1, representing the probability that hypokalemia was detected from the ECG. BN: Batch normalization; Conv: Convolutional layer; CNN: Convolutional neural networks; DLM: Deep learning model; ECG: Electrocardiography; RELU: Rectified linear units.



**Figure 2:** Study flowchart. Hospital A: The Second Affiliated Hospital of Nanchang University. Hospital B: Jiangling branch of the Second Affiliated Hospital of Nanchang University. DLM: Deep learning model; ECG: Electrocardiogram.

mance. As the DLM was developed to rapidly screen patients with potential hypokalemia, we evaluated the specificity, positive predictive value, and negative predictive value at a cutoff point selected for high sensitivity in the development data.<sup>[11]</sup> Except for the AUC, all the diagnostic performance indicators were based on an accurate 95% confidence interval (CI). The reliability interval of the AUC was determined by using the pROC software package in R (R Foundation) to perform the Sun and Su optimization of the Delong method. A bilateral  $P < 0.05$  indicates statistical significance.<sup>[12]</sup> R software, version 4.0 (R Foundation) was used to perform the analyses.

**Results**

The incidence of hypokalemia in the emergency department patients was approximately 1.49% in this study. A total of 9908 ECGs from patients who admitted at the Jiangling branch of the Second Affiliated Hospital of Nanchang University were included in this study, of which the training data set consisted of 6904 ECGs from 5897 patients, the internal validation data set consisted of 1726 ECGs from 986 patients, and the external validation data set consisted of 1278 ECGs from 1278 patients. Table 1 shows the baseline characteristics of

the study population. A total of 8251 patients with an average age of 64.3 years were included in the study. The average blood potassium concentration of patients with hypokalemia was 2.89 mmol/L, and the average blood potassium drawing time after performing the ECG was 51.3 min. Hypokalemia ECGs were more likely to be recorded in patients with hypomagnesemia or NT pro-BNP >300 pg/mL.

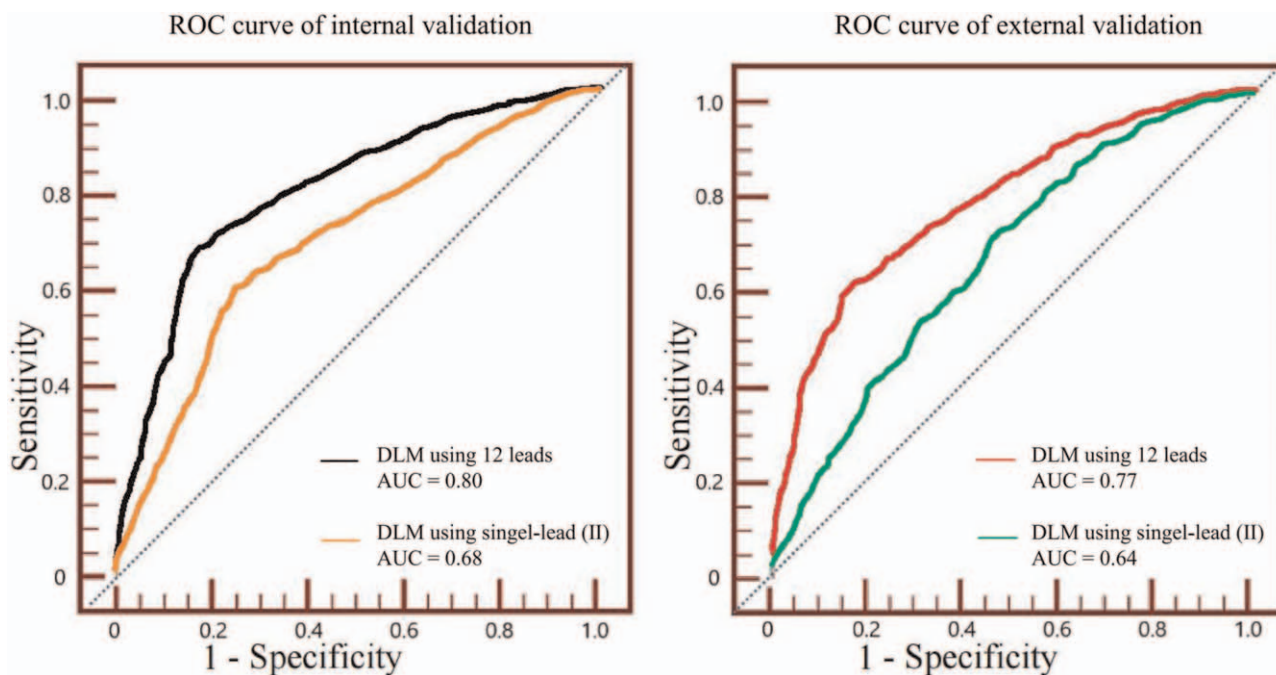
The DLM performed well in identifying hypokalemia for the internal and external validation data sets [Figure 3]. Using 12 ECG leads to detect hypokalemia resulted in AUCs of the DLM of 0.80 (95% CI: 0.77–0.82) and 0.77 (95% CI: 0.75–0.79) for the internal and external validations, respectively. However, the AUC was 0.12 to 0.13 lower using a single ECG lead II. Using an optimal operating point with the 12 ECG leads resulted in a sensitivity of 71.43% (95% CI: 69.35–73.13) and a specificity of 77.15% (95% CI: 75.36–80.16) for the internal validation data set and a sensitivity of 70.01% (95% CI: 67.63–73.34) and a specificity of 69.14% (95% CI: 67.25–72.86) for the external validation data set [Table 2]. When the blood potassium concentration is lower than 2.6 mmol/L, the recognition accuracy rate of DLM is 0.72. When the blood potassium concentration is between 2.6 and 3.5 mmol/L, the recognition accuracy rate

**Table 1: Baseline characteristics of development and validation data sets.**

Characteristics	ECGs			Statistics	P value
	All (n = 6973)	Hypokalemia (n = 3082)	Non-hypokalemia (n = 3891)		
Development data set					
Men	3905 (56.0)	1898 (61.6)	2254 (57.9)	4.849*	0.03
Age (years)	63.8 ± 14.3	66.2 ± 14.2	62.1 ± 13.3	-13.323 <sup>†</sup>	<0.01
Serum potassium level (mmol/L)	3.92 (2.51, 4.82)	2.75 (2.02, 3.11)	4.16 (3.73, 5.10)	27.132*	<0.01
NT pro-BNP >300 pg/mL	1220 (17.5)	804 (26.1)	416 (10.7)	15.672*	<0.01
FT <sub>4</sub> >1.80 or FT <sub>3</sub> >4.20 pg/mL	413 (5.9)	175 (5.7)	238 (6.1)	9.976*	<0.01
Serum magnesium level <0.75 mmol/L	125 (1.8)	78 (2.5)	47 (1.2)	7.852*	<0.01
Serum potassium level <2.60 mmol/L	960 (24.6)	960 (50.6)	0	-	-
Serum potassium test report time after ECG (h)	0.89 (0.52, 1.21)	0.84 (0.51, 1.12)	0.90 (0.53, 1.19)	1.282 <sup>‡</sup>	0.20

Characteristics	ECGs			Statistics	P value
	All (n = 1278)	Hypokalemia (n = 639)	Non-hypokalemia (n = 639)		
External validation data set					
Men	703 (55.0)	387 (60.6)	357 (55.9)	4.782*	0.03
Age (years)	64.8 ± 14.6	67.2 ± 14.8	63.1 ± 12.4	-10.367 <sup>†</sup>	<0.01
Serum potassium level (mmol/L)	3.94 (2.61, 4.84)	2.78 (2.12, 3.17)	4.24 (3.71, 5.08)	30.932 <sup>‡</sup>	<0.01
NT pro-BNP >300 pg/mL	192 (15.0)	108 (16.9)	84 (13.1)	10.365*	<0.01
FT <sub>4</sub> >1.80 or FT <sub>3</sub> >4.20 pg/mL	48 (3.8)	19 (3.0)	29 (4.5)	8.752*	<0.01
Serum magnesium level <0.75 mmol/L	15 (1.2)	9 (1.4)	6 (0.9)	3.845*	0.05
Serum potassium level <2.60 mmol/L	315 (24.6)	315 (49.3)	0	-	-
Serum potassium test report time after ECG (h)	0.82 (0.48, 1.31)	0.79 (0.47, 1.24)	0.83 (0.52, 1.17)	1.484 <sup>‡</sup>	0.14

Data are expressed as n (%), median (inter-percentile ranges), or mean ± standard deviation. \*Chi-square test; <sup>†</sup>t-test; <sup>‡</sup>Mann-Whitney U test. ECG: Electrocardiogram; FT: Free thyroxine; NT pro-BNP: N-terminal pro-B-type natriuretic peptide.



**Figure 3:** Performances of DLM for screening hypokalemia. AUC: Area under the receiver operating characteristic curve; DLM: Deep learning model; ROC: Receiver operating characteristic curve.

of DLM is 0.76. Therefore, DLM still has a good accuracy rate for ECGs under different blood potassium concentrations [Supplementary Figure 1, <http://links.lww.com/CM9/A779>].

To further evaluate the DLM performance under hypothetical confounding factors, we constructed a verification data set comprising 176 randomly selected samples with atrial fibrillation (AF), complete left bundle branch block



**Table 2: Validation data set performance for hypokalemia from single-lead (II).**

Validation data set	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Internal validation, using 12 leads ECG	0.80 (0.77–0.82)	71.43 (69.35–73.13)	77.15 (75.36–80.16)	75.12 (73.54–76.94)	73.56 (71.87–76.12)
Internal validation, using single-lead ECG (lead II)	0.68 (0.66–0.71)	70.03 (67.82–72.31)	51.95 (48.37–53.23)	68.22 (66.13–71.74)	58.73 (56.21–60.33)
External validation, using 12 leads ECG	0.77 (0.75–0.79)	70.01 (67.63–73.34)	69.14 (67.25–72.86)	79.63 (77.55–82.12)	57.33 (55.23–59.91)
External validation, using single-lead ECG (lead II)	0.64 (0.62–0.67)	70.91 (68.23–74.15)	51.02 (49.13–53.02)	66.23 (64.13–68.45)	56.33 (54.34–59.24)

AUC: Area under the receiver operating characteristic curve; CI: Confidence interval; ECG: Electrocardiogram; NPV: Negative predictive value; PPV: Positive predictive value

**Table 3: Results of DLM in identifying hypokalemia in ECG with confounding factors from 12 ECG leads.**

Independent variables	Number of ECG	Results of identifying		
		Hypokalemia, <i>n</i> (%)	Non-hypokalemia, <i>n</i> (%)	Accuracy, %
AF	Hypokalemia ( <i>n</i> = 36)	26 (72.2)	10 (27.8)	72.1
	Non-hypokalemia ( <i>n</i> = 75)	21 (28.0)	54 (72.0)	
CRBBB	Hypokalemia ( <i>n</i> = 20)	11 (55.0)	9 (45.0)	55.4
	Non-hypokalemia ( <i>n</i> = 36)	16 (44.0)	20 (56.0)	
CLBBB	Hypokalemia ( <i>n</i> = 2)	0 (0)	2 (100)	16.7
	Non-hypokalemia ( <i>n</i> = 4)	3 (75.0)	1 (25.0)	
Pacing	Hypokalemia ( <i>n</i> = 2)	2 (100)	0 (0)	100.0
	Non-hypokalemia ( <i>n</i> = 1)	0 (0)	1 (100)	

AF: Atrial fibrillation; CRBBB: Complete right bundle branch block; CLBBB: Complete left bundle branch block; DLM: Deep learning model; ECG: Electrocardiogram.

(CLBBB), complete right bundle branch block, and pacing ECG. In this data set, the overall recognition accuracy rate of DLM is 61.1%. The verification results are shown in Table 3. The sensitivity and specificity of this model in identifying hypokalemia from AF ECGs were 74.2% and 72.0%, respectively, with an accuracy of 72.1%. The model performed best for pacing ECGs and worst for CLBBB ECGs.

**Discussion**

Over the past 10 years, various DLMs have been increasingly applied in research on cardiovascular disease, such as for the prediction of left ventricular systolic function, AF, and cardiac arrest,<sup>[11,13,14]</sup> especially during the Coronavirus disease 2019 (COVID-19) epidemic. Clinicians have found AI extremely useful for identifying patients with COVID-19 and predicting the severity and progress of the disease.<sup>[15,16]</sup> The aforementioned studies have shown that a CNN-based DLM can confer strong recognition or prediction ability to a machine.

A DLM for screening hypokalemia in emergency patients using 12 ECG leads was developed and validated in this study. Using the 12 ECG leads resulted in an AUC of the DLM of 0.80 for the internal validation data set and 0.77 for the external validation data set (not used for DLM development), which indicates good and stable model performance for hypokalemia screening. The model

outperformed other common screening tests, such as fecal occult blood testing for detecting colorectal neoplasia (AUC 0.71; overall sensitivity, 29%).<sup>[17]</sup> However, lower model performance was obtained using a single-lead ECG (lead II). This result may have been obtained because of the relatively few data used. Extending the ECG monitoring time could gradually increase the quantity of acquired data and improve the detection performance. Unlike previous studies in which CNNs have been used to construct DLMs to screen serum ion concentrations,<sup>[7]</sup> complete 12-lead ECG data were used in this study to develop the DLM. Thus, the DLM detection performance is lower than that reported in previous studies but may be more reliable.

Although the serum potassium concentration can be obtained relatively rapidly by venous blood measurement in a hospital, hypokalemia diagnosis outside the hospital (such as in community clinics) remains challenging because patients with hypokalemia usually do not exhibit characteristic symptoms. Using ECGs to non-invasively screen patients for hypokalemia can be a powerful facilitator for early detection of this disease and potentially improve care and outcomes. Moreover, many wearable devices for monitoring ECGs have been developed over the past few years.<sup>[18]</sup> Therefore, the serum potassium concentration can be dynamically detected at home, which is highly beneficial for patients prone to hypokalemia. However, whether a similar DLM performance would be obtained using wearable ECG inputs remains to be determined.

The most important ability of a CNN is the extraction of features from various types of data, such as images, two-dimensional data, and waveforms, as well as algorithm generation. Traditional methods use a standard regression model to estimate the potassium content, where the *T*-wave width, *T*-wave amplitude, *T*-wave slope, and *U*-wave value are considered to be important indexes of changes in blood potassium levels.<sup>[19,20]</sup> However, CNN is peculiar in precluding the inference of which feature information is extracted by the DLM.<sup>[21]</sup> We only know that the DLM can screen for hypokalemia based on characteristic changes in ECGs that humans have not yet discerned. Although some researchers have used visualization technology to determine the image area where DLM is used for decision-making, this area still cannot be quantified to enable humans to make the same judgment.<sup>[11]</sup> Therefore, a visual analysis of the DLM was not performed in this study.

Overfitting models often only have a good recognition rate for specific data sets, but our DLM has a high recognition rate regardless of whether it is in the internal data validation set or the external validation data set; even in a data set that is full of potential confounding factors, there is still no recognition rate <60%. Therefore, the DLM model has stable performance and no overfitting. Among the hypothesized confounding factors, CLBBB and pacing rhythm may have the largest impact on the detection of hypokalemia by DLM. This suggests that we need to be more cautious about the judgment results of DLM when we encounter these two kinds of ECG in clinic. It is worth mentioning that our model still achieved good results in the ECG with features of AF. In addition to the considered confounding factors, the concentrations of serum calcium, troponin, creatinine, and free thyroxine may also obscure the characteristics of hypokalemia ECGs and interfere with the DLM extraction of hypokalemia characteristics, thereby affecting the model used to screen hypokalemia, which needs further research and analysis.

Of course, using AI as a preliminary screening tool for hypokalemia constitutes a qualitative early-warning diagnostic method, regardless of whether the evaluation result is a true positive, whereas biochemical testing remains the gold standard for an unambiguous diagnosis of hypokalemia. In our study, the highest-performing DLM had a false positive rate of 22.9% and a specificity of only 77.1%. This result may be caused by false positive patients under the gold standard test.<sup>[22]</sup> In addition, the potassium level detected by the DLM in patients may better reflect the risk of arrhythmia than blood tests. An ECG reflects the response of heart tissue to the blood potassium level and is thus a direct response based on the serum potassium concentration near the actual myocardium.<sup>[23]</sup> The DLM might be more physiological tool than a blood test from this perspective.

Our study has some limitations at this stage. First, a retrospective study was performed using conventional 12-lead ECG. Prospective studies must be conducted to correlate the DLM with enhanced hypokalemia detection and improved outcomes. Note that the DLM was developed and verified using 12-lead ECG data obtained

in the environment of a hospital ranked among the top three of all hospitals in China. Therefore, prospective testing is required to analyze the DLM performance based on ECG data obtained in a home environment. Similarly, further testing of the detection performance using ECG data from wearable devices is also required. Third, the DLM performance must be further enhanced before application as a reliable detection tool for the serum potassium concentration. Fortunately, the popularity of AI and continuous optimization of deep learning algorithms make it likely that we will develop a better DLM to screen hypokalemia in the near future. Fourth, we only used 8630 ECG data to develop the DLM because of the limited computing power of existing machines. The relatively few ECGs notwithstanding, we used all of the 12-lead ECG data. Thus, the overall quantity of data used in the analysis was not less than that used in studies based on only 4- or 6-lead ECG data.<sup>[24]</sup> The difficulty of limited computing power will be resolved with the upgrading of equipment and the use of more advanced computers. Fifth, the influence of arterial blood gas and blood pH on the ECG cannot be ignored, but unfortunately the above data is not stored electronically in our hospital; hence the influence on the DLM model cannot be further evaluated. In the follow-up research, we will pay attention to this part of the data and collect it manually. Finally, the decision-making process of DLM needs to be further explored. Explainable AI has recently attracted considerable interest in medicine and has been studied and reported on. This consideration motivates our next research direction. We expect to finally uncover the mystery of CNNs and understand their detailed decision-making methods in the near future.

In conclusion, a CNN-based DLM exhibits good performance in screening hypokalemia using 12-lead ECGs and can provide more rapid serum potassium detection capabilities and dynamic detection capabilities for emergency patients than current methods. However, a prospective study needs to be conducted to determine whether the DLM can improve the clinical outcomes of emergency patients.

### Acknowledgements

The authors thank the Translational Medical College of Nanchang University for assisting in developing this deep learning model and also thank Dr. Libin Deng for his guidance.

### Funding

This work was supported by the National Natural Science Foundation of China (No. 81360025).

### Conflicts of interest

None.

### References

1. de Moraes AG, Surani S. Effects of diabetic ketoacidosis in the respiratory system. *World J Diabetes* 2019;10:16–22. doi: 10.4239/wjd.v10.i1.16.

2. Skogestad J, Aronsen JM. Hypokalemia-induced arrhythmias and heart failure: new insights and implications for therapy. *Front Physiol* 2018;9:1500. doi: 10.3389/fphys.2018.01500.
3. Petit PF, Hantson P, Jadoul M, Gillion V. The case | severe hypokalemia complicated by a syncope. *Kidney Int* 2018;94:225–226. doi: 10.1016/j.kint.2017.12.009.
4. Surawicz B. Relationship between electrocardiogram and electrolytes. *Am Heart J* 1967;73:814–834. doi: 10.1016/0002-8703(67)90233-5.
5. Wrenn KD, Slovis BS, Slovis CM. The ability of physicians to predict electrolyte deficiency from the ECG. *Ann Emerg Med* 1990;19:580–583. doi: 10.1016/s0196-0644(05)82194-8.
6. Ting DSW, Cheung CY, Lim G, Tan GSW, Quang ND, Gan A, *et al.* Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA* 2017;318:2211–2223. doi: 10.1001/jama.2017.18152.
7. Galloway CD, Valys AV, Shreibati JB, Treiman DL, Petterson FL, Gundotra VP, *et al.* Development and validation of a deep-learning model to screen for hyperkalemia from the electrocardiogram. *JAMA Cardiol* 2019;4:428–436. doi: 10.1001/jamacardio.2019.0640.
8. Halcox JPP, Wareham K, Cardew A, Gilmore M, Barry JP, Phillips C, *et al.* Assessment of remote heart rhythm sampling using the alivecor heart monitor to screen for atrial fibrillation: the REHEARSE-AF study. *Circulation* 2017;136:1784–1794. doi: 10.1161/CIRCULATIONAHA.117.030583.
9. Yasin OZ, Attia Z, Dillon JJ, DeSimone CV, Sapir Y, Dugan J, *et al.* Noninvasive blood potassium measurement using signal-processed, single-lead ECG acquired from a handheld smartphone. *J Electrocardiol* 2017;50:620–625. doi: 10.1016/j.jelectrocard.2017.06.008.
10. Zheng W, Hong Q, Zhang X, Geng X, Cai G, Chen X, *et al.* Clinical analysis of a hypokalemic salt-losing tubulopathy case. *Chin Med J* 2016;129:601–603. doi: 10.4103/0366-6999.176992.
11. Kwon JM, Kim KH, Jeon KH, Lee SY, Park J, Oh BH. Artificial intelligence algorithm for predicting cardiac arrest using electrocardiography. *Scand J Trauma Resusc Emerg Med* 2020;28:98. doi: 10.1186/s13049-020-00791-0.
12. Ding L, Liu GW, Zhao BC, Zhou YP, Li S, Zhang ZD, *et al.* Artificial intelligence system of faster region-based convolutional neural network surpassing senior radiologists in evaluation of metastatic lymph nodes of rectal cancer. *Chin Med J* 2019;132:379–387. doi: 10.1097/CM9.0000000000000095.
13. Shameer K, Johnson KW, Glicksberg BS, Dudley JT, Sengupta PP. Machine learning in cardiovascular medicine: are we there yet? *Heart* 2018;104:1156–1164. doi: 10.1136/heartjnl-2017-311198.
14. Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, *et al.* An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet* 2019;394:861–867. doi: 10.1016/S0140-6736(19)31721-0.
15. Jiao Z, Choi JW, Halsey K, Tran TML, Hsieh B, Wang D, *et al.* Prognostication of patients with COVID-19 using artificial intelligence based on chest x-rays and clinical data: a retrospective study. *Lancet Digit Health* 2021;3:e286–e294. doi: 10.1016/S2589-7500(21)00039-X.
16. Vantaggiato E, Paladini E, Bougourzi F, Distante C, Hadid A, Taleb-Ahmed A. COVID-19 recognition using ensemble-CNNs in two new chest x-ray databases. *Sensors (Basel)* 2021;21:1742. doi: 10.3390/s21051742.
17. Haug U, Kuntz KM, Knudsen AB, Hundt S, Brenner H. Sensitivity of immunochemical faecal occult blood testing for detecting left- vs right-sided colorectal neoplasia. *Br J Cancer* 2011;104:1779–1785. doi: 10.1038/bjc.2011.160.
18. Steinhubl SR, Waalen J, Edwards AM, Ariniello LM, Mehta RR, Ebner GS, *et al.* Effect of a home-based wearable continuous ECG monitoring patch on detection of undiagnosed atrial fibrillation: the mSToPS randomized clinical trial. *JAMA* 2018;320:146–155. doi: 10.1001/jama.2018.8102.
19. Velagapudi V, O'Horo JC, Vellanki A, Baker SP, Pidikiti R, Stoff JS, *et al.* Computer-assisted image processing 12 lead ECG model to diagnose hyperkalemia. *J Electrocardiol* 2017;50:131–138. doi: 10.1016/j.jelectrocard.2016.09.001.
20. Attia ZI, DeSimone CV, Dillon JJ, Sapir Y, Somers VK, Dugan JL, *et al.* Novel bloodless potassium determination using a signal-processed single-lead ECG. *J Am Heart Assoc* 2016;5:e002746. doi: 10.1161/JAHA.115.002746.
21. Holzinger A, Carrington A, Muller H. Measuring the quality of explanations: the system causability scale (SCS): comparing human and machine explanations. *Kunstliche Intell (Oldenbourg)* 2020;34:193–198. doi: 10.1007/s13218-020-00636-z.
22. Friedman PA, Scott CG, Bailey K, Baumann NA, Albert D, Attia ZI, *et al.* Errors of classification with potassium blood testing: the variability and repeatability of critical clinical tests. *Mayo Clin Proc* 2018;93:566–572. doi: 10.1016/j.mayocp.2018.03.013.
23. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011;12:77. doi: 10.1186/1471-2105-12-77.
24. Cho Y, Kwon JM, Kim KH, Medina-Inojosa JR, Jeon KH, Cho S, *et al.* Artificial intelligence algorithm for detecting myocardial infarction using six-lead electrocardiography. *Sci Rep* 2020;10:20495. doi: 10.1038/s41598-020-77599-6.

---

**How to cite this article:** Wang CX, Zhang YC, Kong QL, Wu ZX, Yang PP, Zhu CH, Chen SL, Wu T, Wu QH, Chen Q. Development and validation of a deep learning model to screen hypokalemia from electrocardiogram in emergency patients. *Chin Med J* 2021;134:2333–2339. doi: 10.1097/CM9.0000000000001650