Serum Potassium Monitoring using AI-enabled Smart Watch Electrocardiograms

I-Min Chiu[1,2,3], Po-Jung Wu[4], Huan Zhang[3], J. Weston Hughes[5], Albert J Rogers[6], Laleh Jalilian[7], Marco Perez[6], Chun-Hung Richard Lin[3], Chien-Te Lee[8], James Zou[9], and David Ouyang[1,*]

1. Department of Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA
2. Department of Emergency Medicine, Chang Gung Memorial Hospital Kaohsiung Branch, Kaohsiung, Taiwan
3. Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan
4. Division of Nephrology, Department of Internal Medicine, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan
5. Department of Computer Science, Stanford University, Palo Alto, CA, USA
6. Department of Medicine and Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA, USA
7. Department of Anesthesiology and Perioperative Medicine, University of California Los Angeles, Los Angeles, CA
8. Division of Nephrology, Department of Internal Medicine, Kaohsiung Chang Gung Memorial Hospital and Kaohsiung Municipal Feng-Shan Hospital, Kaohsiung, Taiwan
9. Department of Biomedical Data Science, Stanford University, Palo Alto, CA, USA

Short title: AI-Based Smartwatch EKG for Potassium Monitoring

Word count: 3,731

Funding Disclosure:

I.M.C receive funding from National Science and Technology Council in Taiwan (111-2221-E-182A-008-); D.O and I.M.C receive funding from Apple and the NIH.

Address for correspondence:

David Ouyang, MD

127 S. San Vicente Blvd., Suite A3600

Los Angeles, CA 90048

david.ouyang@cshs.org

Abstract

Background

Hyperkalemia, characterized by elevated serum potassium levels, heightens the risk of sudden cardiac death, particularly increasing risk for individuals with chronic kidney disease and end-stage renal disease (ESRD). Traditional laboratory test monitoring is resource-heavy, invasive, and unable to provide continuous tracking. Wearable technologies like smartwatches with electrocardiogram (ECG) capabilities are emerging as valuable tools for remote monitoring, potentially allowing for personalized monitoring with AI-ECG interpretation.

Objectives

To develop an AI-ECG algorithm to predict serum potassium level in ESRD patient with smartwatch generated ECG waveforms.

Methods

A cohort of 152,508 patients with 293,557 ECGs paired serum potassium levels obtained within one hour at Cedars Sinai Medical Center (CSMC) was used to train an AI-ECG model (‘Kardio-Net’) to predict serum potassium level. The model was further fine-tuned on 4,337 ECGs from 1,463 patients with ESRD using inputs from 12-lead and single-lead ECGs. Kardio-Net was evaluated in held-out test cohorts from CSMC and Stanford Healthcare (SHC) as well as a prospective international cohort of 40 ESRD patients with smartwatch ECGs at Chang Gung Memorial Hospital (CGMH).

Results

The Kardio-Net, when applied to 12-lead ECGs, identified severe hyperkalemia (>6.5 mEq/L) with an AUC of 0.852 (95% CI 0.745–0.956) and a mean absolute error (MAE) of 0.527 mEq/L. In external validation at SHC, the model achieved an AUC of 0.849 (95% CI 0.823-0.875) and an MAE of 0.599 mEq/L. For single-lead ECGs, Kardio-Net detected severe hyperkalemia with an AUC of 0.876 (95% CI 0.765–0.987) in the primary cohort and had an MAE of 0.575 mEq/L. In the external SHC validation, the AUC was 0.807 (95% CI 0.778–0.835) with an MAE of 0.740 mEq/L. Using prospectively obtained smartwatch data, the AUC was 0.831 (95% CI 0.693–0.975), with an MAE of 0.580 mEq/L.

Conclusions

We validate a deep learning model to predict serum potassium levels from both 12-lead ECGs and single-lead smartwatch data, demonstrating its utility for remote monitoring of hyperkalemia.

Condensed Abstract

Hyperkalemia significantly increases the risk of sudden cardiac death in end-stage renal disease (ESRD) patients. We developed 'Kardio-Net,' an AI-driven ECG model, using data from 152,508 patients at Cedars Sinai Medical Center, and refined it with ECGs from 1,463 ESRD patients using inputs from 12-lead and single-lead ECGs. This model facilitates continuous and non-invasive potassium monitoring, leveraging both traditional and smartwatch-generated ECGs. Tested across

various cohorts, including a prospective smartwatch group, Kardio-Net achieved an AUC range of 0.807 to 0.876, demonstrating its effectiveness for real-time hyperkalemia monitoring.

Keywords: Potassium Monitoring, Hyperkalemia, Artificial Intelligence, Smart Watch, Apple Watch, Electrocardiograms

Abbreviations:

ESRD = End-Stage Renal Disease

ECG = Electrocardiogram

AI = Artificial Intelligence

DL = Deep Learning

SHC = Stanford HealthCare

CGMH = Chang Gung Memorial Hospital

CNN = Convolutional Neural Network

AUC = Area Under the receiver operating characteristic Curve

MAE = Mean Absolute Error

IQR = Interquartile Range

Introduction

Hyperkalemia, or elevated serum potassium, presents a severe and potentially life-threatening risk, in particular for patients with acute kidney injury or chronic kidney disease. Patients with end-stage renal disease (ESRD)[1] are at increased risk for sudden cardiac death due to hyperkalemia[2-3] and require vigilant electrolyte monitoring for hemodialysis[4]. Meanwhile, phlebotomy for serum potassium monitoring is resource-intensive, invasive, and limited in ability to monitor in real time. A noninvasive approach for potassium monitoring could greatly improve the detection of potential lethal electrolyte imbalances. Remote monitoring with smartwatch devices presents a potential solution by offering the capability for remote health monitoring.

There are well known changes in electrocardiogram (ECG) with changes in serum potassium, including the peaking of T waves, QRS prolongation, and PR shortening[5]. Despite recognizable changes, hyperkalemia associated changes are subtle, leading to low sensitivity of physician readers[6]. Recent advancements in artificial intelligence (AI), especially deep learning (DL), have shown significant promise in enhancing the analysis and interpretation of ECG signals[7–13]. Research in the past few years has been instrumental in improving the sensitivity for detecting hyperkalemia, demonstrating the potential of AI to augment clinical decision-making in identifying this dangerous

electrolyte imbalance[14–17]. Recent work have shown the ability for AI-ECG applications originally developed using 12-lead ECGs[18] to be optimized for single lead smartwatch ECGs[13].

We hypothesized an AI-enhanced ECG model (Kardio-Net) can accurately identify hyperkalemia, including in patients with ERSD using ECG data collected from a smartwatch. To test our hypothesis, we develop, validate, and test a model for predicting potassium levels across three international institutions using both 12-lead and smartwatch ECGs in ESRD patients.

Method

Patient identification and data sources

A primary cohort of patients with paired ECGs and serum potassium levels was curated from the CSMC between 2008 to 2022. A total of 153,971 adult patients had at least one 12-lead ECG and serum potassium test within one hour, leading to 297,914 twelve-lead ECGs and corresponding potassium values. ECGs were mapped to the temporally closest potassium test. Given an interest in assessing performance in the ESRD population, all recordings from non ESRD patients (n = 152,508) were used for initial model training, and the model was then finetuned on ESRD patients alone (n = 1,463). The cohorts were randomly split at the patient level into 80% for training set, 10% for validation, and 10% reserved as a held-out test set. For external validation, we identified patient cohorts from two other healthcare systems. We identified 7,586 ECGs among 3,107 ESRD patients at Stanford HealthCare (SHC) from 8/2005 to 6/2018 (Figure 1).

Additionally, a prospective cohort collected at Chang Gung Memorial Hospital (CGMH) from October 2022 to July 2023 aimed to assess the performance of Kardio-Net in detecting hyperkalemia using smartwatch-generated ECG waveforms in ESRD patients. The study recruited adult ESRD patients at CGMH who were aged over 20 years, scheduled for routine hemodialysis in outpatient department and were able to independently sign the informed consent. Prior to participation, written consent was secured from each individual. This study was approved by the institutional review boards of the CSMC, SHC, and CGMH. Patients' informed consent was waived at CSMC and SHC due to the retrospective nature of the study using de-identified ECG and electronic health record data.

Model Development

We developed a convolutional neural network (CNN) designed for ECG interpretation, Kardio-Net, with the capability to integrate digitized waveform for predicting serum potassium levels. All ECGs obtained from CSMC were recorded using a GE machine, adhering to the standard 10-second, 12-lead ECG protocol. These ECG waveform data were captured at a sampling rate of 500 Hz and represented as 10-second, 12×5000 matrices of amplitude values. In external validation at SHC, ECGs were stored using the Phillips TraceMaster system with the same sampling rate with recorded also as 12×5000 matrices and were independently used as input examples for external validation.

Kardio-Net was trained on PyTorch starting from random initialization. Training utilized a mean square error loss for up to 100 epochs, employing an ADAM optimizer with a starting learning rate of 1e-2. Early stopping based on minimizing validation cohorts mean absolute error (MAE). For the fine-tuning in ESRD patients, weights from the initial model were used as the starting point and trained using the same loss function and optimizer with a learning rate of 1e-4. To ensure compatibility with the smartwatch ECG data, we trained the Kardio-Net by utilized lead I from the 12-lead ECGs, as it most closely resembles the ECG vector of smartwatch recordings. For Kardio-Net inference on ECG waveforms smartwatch, we split the waveform data into consecutive 5-second intervals. During the inference, a 30-second ECG waveform was split at 1-second intervals, yielding 26 5-second waveform data inputs. These were then averaged to formulate the final prediction.

Validation on smartwatch ECG

After consent and enrollment of the patients, digital ECG recordings were captured by the Apple Watch ECG within 5 minutes before their routine serum potassium tests, which were conducted every 2 to 4 weeks, prior to hemodialysis. The smartwatch ECG waveforms, initially recorded at a sampling rate of 512 Hz, were resized to 500 Hz to match the sampling rate of ECGs in CSMC. The digitized ECG waveforms stored on the hospital server via the Apple Developer HealthKit API for subsequent analysis[19]. This process yielded a validation dataset of 589 ECGs waveforms paired with its corresponding potassium levels. Given that smartwatch waveforms can contain artifacts and baseline wandering distinct from 12-lead ECGs, we developed a noise ECG classifier to exclude excessively noisy waveforms before averaging the predictions for final analysis (See Supplemental 1). 23 (3.9%) ECGs were excluded from the noise classifier. A domain adaptation technique was employed to facilitate the Kardio-Net's application to Apple Watch ECG waveforms[20]. This adjustment aims to mitigate domain discrepancies, thereby enhancing Kardio-Net‘s performance on Apple Watch ECGs (See Supplemental 2).

Statistical analysis

All analyses were conducted on the held-out test dataset, external validation set, and smartwatch dataset, which were not used during the model's training. The primary metric for assessing Kardio-Net's performance in predicting serum potassium levels were the MAE and Pearson correlation. Additionally, the model's ability to identify severe hyperkalemia, defined as potassium levels greater than 6.5 mEq/L, was evaluated using the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). These metrics allow for a comprehensive assessment of the model's classification accuracy and precision in detecting cases of hyperkalemia. For each metric, we calculated two-sided 95% confidence intervals using 1000 bootstrapped samples to ensure robust statistical inference. In addition, we utilized a Bland-Altman plot to visually assess the agreement between model predictions and actual serum potassium levels, providing further insights into the accuracy and bias of the

predictive models across different ranges of potassium levels. The modeling pipeline was implemented using Python (3.8) with PyTorch (2.0) as the deep learning framework. Signal processing and data analysis were facilitated by Python libraries such as SciPy (1.11), Scikit-learn (1.3.2), pandas (2.0.2), and matplotlib (3.7.1).

Result

Primary cohort characteristics

We identified a total of 153,971 patients at Cedars-Sinai Medical Center who had at least one pair of ECG and serum potassium measurement within 1 hour, resulting in 293,557 ECG recordings. 1,463 patients diagnosed with ESRD contributed 4,337 ECGs were selected as ESRD cohort, where other 293,557 ECGs from 152,508 patients formed the general cohort. The ESRD cohort was divided into a training set of 1,170 patients, a validation set of 146 patients, and a test set of 147 patients. The median age within the ESRD group was 62 years (interquartile range [IQR]: 50-72), with 57.8% (846 patients) being male. The median serum potassium level was 4.6 mEq/L (IQR: 4.1-5.2), with 190 ECGs (4.4%) indicating a potassium level higher than 6.5 mEq/L. Additional demographic and clinical characteristics are detailed in Table 1. Baseline characteristics of patients between hyperkalemia and normokalemia is demonstrated in supplemental table 1.

Kardio-Net performance in the primary cohort

The initial deployment of our 12-lead Kardio-Net model, which was pre-trained on a general cohort, achieved an MAE of 0.393 mEq/L and an AUC of 0.891 (95% CI 0.855–0.927) in predicting severe hyperkalemia (supplemental table 2). Upon validating in the test set of ESRD cohort, the pre-trained model exhibited an MAE of 0.593 mEq/L with predictions concentrating around the 4.0 - 5.0 mEq/L level (Extended Figure 3). After fine-tuning this model with training set in ESRD cohort, there was a notable improvement in prediction accuracy, with the MAE reducing to 0.527 mEq/L. The Pearson correlation coefficient reached 0.674, with a p-value < 0.001, indicating a strong and statistically significant correlation. Kardio-Net achieved discrimination of severe hyperkalemia (potassium > 6.5 mEq/L) with an AUC of 0.852 (95% CI 0.745–0.956) (Figure 2). It predicted severe hyperkalemia with specificity of 0.993, sensitivity of 0.350, PPV of 0.700, and NPV of 0.972 (Table 2).

To adapt to apple watch, we trained a single-lead ECG model using only lead I data through the same process. The model achieved an AUC of 0.876 (95% CI 0.765–0.987) on severe hyperkalemia detection, MAE of 0.575 mEq/L, and correlation coefficient of 0.554 (p-value < 0.001). Further evaluation of Kardio-Net's performance across various patient subgroups, including older individuals and those with diabetes, hypertension, atrial fibrillation, heart failure, and coronary artery disease, was conducted. These groups are generally considered high-risk or exhibit ECG morphology changes. The model's discrimination AUC for detecting severe hyperkalemia varied from 0.725 to 0.983 across these subgroups. Detailed subgroup results are shown in Figure 3.

Kardio-Net performance in the external validation dataset.

The external validation cohort in SHC consisted of a total of 7,586 ECGs among 3,107 patients. The median age was 60 years (IQR: 49-70), with 58.6% (1,820 patients) being male. Atrial fibrillation is present in 25.3% of cases, heart failure in 37.0%, CAD in 35.0%, hypertension in 70.1%, and diabetes mellitus in 48.7%. The median serum potassium level was 4.5 mEq/L (IQR: 4.0-5.1), with 340 ECGs (4.5%) mapping to a severe hyperkalemia result. Other demographic and clinical characteristics are presented in Table 1.

In the external validation dataset, the performances of our 12-lead and single-lead Kardio-Net were consistent with the primary cohort. The 12-lead model demonstrated an AUC of 0.849 (95% CI 0.823-0.875) for hyperkalemia discrimination, achieved an MAE of 0.599 mEq/L and a correlation coefficient of 0.612 (p-value < 0.001). The model predicted severe hyperkalemia with specificity of 0.980, sensitivity of 0.274, PPV of 0.394, and NPV of 0.966. The single-lead model yielded an AUC

of 0.807 (95% CI 0.778–0.835), an MAE of 0.740, and a correlation coefficient of 0.475 (p-value<0.001) in predicting potassium levels. Further analysis in external validation was also conducted using the same subgroup to primary cohort. Kardio-Net's discrimination AUC for detecting severe hyperkalemia varied from 0.780 to 0.860 across these subgroups with consistent performance. The agreement between predicted and actual serum potassium levels in different cohorts, are illustrated in the Bland-Altman plot shown in Extended Figure 4.

Validation in Smartwatch ECG Waveforms

In the prospective international cohort, 589 ECG waveforms were collected from 40 CGMH patients. The median age was 65 years ([IQR]: 58-71), with 16 male patients (40.0%). The median serum potassium level was 4.9 mEq/L (IQR: 4.4-5.3), with 8 ECGs (1.5%) associated with severe hyperkalemia. 23 (3.9%) ECGs were excluded for excessive noise and baseline wander. Single-lead Kardio-Net demonstrated an MAE of 0.580 mEq/L and a correlation coefficient of 0.311 (p-value < 0.001) in predicting serum potassium base on smartwatch ECG. The model achieved an AUC of 0.831 (95% CI 0.693–0.975) for severe hyperkalemia detection using smartwatch ECG waveforms.

A beat-level analysis was performed to interpret Kardio-Net's predictions on a smartwatch (Figure 4). We segmented the original ECG waveform into heartbeat cycles based on the R-point, with 200 samples before and 300 samples after. Averaging heartbeat waveforms corresponding to different serum potassium levels from laboratory tests (Figure 4A) and model predictions (Figure 4B), we observed changes in ECG morphology associated with higher serum potassium levels. These changes included elevated T-waves and widened QRS complexes.

Discussion

Hyperkalemia is a serious, potentially life-threatening electrolyte disturbance that significantly increases the risk for patients with ESRD, tripling their odds of mortality within one day when compared to those with normal potassium levels[21,22]. In this study, we developed a deep learning model, Kardio-Net, to estimate serum potassium levels using both standard 12-lead ECGs and single-lead ECGs including those from wearable devices like the Apple Watch. Across three geographically distinct institutions, our model was able to detect hyperkalemia in ESRD patients. These findings suggest potential for ongoing potassium monitoring in the periods between regular hemodialysis (Central Illustration).

Deep learning models have previously shown proficiency in detecting alterations in ECGs that indicate electrolyte imbalances[9,14,15,23,24]. Several studies significantly advanced the field by employing deep learning techniques to screen for hyperkalemia across patients in outpatient clinic and emergency department, but not exclusively focusing on ESRD patients. Galloway et al[14] utilized 2-lead and 4-lead ECG configurations, achieving an AUC of 0.853 to 0.901. Lin et al.[25], utilizing a

12-lead ECG, highlighted the clinical utility of AI in emergency settings, reflecting high performance metrics.

These earlier studies have suggested that certain ECG features can indicate potassium changes. However, these signs are often not as clear or accurate in patients with ESRD that their cardiac myocyte are less sensitive to changes in potassium, making the usual signs of high potassium less apparent[26–28]. Our study diverges by specifically focusing on an ESRD population, utilizing both 12-lead and single-lead ECG inputs to evaluate their efficacy, and demonstrating the capability to detect severe hyperkalemia using smartwatch devices. The 12-lead model demonstrated a competitive performance with an AUC of 0.849-0.852, closely aligning with Lin's findings, while the single-lead model, suitable for continuous monitoring, showed robust performance with an AUC of 0.831 on wearable device.

Further, our application of transfer learning demonstrates an effective strategy in medical research, particularly where data can be sparse and challenging to collect in certain population. By initially training our deep learning model, Kardio-Net, on a general patient cohort and subsequently fine-tuning it on data exclusively from ESRD patients, we enhance the model's ability to recognize and predict complex, subtle ECG changes associated with hyperkalemia in this population. This approach not only retains the broad learning from the larger general population but also finely adjusts its focus to the specific needs of the ESRD population. Such strategies are well-documented and increasingly vital in biomedical applications of deep learning, where precision in medical treatment and monitoring is paramount[29].

In further analysis, the Bland-Altman plot showed that when serum potassium levels exceed 6.0 mEq/L, significant differences often occur between the predicted and actual serum potassium values. These differences could be due to a combination of factors, including the model's predictive capabilities and the intrinsic variability of potassium levels in ESRD. We observed significant inter-patient variability of ECG changes among ESRD patients, which likely contributed to the challenges in accurately predicting severe hyperkalemia. A critical aspect of this variability is the reduced responsiveness of cardiac myocytes to changes in serum potassium levels, which may obscure the typical ECG indicators of hyperkalemia[28,30]. As a result, ECGs of ESRD patients often do not show the expected changes when serum potassium levels increase. Despite these physiological differences, our model effectively detected important ECG changes. Beat-level analysis (Figure 4) revealed that as Kardio-Net predicted higher serum potassium levels, it identified related changes in ECG morphology, such as elevated T-waves and widened QRS complexes. These findings emphasize the model's ability to capture relevant features even in single-lead ECGs, indicating that despite individual variability in response to potassium changes, some consistent patterns can still be detected and are clinically important.

Previous studies have shown the ability of DL models to detect hyperkalemia, but they also found that as the number of input leads decreased, the overall performance deteriorated[14]. Furthermore, the efficacy of algorithms specifically designed for wearable devices like the Apple Watch, where the signal-to-noise ratio is typically lower, had not been extensively tested[31]. Nevertheless, we adopted several approaches to enhance single-lead Kardio-Net performance as cited in previous literature, including a moving window average prediction from 5-s ECG[32], and noise detection and removal[33]. These methods contributed the consistency of performance transit from 12-lead ECG to single lead ECG. This approach ensured the Kardio-Net's prediction of serum potassium levels remained stable, with the MAE only slightly higher comparing 12-lead to single-lead ECGs.

Study Limitations

Several limitations of the study merit consideration. First, the retrospective design of training data cohort may introduce selection bias of individuals who would receive 12-lead ECGs and serum electrolyte checks within 1 hour. This clinical setting might not reflect the variability found in broader outpatient settings or during daily activities where wearable technologies are used. However, we were able to evaluate the model's performance on a prospectively collected cohort of patients immediately prior to hemodialysis. Despite consisting of a relatively small number of patients, the prospective international cohort, which utilized smartwatch-generated waveforms from diverse racial groups, supports the generalizability of Kardio-Net. The use of real-world clinical data and the innovative application of deep learning techniques to a critical clinical biomarker substantially increase the relevance and potential impact of our findings in patient care management.

Conclusion

Our research proposed a deep learning model, Kardio-Net, is capable of accurately predicting serum potassium levels in ESRD patients by analyzing both standard 12-lead ECGs and single-lead ECG data from wearable devices like the Apple Watch. The ability of this model to consistently identify hyperkalemia and its minimal increase in prediction error when moving from multi-lead to single-lead data supports its potential as a practical tool for continuous monitoring. Such technology could markedly enhance patient outcomes by facilitating the immediate detection of electrolyte imbalances, bridging the monitoring gap for patients undergoing regular hemodialysis.

Perspectives

Competency in Medical Knowledge:

The validation of an AI-ECG algorithm for real-time hyperkalemia monitoring via smartwatches offers a non-invasive, continuous potassium level assessment for high-risk patients.

Translational Outlook:

Translating this AI-ECG innovation into clinical practice requires efficacy evaluations across diverse populations and integration with healthcare systems to better patient outcomes.

Reference

1. Anon. *How common is hyperkalaemia? A systematic review and meta-analysis of the prevalence and incidence of hyperkalaemia reported in observational studies Toby Humphrey, Mogamat Razeen Davids, Mogamat-Yazied Chothia, Roberto Pecoits-Filho, Carol Pollock*. Glen James

2. Anon. *Serum Potassium Levels and Mortality in Hemodialysis Patients: A Retrospective Cohort Study Topic Article Package: Topic Article Package: Diabetes Subject Area: Nephrology Content Sponsor: Karger OLA Akeem A. Yusuf; Yan Hu; Bhupinder Singh; José A. Menoyo; James*. B. Wetmore

3. Huang C-W, Lee M-J, Lee P-T, et al. Low potassium dialysate as a protective factor of sudden cardiac death in hemodialysis patients with hyperkalemia. *PLoS One*. 2015;10:e0139886.

4. Bansal S, Pergola PE. Current management of hyperkalemia in patients on dialysis. *Kidney Int Rep*. 2020;5:779–789.

5. Surawicz B. Relationship between electrocardiogram and electrolytes. *Am Heart J*. 1967;73:814–834.

6. Wrenn KD, Slovis CM, Slovis BS. The ability of physicians to predict hyperkalemia from the ECG. *Ann Emerg Med*. 1991;20:1229–1232.

7. Hannun AY, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med*. 2019;25:65–69.

8. Hughes JW, Olgin JE, Avram R, et al. Performance of a convolutional neural network and explainability technique for 12-lead electrocardiogram interpretation. *JAMA Cardiol*. 2021;6:1285–1295.

9. Chiu I-M, Cheng C-Y, Chang P-K, Li C-J, Cheng F-J, Lin C-HR. Utilization of personalized machine-learning to screen for dysglycemia from ambulatory ECG, toward noninvasive blood glucose monitoring. *Biosensors (Basel)*. 2022;13:23.

10. Yuan N, Duffy G, Dhruva SS, et al. Deep learning of electrocardiograms in sinus rhythm from US veterans to predict atrial fibrillation. *JAMA Cardiol*. 2023. Published onlineOctober 18, 2023. https://doi.org/10.1001/jamacardio.2023.3701.

11. Theurer J, Yuan. Electrocardiographic deep learning for predicting post-procedural mortality: a model development and validation studyD Ouyang. *The Lancet Digital Health*. 6:e70–e78.

12. Tison GH, Sanchez JM, Ballinger B, et al. Passive detection of atrial fibrillation using a commercially available smartwatch. *JAMA Cardiol*. 2018;3:409.

13. Attia ZI, Harmon DM, Dugan J, et al. Prospective evaluation of smartwatch-enabled detection of left ventricular dysfunction. *Nat Med*. 2022;28:2497–2503.

14. Galloway CD, Valys AV, Shreibati JB, et al. Development and validation of a deep-learning model to screen for hyperkalemia from the electrocardiogram. *JAMA Cardiol*. 2019;4:428–436.

15. Lin C-S, Lin C, Fang W-H, et al. A deep-learning algorithm (ECG12Net) for detecting hypokalemia and hyperkalemia by electrocardiography: algorithm development. *JMIR medical informatics*. 2020;8:e15931.

16. Chiu I-M, Cheng J-Y, Chen T-Y, et al. Using deep transfer learning to detect hyperkalemia from ambulatory electrocardiogram monitors in intensive care units: Personalized medicine approach. *J Med Internet Res*. 2022;24:e41163.

17. Kim D, Jeong J, Kim J, et al. Hyperkalemia detection in emergency departments using initial ECGs: A smartphone AI ECG analyzer vs. Board-certified physicians. *J Korean Med Sci*. 2023;38.

18. Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med*. 2019;25:70–74.

19. Anon. HKElectrocardiogram. *Apple Developer Documentation* Accessed March 26, 2024. https://developer.apple.com/documentation/healthkit/hkelectrocardiogram.

20. Chen M, Wang G, Ding Z, Li J, Yang H. Unsupervised domain adaptation for ECG arrhythmia classification. *Annu Int Conf IEEE Eng Med Biol Soc*. 2020;2020:304–307.

21. Einhorn LM, Zhan M, Hsu VD, et al. The frequency of hyperkalemia and its significance in chronic kidney disease. *Arch Intern Med*. 2009;169:1156–1162.

22. Luo J, Brunelli SM, Jensen DE, Yang A. Association between serum potassium and outcomes in patients with reduced kidney function. *Clin J Am Soc Nephrol*. 2016;11:90–100.

23. Cordeiro R, Karimian N, Park Y. Hyperglycemia identification using ECG in deep learning era. *Sensors (Basel)*. 2021;21:6263.

24. Kwon J-M, Jung M-S, Kim K-H, et al. Artificial intelligence for detecting electrolyte imbalance using electrocardiography. *Ann Noninvasive Electrocardiol*. 2021;26:e12839.

25. Lin C, Chau T, Lin C-S, et al. Point-of-care artificial intelligence-enabled ECG for dyskalemia: a retrospective cohort analysis for accuracy and outcome prediction. *NPJ Digit Med*. 2022;5:8.

26. Rafique Z, Hoang B, Mesbah H, et al. Hyperkalemia and electrocardiogram manifestations in end-stage renal disease. *Int J Environ Res Public Health*. 2022;19:16140.

27. Rafique Z, Aceves J, Espina I, Peacock F, Sheikh-Hamad D, Kuo D. Can physicians detect hyperkalemia based on the electrocardiogram? *Am J Emerg Med*. 2020;38:105–108.

28. Aslam S, Friedman EA, Ifudu O. Electrocardiography is unreliable in detecting potentially lethal hyperkalaemia in haemodialysis patients. *Nephrol Dial Transplant*. 2002;17:1639–1642.

29. Kim HE, Cosa-Linan A, Santhanam N, Jannesari M, Maros ME, Ganslandt T. Transfer learning for medical image classification: a literature review. *BMC Med Imaging*. 2022;22:69.

30. Nemati E, Taheri S. Electrocardiographic manifestations of hyperkalemia in hemodialysis patients. *Saudi J Kidney Dis Transpl*. 2010;21:471–477.

31. Canali S, Schiaffonati V, Aliverti A. Challenges and recommendations for wearable devices in digital health: Data quality, interoperability, health equity, fairness. *PLOS Digit Health*. 2022;1:e0000104.

32. Gutierrez Maestro E, De Almeida TR, Schaffernicht E, Martinez Mozos Ó. Wearable-based intelligent emotion monitoring in older adults during daily life activities. *Appl Sci (Basel)*. 2023;13:5637.

33. Vijayakumar V, Ummar S, Varghese TJ, Shibu AE. ECG noise classification using deep learning with feature extraction. *Signal Image Video Process*. 2022;16:2287–2293.

Figure Legend

Central Illustration: Kardio-Net is a convolutional neural network (CNN) designed to predict serum potassium levels in patients with end-stage renal disease (ESRD) using ECG data. Capable of processing both 12-lead and single-lead ECGs, this model adapts to traditional and smartwatch-based systems to provide continuous monitoring. Evaluated on retrospective and prospective cohorts, including smartwatch ECGs, Kardio-Net demonstrates robust performance with an AUC ranging from 0.807 to 0.876 in detecting hyperkalemia.

Figure 1: Selection of study subjects in the primary and external validation cohorts.

Figure 2: Deep learning model performance in internal and external validations. Figure 2A displays a scatter plot of serum potassium predictions from the 12-lead ECG model alongside the ROC curve for hyperkalemia recognition. Figure 2B presents the results from the single-lead ECG model.

Figure 3: Subgroup analysis of hyperkalemia detection performance. Figure 3A illustrates the AUC for different subgroups within CSMC, and Figure 3B for SHC.

Figure 4. Correlated of ECG Waveforms with Serum Potassium Levels. Figure 4A shows mean ECG waveforms associated with measured serum potassium concentrations. The graph displays waveforms for potassium levels categorized into four ranges: 3.5 to 4.5 mEq/L (blue), 4.5 to 5.5 mEq/L (orange), 5.5 to 6.5 mEq/L (green), and >6.5 mEq/L (red), demonstrating the expected ECG changes corresponding to each potassium range. Figure 4B shows mean ECG waveforms corresponding to serum potassium level as predicted by Kardio-Net.

Extended Figure 1: Examples of noisy and non-noisy ECG segments

Extended Figure 2: Illustration of domain adaptation method

Extended Figure 3: Scatter plots and MAE illustrating improved performance of the model in predicting serum potassium levels in the primary ESRD patient cohort after fine-tuning compared to the pre-training phase, using data from 12-lead (A) and single-lead (B) ECGs.

Extended Figure 4. Bland-Altman Plot of ECG Prediction Accuracy Across Different Cohorts

Table 1. Baseline Characteristics of Patients

	CSMC				
Characteristic	Train	Val	Test	Apple watch	Stanford
Number of patients	1,170	146	147	40	3107
Number of ECGs	3,451	421	465	566	7586
Sex (% Male)	682 (58.3%)	81 (55.5%)	83 (56.5%)	16 (40.0%)	1820 (58.6%)
Age, median (IQR), y	62 (50-72)	61 (50-69)	60 (50-72)	65 (58-71)	60 (49-70)
Race					
Caucasian, n (%)	621 (53.1%)	76 (52.1%)	73 (49.7%)	0	1147 (36.9%)
Black, n (%)	304 (26.0%)	39 (26.7%)	41 (27.9%)	0	308 (9.9%)
Asian, n (%)	140 (12.0%)	19 (13.0%)	24 (16.3%)	40 (100%)	542 (17.4%)
Others, n (%)	105 (9.0%)	12 (8.2%)	9 (6.1%)	0	1110 (35.7%)
Clinical History					
Atrial Fibrillation	47 (4.0%)	6 (4.1%)	7 (4.8%)	3 (7.5%)	786 (25.3%)
Heart Failure	222 (19.0%)	16 (11.0%)	26 (17.7%)	4 (10.0%)	1149 (37.0%)
CAD	334 (28.5%)	37 (25.3%)	42 (28.6%)	9 (22.5%)	1087 (35.0%)
Hypertension	534 (45.6%)	63 (43.2%)	79 (53.7%)	26 (65.0%)	2179 (70.1%)
Diabetes Mellitus	401 (34.3%)	45 (30.8%)	56 (38.1%)	15 (37.5%)	1542 (48.7%)
Potassium (mEq/L)					
median (IQR)	4.6 (4.1-5.2)	4.5 (4.0-5.2)	4.5 (4.0-5.1)	4.9 (4.4-5.3)	4.5 (4.0-5.1)
>6.5, n (%)	161 (4.7%)	12 (2.9%)	17 (3.7%)	8 (1.5%)	340 (4.5%)

Table 2. Performance Metrics to detect Severe Hyperkalemia using 12 leads and Single lead ECG

	12 leads		Single lead		
	CSMC	SHC	CSMC	SHC	Apple Watch
Sensitivity	0.350	0.274	0.133	0.297	0.125
Specificity	0.993	0.980	1.0	0.976	1.0
PPV	0.700	0.394	1.0	0.370	1.0
NPV	0.972	0.966	0.972	0.967	0.988

PPV= Positive Predictive Value; NPV=Negative Predictive Value

Figure 1.

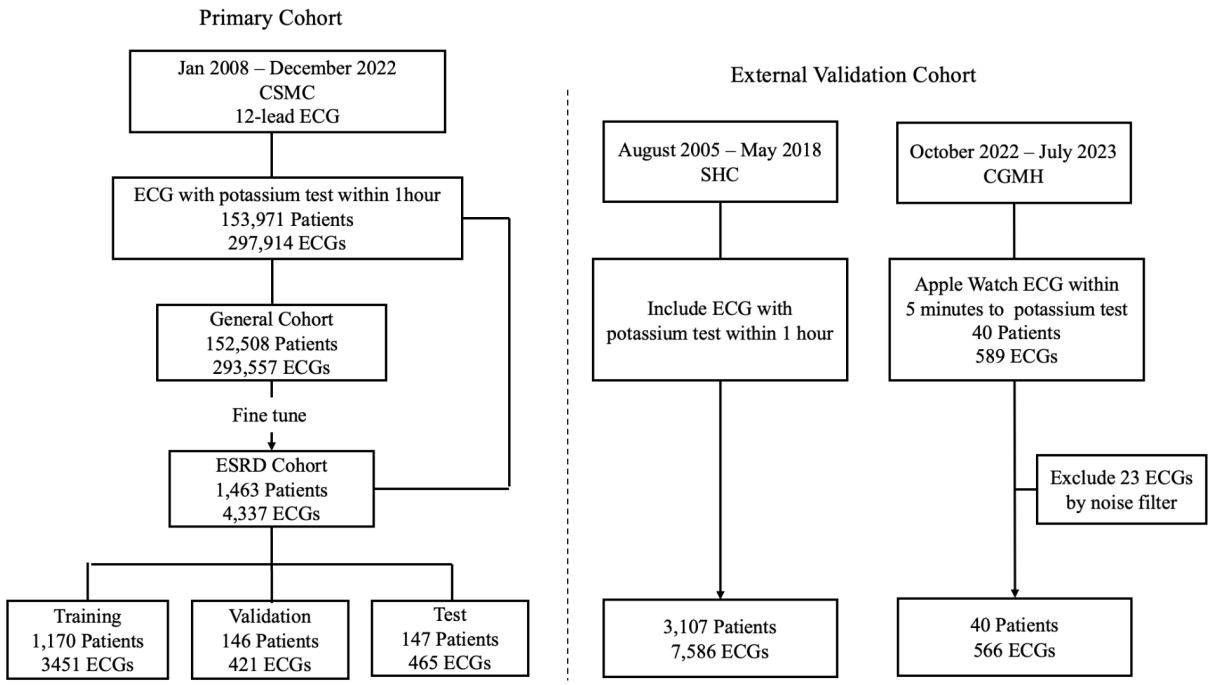


Figure 2.

A	

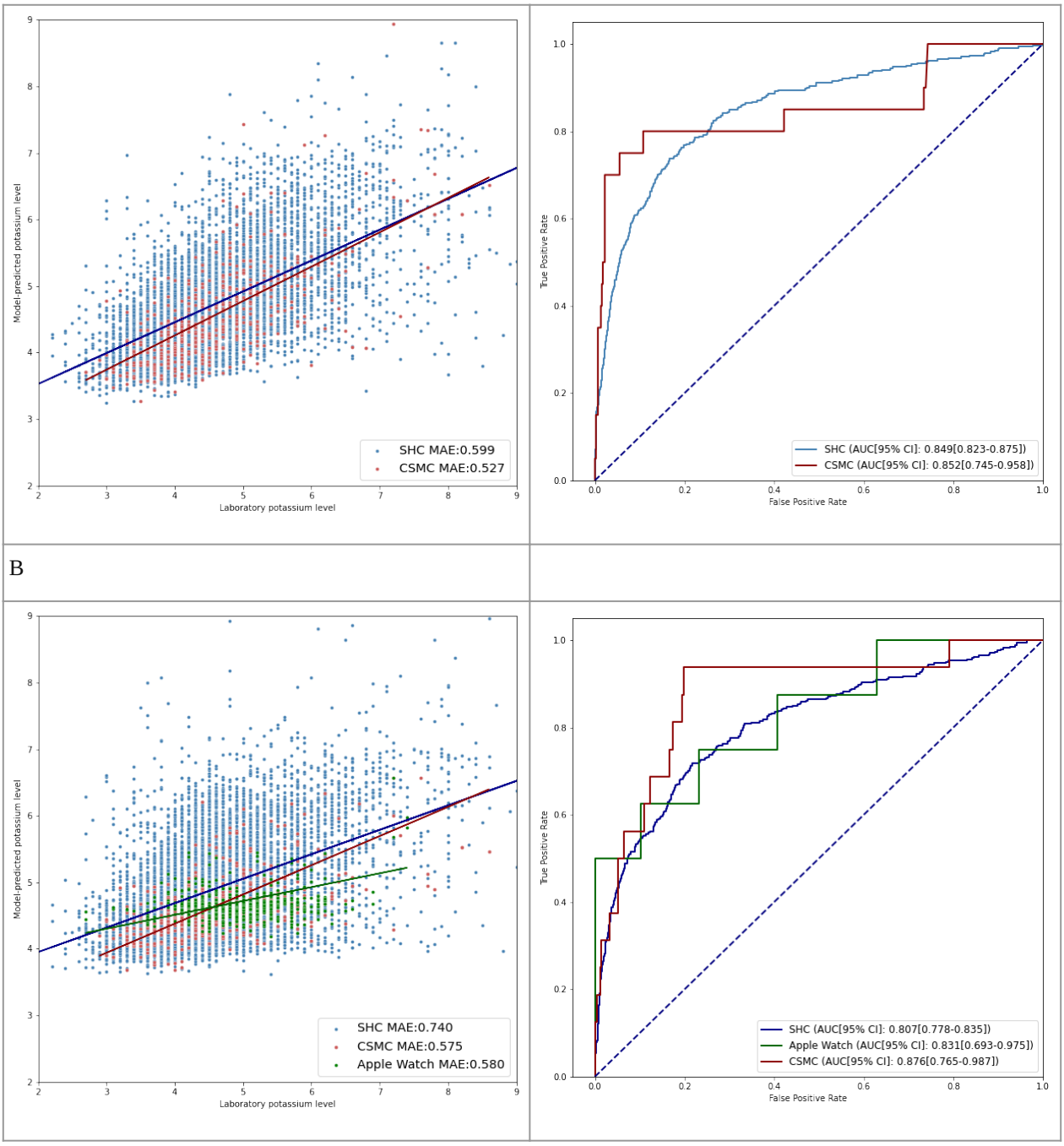

Model-predicted potassium level
Laboratory potassium level
SHC MAE:0.599
CSMC MAE:0.527
True Positive Rate
False Positive Rate
SHC (AUC[95% CI]: 0.849[0.823-0.875])
CSMC (AUC[95% CI]: 0.852[0.745-0.958])
B
Model-predicted potassium level
Laboratory potassium level
SHC MAE:0.740
CSMC MAE:0.575
Apple Watch MAE:0.580
True Positive Rate
False Positive Rate
SHC (AUC[95% CI]: 0.807[0.778-0.835])
Apple Watch (AUC[95% CI]: 0.831[0.693-0.975])
CSMC (AUC[95% CI]: 0.876[0.765-0.987])

Figure 3.

A	B

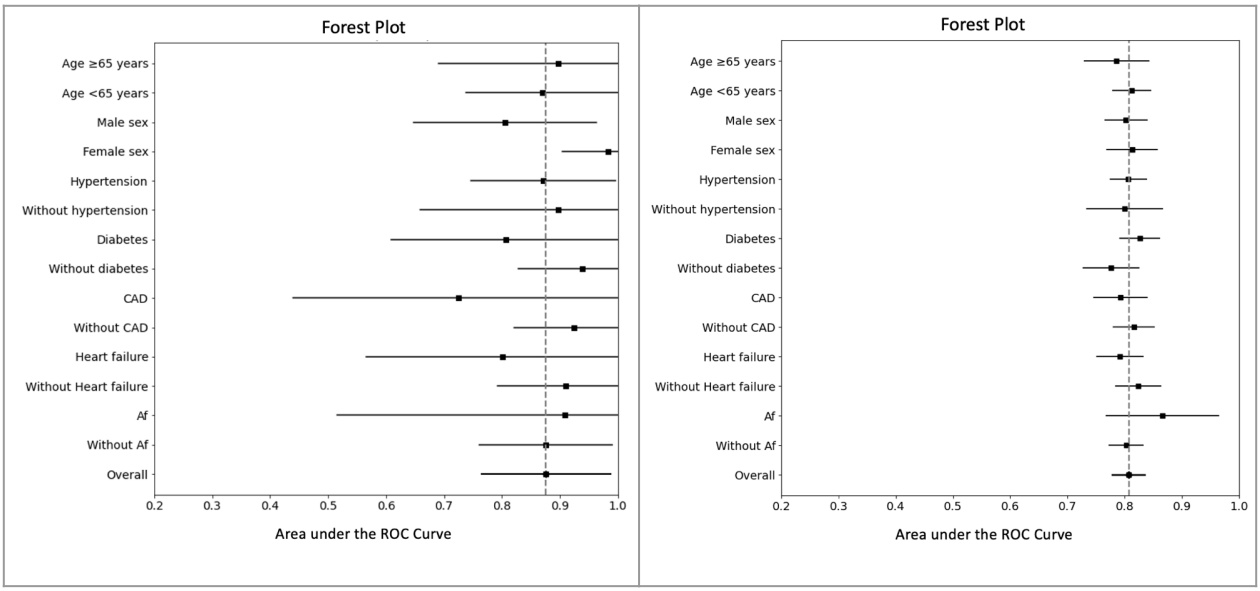
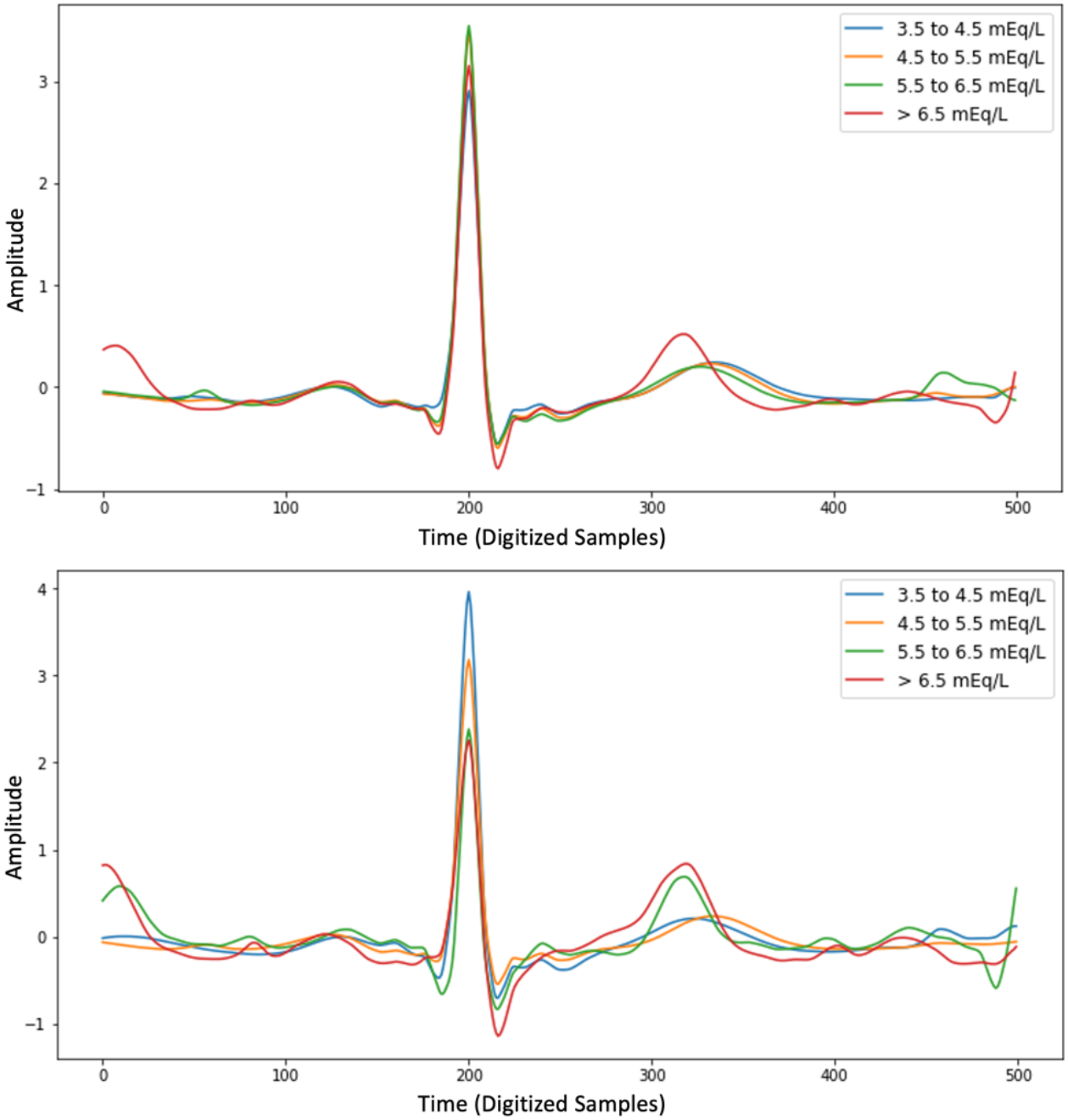

Forest Plot
Age ≥65 years
Age <65 years
Male sex
Female sex
Hypertension
Without hypertension
Diabetes
Without diabetes
CAD
Without CAD
Heart failure
Without Heart failure
Af
Without Af
Overall
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Area under the ROC Curve
Forest Plot
Age ≥65 years
Age <65 years
Male sex
Female sex
Hypertension
Without hypertension
Diabetes
Without diabetes
CAD
Without CAD
Heart failure
Without Heart failure
Af
Without Af
Overall
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Area under the ROC Curve

Figure 4.

A B



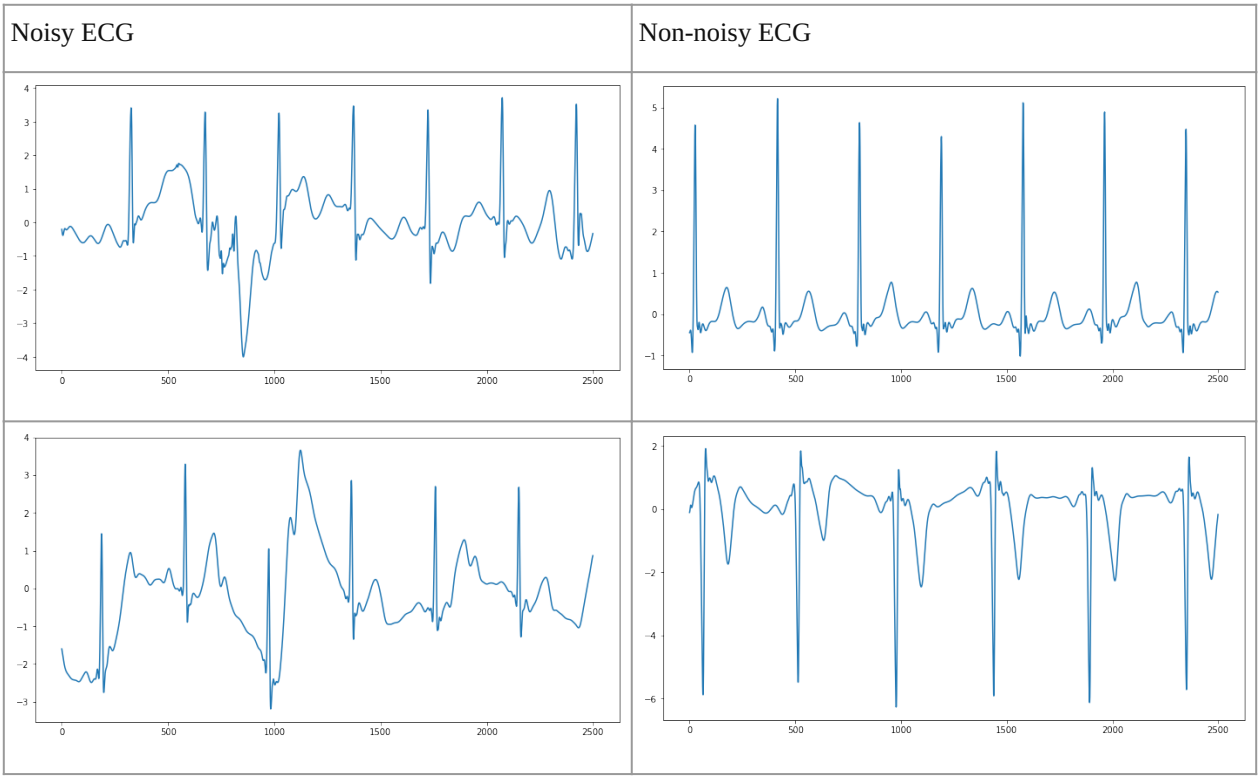
Supplemental 1. Description of Noise Classifier

In this study, we developed an ECG processing pipeline specifically designed to manage raw waveforms from a single lead (Lead I), incorporating a noise classifier to enhance quality control for wearable device applications. To train the noise classifier, we manually labeled 1,000 5-second ECG segments collected from Apple Watch data, identifying 293 segments as noisy. In Extended Figure 1, we provide examples of noisy and non-noisy ECG segments to illustrate the criteria used for classification by our noise classifier. This figure visually differentiates the characteristics of ECG waveforms that were labeled as 'noisy' from those deemed 'non-noisy,' offering insights into the data

quality assurance process implemented in our study. We employed the same CNN architecture used in the primary study to train the classification model, which effectively identified noisy ECGs, achieving an AUC of 0.985 and an accuracy of 0.96.

During the model inference phase, aimed at predicting serum potassium levels from wearable devices, we applied a cutoff threshold of 0.7 to classify ECGs as noisy. ECGs with model prediction probabilities greater than 0.7 were considered noisy and excluded from further analysis. Using this method, out of 589 30-second ECGs assessed in the prospective cohort, 23 were excluded due to noise.

Extended Figure 1.



Supplemental 2. Description of Domain Adaptation Method

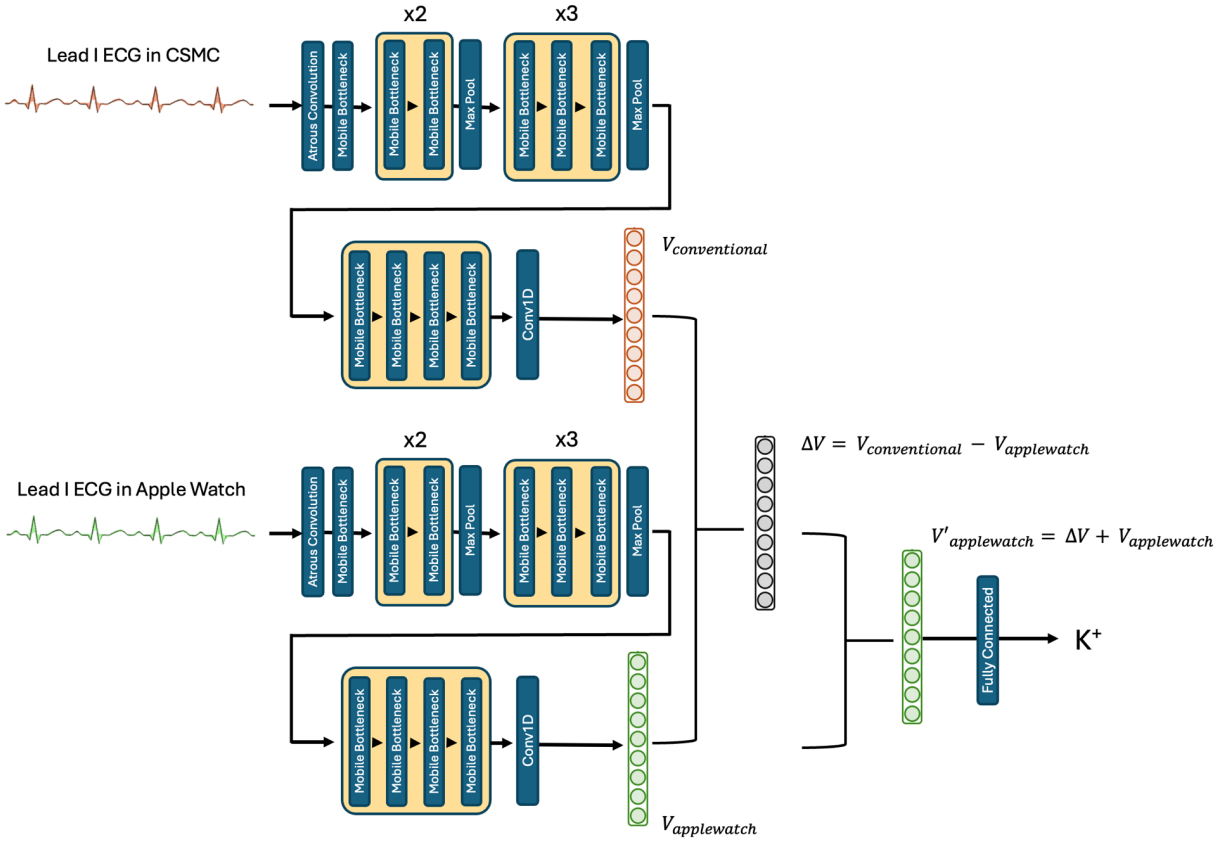
A domain adaptation technique (see extended figure 2) was employed to facilitate the Kardio-Net's application to Apple Watch ECG waveforms. This process hinges on aligning the feature spaces of conventional ECGs and Apple Watch ECGs. From the trained model, we extracted represent the array vector derived from the final average pooling. Similarly, let denote the array vector from the final average pooling layer when the model infers on Apple Watch ECGs.

The adaptation vector, , is calculated as the difference between these two vectors:

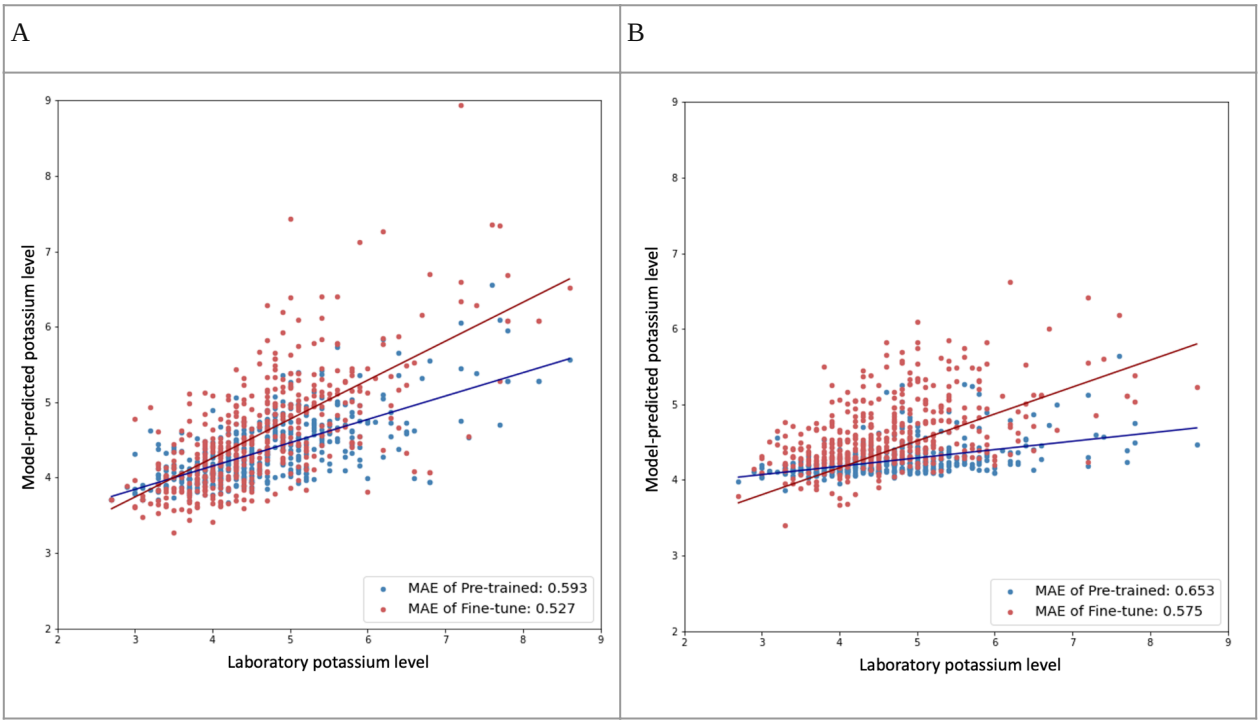
During subsequent inferences on Apple Watch ECG data, we adjust to a new vector, , by applying :

This adjustment aims to mitigate domain discrepancies, thereby enhancing Kardio-Net‘s performance on Apple Watch ECGs.

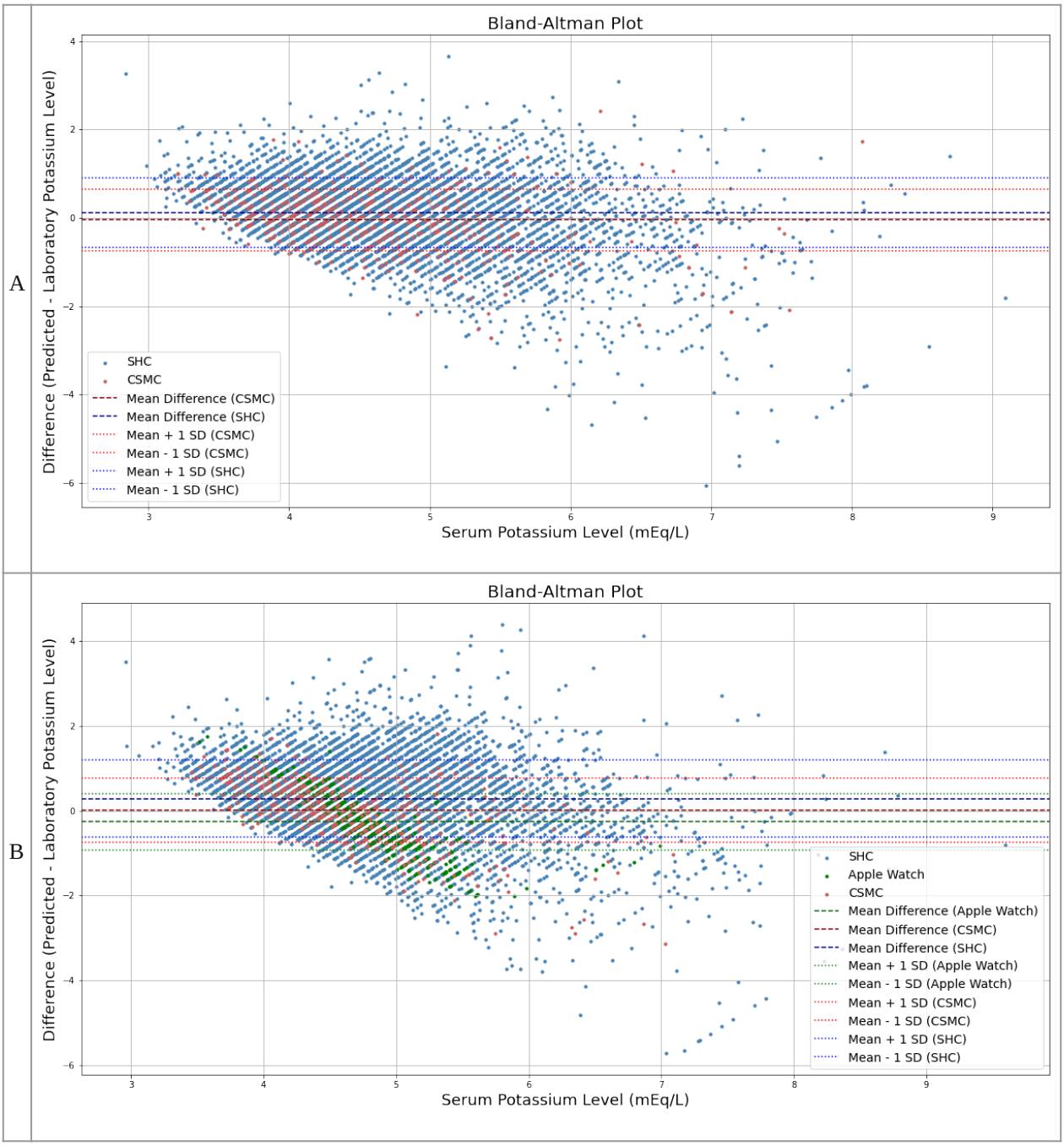
Extended Figure 2.



Extended Figure 3



Extended Figure 4



Supplemental Table 1. Baseline Characteristics of Patients between Hyperkalemia and Normokalemia

Characteristic	Normokalemia	Hyperkalemia
Number of patients	1,446	136
Number of ECGs	4,147	190
Sex (% Male)	841 (58.2%)	86 (63.2%)
Age, median (IQR), y	62 (50-72)	59 (45-68)

Race		
Caucasian, n (%)	760 (52.6%)	70 (51.5%)
Black, n (%)	384 (26.6%)	42 (30.9%)
Asian, n (%)	182 (12.6%)	17 (12.5%)
Others, n (%)	120 (8.3%)	7 (5.2%)
Clinical History		
Atrial Fibrillation	59 (4.1%)	4 (2.9%)
Heart Failure	261 (18.0%)	30 (22.0%)
CAD	412 (28.5%)	45 (33.1%)
Hypertension	666 (46.1%)	70 (51.5%)
Diabetes Mellitus	497 (34.4%)	42 (30.9%)

Supplemental Table 2: Comparative Performance of Regression and Binary Classification Models for Predicting Severe Hyperkalemia in the Primary Cohort

	Regression Model	Binary Classification Model
General cohort	0.891 (0.855–0.927)	0.877 (0.840–0.924)
ESRD cohort	0.852 (0.745–0.956)	0.800 (0.710–0.877)

Supplemental Table 3. Model Performance Across Different Cohorts for Predicting Serum Potassium Levels and Severe Hyperkalemia

	CSMC	SHC	Apple Watch
12 Leads			
MAE	0.527	0.599	
AUC (95% CI)	0.852 (0.745–0.956)	0.849 (0.823-0.875)	
Single Lead (10-seconds ECG)*			
MAE	0.625		
AUC (95% CI)	0.731 (0.643-0.822)		
Single Lead (5-seconds ECG)			

MAE	0.575	0.740	0.580
AUC (95% CI)	0.876 (0.765–0.987)	0.807 (0.778–0.835)	0.831 (0.693–0.975)

*The result of using whole 10-seconds ECG as input for single-lead model in the primary cohort.

Supplemental Table 4. Performance Metrics for Detecting Severe Hyperkalemia Using 12-Lead and Single-Lead ECGs with Lower Prediction Thresholds

	12 leads		Single lead		
	CSMC	SHC	CSMC	SHC	Apple Watch
Sensitivity	0.700	0.500	0.467	0.682	0.500
Specificity	0.978	0.945	0.939	0.806	1.0
PPV	0.583	0.299	0.206	0.142	1.0
NPV	0.987	0.976	0.981	0.982	0.993