

RESEARCH

Open Access



# Chest-X-ray based deep feature-driven framework for anemia detection insights from a large hospital cohort

Hao-Chun Liao<sup>1,4</sup>, Dung-Jang Tsai<sup>3,4</sup>, Shan-Yueh Chang<sup>5</sup>, Chin Lin<sup>3,4,7</sup>, Shih-Hua Lin<sup>6</sup> and Wen-Hui Fang<sup>2,4,7\*</sup>

\*Correspondence:  
Wen-Hui Fang  
rumaf.fang@gmail.com

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

## Abstract

**Background** Anemia is a common global health problem traditionally diagnosed through laboratory hemoglobin testing, which requires invasive blood sampling. This study evaluated the feasibility of using a chest X-ray (CXR)-based deep feature-driven framework for noninvasive anemia prediction.

**Patients and methods** We conducted a retrospective cohort study using data from Tri-Service General Hospital, Taipei, Taiwan, collected between June 2016 and February 2022. A total of 305,793 adults aged  $\geq 20$  years who had at least one CXR were included. The dataset was divided into development, tuning, internal validation, and external validation sets. A two-stage framework leveraging a pretrained Vision Transformer (ViT-B/32) encoder for feature extraction, followed by a downstream logistic regression classifier, was developed to predict anemia ( $Hb \leq 10$  g/dL) from CXR images without model fine-tuning.

**Results** The framework for predicting anemia showed strong performance, with an AUC of 0.845 in the internal validation set (sensitivity 68.5%, specificity 84.7%) and 0.852 in the external set (sensitivity 71.5%, specificity 83.2%). Subgroup analysis revealed better diagnostic accuracy with computed radiography, posteroanterior view, and in males aged 55 to 64, while younger females had lower accuracy. The framework, especially when combined with clinical variables, outperformed other clinical models.

**Conclusion** This study presents a deep feature extraction approach for opportunistic moderate-to-severe anemia screening using chest X-rays. With further validation, this model may facilitate early identification of anemia in routine clinical practice.

**Keywords** Artificial intelligence, Deep learning, Anemia, Chest-X-ray

## Introduction

Anemia is a widespread global health issue, defined by a reduction in the number of red blood cells (RBCs) or a decrease in hemoglobin concentration within the cells, falling below normal levels [1]. According to the World Health Organization (WHO), in 2019, anemia affected 30% (571 million) of women aged 15–49 years, 37% (32 million)



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

of pregnant women, and 40% of children aged 6–59 months [2]. Additionally, over 20% of individuals aged 85 years and older suffer from anemia [3]. It is a common condition across different nations and age groups. Anemia can be triggered by various acute and chronic conditions, surgeries, menstrual blood loss, gastrointestinal bleeding, infections, immune-mediated disorders, and inherited diseases, causing fatigue, cognitive impairment, multi-organ dysfunction, irreversible complications, and even life-threatening conditions [4].

Anemia typically presents with nonspecific symptoms, and a wide range of signs may suggest its presence [4]. As a result, diagnosing anemia based on physical examination alone is challenging. Historically, anemia diagnosis has relied on laboratory assessments of hemoglobin concentration, which are generally conducted through blood tests. While this approach is highly accurate, it is invasive, costly, and dependent on specialized equipment and trained professionals [5]. These resources may not always be accessible, especially in remote or underdeveloped healthcare settings. In clinical practice, we frequently observe that elderly individuals living in rural or underserved areas often lack access to a complete laboratory infrastructure for routine blood testing. Due to logistical and resource constraints, healthcare services in such regions are typically limited to mobile units—such as portable electrocardiogram or X-ray vehicles—that provide basic diagnostic support but are insufficient for comprehensive hematological evaluation. As a result, early identification and monitoring of anemia in these populations remain a significant challenge.

Anemia leads to insufficient oxygen delivery due to a lack of adequate hemoglobin, resulting in a series of cardiovascular changes that ultimately cause left ventricular hemodynamic alterations [6, 7]. Chest X-ray (CXR) is a commonly used tool in hospitals due to its low cost and the availability of extensive databases. However, because anatomical structures are superimposed along the projection direction, interpreting chest X-rays can be challenging. This effect makes detecting subtle abnormalities at specific locations difficult [8]. For human radiologists, it is nearly impossible to interpret anemia from CXR images alone. However, by evaluating signs like the “aortic ring sign” and the “interventricular septum sign,” radiologists can use computed tomography (CT) to assess anemia [9]. Recent studies have explored the use of electrocardiogram (EKG)-based deep learning models (DLM) to predict anemia by analyzing the cardiovascular changes that result in corresponding electrical changes in the heart [10]. Compared to EKG, CXR is theoretically a more intuitive method for observing anemia-related volume and ventricular changes. Additionally, DLM has been widely used in analyzing CXR images for conditions like tuberculosis, lung cancer, COVID-19, and cardiomegaly, although none of these studies have focused on anemia detection [11]. Recent research has also successfully used CXR-based DLM to predict cardiovascular age, further suggesting that DLM can assess cardiovascular status from CXR images [12]. Given the success of EKG-based DLM in predicting anemia and prior studies demonstrating the capability of CXR-based DLM in assessing cardiovascular conditions, we hypothesize that DLM could capture predictive indicators of anemia from CXR images.

Based on the aforementioned discussion, CXRs hold clear potential for capturing subtle cardiovascular changes, and its application within DLM has already been demonstrated in various medical conditions. Building on this foundation, we conducted a retrospective cohort study and developed a deep feature-driven framework to predict

anemia in order to explore these predictive associations. Additionally, we performed an in-depth analysis of subgroups in an effort to identify factors that may affect predictive accuracy and explored the associations between relevant characteristics. Finally, we performed a comparative analysis between the predictive capabilities of the framework and traditional clinical data. Specifically, we sought to determine whether the framework, when combined with imaging data from CXR, could outperform predictions made using conventional clinical information alone. Our goal was to assess whether the integration of machine learning with imaging techniques provides an advantage for clinical anemia scenario.

## Method

### Data source and population

The database utilized for this research was sourced from Tri-Service General Hospital in Taipei, Taiwan, and the study received ethical approval from the hospital's institutional review board (IRB No. C202305019). The CXRs were selected according to the following criteria: (1) Data collection occurred between June 2016 and February 2022; (2) Patients younger than 20 years old were excluded; and (3) The CXRs originated from the emergency department (ER), outpatient department (OPD), or inpatient department (IPD). Each patient contributed only one CXR, chosen at a randomized point in time to ensure a representative sample, as previously applied in other studies. This approach was chosen to provide a more representative sample, as opposed to using only the most recent X-ray, and better simulates the scenario where a patient may present at an outpatient clinic at any given time. Additionally, previous studies have confirmed that despite the randomness in selection, the model's performance remains consistent [13].

The database was divided into four subsets: development, tuning, internal validation, and external validation sets. Their baseline characteristics and image features are presented in Table 1, with the process of data separation illustrated in Fig. 1. A total of 305,793 CXRs met the inclusion criteria, and the corresponding demographic details were analyzed. To facilitate the separation process, 47,898 CXRs from the Jingjhou branch were assigned to the external validation set, ensuring that no data from this branch participated in the model training phase, allowing for independent validation, which was also widely used in previous study [14]. For model development, 135,867 CXRs from branches outside Jingjhou were assigned to the development set, 54,405 to the tuning set for parameter optimization, and 67,623 to the internal validation set. The development and tuning sets are both involved in the training process. The development set serves as the primary dataset for training, while the tuning set is used to fine-tune the model and adjust its training trajectory. Since both of them are already part of the training process, it cannot be considered for independent case predictions and, therefore, will not be included in the result presentation.

### Observation variables

Our main observational variables included both disease-related and basic patient characteristics.

Disease characteristics included heart failure (HF), diabetes mellitus (DM), chronic kidney disease (CKD), coronary artery disease (CAD), hypertension (HTN), Hyperlipidemia (HLP), and Chronic Obstructive Pulmonary Disease (COPD).

**Table 1** Patient characteristics and laboratory results in the development, tuning, internal validation and external validation sets

| Variable                              | Development set<br>(n = 135,867) | Tuning set<br>(n = 54,405) | Internal validation set<br>(n = 67,623) | External validation set<br>(n = 47,898) | p-value |
|---------------------------------------|----------------------------------|----------------------------|-----------------------------------------|-----------------------------------------|---------|
| <b>Main variable</b>                  |                                  |                            |                                         |                                         |         |
| Anemia                                | 12348(9.1%)                      | 4954(9.1%)                 | 5939(8.8%)                              | 4294(9.0%)                              | 0.112   |
| <b>Basic Demographics</b>             |                                  |                            |                                         |                                         |         |
| Age                                   | 51.4 $\pm$ 19.1                  | 51.5 $\pm$ 19.1            | 50.9 $\pm$ 18.9                         | 52.2 $\pm$ 20.8                         | <0.001  |
| Gender(male)                          | 71323(52.5%)                     | 28493(52.4%)               | 34806(51.5%)                            | 24436(51.0%)                            | <0.001  |
| PA view                               | 127729(94.0%)                    | 51195(94.1%)               | 63805(94.4%)                            | 45255(94.5%)                            | <0.001  |
| <b>Acquisition location</b>           |                                  |                            |                                         |                                         |         |
| Emergency Room                        | 41238(30.4%)                     | 16613(30.5%)               | 20769(30.7%)                            | 17583(36.7%)                            | <0.001  |
| Inpatient Department                  | 50797(37.4%)                     | 20526(37.7%)               | 25233(37.3%)                            | 15173(31.7%)                            | <0.001  |
| Outpatient Department                 | 43832(32.3%)                     | 17266(31.7%)               | 21621(32.0%)                            | 15123(31.6%)                            | <0.001  |
| <b>Disease history</b>                |                                  |                            |                                         |                                         |         |
| Diabetes Mellitus                     | 13689(10.1%)                     | 5472(10.1%)                | 6177(9.1%)                              | 7231(15.1%)                             | <0.001  |
| Hypertension                          | 2784(2.0%)                       | 1101(2.0%)                 | 1248(1.8%)                              | 1856(3.9%)                              | <0.001  |
| Hyperlipidemia                        | 17918(13.2%)                     | 7204(13.2%)                | 7603(11.2%)                             | 10821(22.6%)                            | <0.001  |
| Chronic Kidney Disease                | 8686(6.4%)                       | 3511(6.5%)                 | 4017(5.9%)                              | 4016(8.4%)                              | <0.001  |
| Heart failure                         | 3926(2.9%)                       | 1585(2.9%)                 | 1658(2.5%)                              | 2280(4.8%)                              | <0.001  |
| Coronary Artery Disease               | 11188(8.2%)                      | 4518(8.3%)                 | 4980(7.4%)                              | 5993(12.5%)                             | <0.001  |
| Chronic Obstructive Pulmonary Disease | 8688(6.4%)                       | 3477(6.4%)                 | 3215(4.8%)                              | 5018(10.5%)                             | <0.001  |

All the *p* values of the characteristics were less than 0.001, except anemia. The mean age ranged from 50.9 to 52.2 years

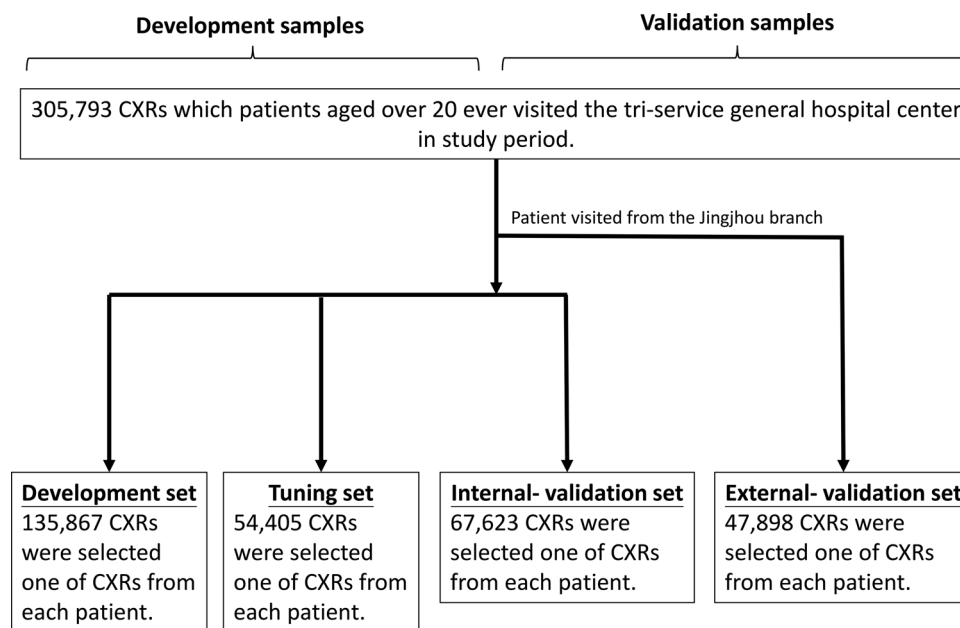


**Fig. 1** Development, tuning, internal validation and external validation set generation. This diagram outlines the dataset creation and analysis strategy, designed to ensure a robust and reliable dataset for training, validating, and testing the network. Each patient's data was assigned to one specific set—development, tuning, or validation—and once assigned, the data remained in that set to prevent any overlap or “cross-contamination.” the external validation set was created using data exclusively from patients who visited the Jingzhou branch, while other patients contributed to different datasets. Further details on the flow and specific use of each dataset can be found in the methods section

Basic patient characteristics included age, sex, acquisition location (e.g., IPD, OPD, or ER), posteroanterior (PA) or anteroposterior (AP) view, and hemoglobin level. The endpoint of this study was moderate-to-severe anemia, defined as hemoglobin  $\leq 10.0$  g/dL, based on prior research and clinical action thresholds [10, 15]. Hemoglobin values were collected within one day before or after the CXR.

The definitions of the other variables were described in Supplemental material.

### The CXR report analysis

To gain a deeper understanding of the relationship between lung structure and anemia, we analyzed radiologist-typed CXR reports. These reports, based on images obtained by radiologic technologists, were first preprocessed to standardize their content. We then extracted and analyzed key terms and phrases identified by the radiologists to explore potential associations between radiographic findings and anemia.

### The implementation of the DLM

The overall modeling framework was consistent with our previous study [16].

The CXR images used in this study were stored in DICOM format with a resolution exceeding  $2000 \times 2000$  pixels. For feature extraction, we adopted a framework based on Pretraining Approach for Small-sample Training Employing Radiographs (PASTER) model [16]. PASTER is a multimodal pretraining approach based on the Contrastive Language-Image Pretraining (CLIP) framework [17], which leverages paired chest radiographs and free-text radiology reports to learn joint image-text representations. Similar techniques have been successfully implemented in models such as ConVIRT and BioMedCLIP, demonstrating superior performance in medical imaging tasks [18–20]. Through large-scale pretraining, the model captures clinically meaningful radiographic patterns and generates transferable visual embeddings for downstream applications. This strategy reduces the reliance on extensive manually annotated datasets while improving generalizability across different clinical scenarios. The design of PASTER is conceptually related to earlier image-text contrastive learning frameworks, such as CheXzero [21], which align radiographic images with textual descriptions to learn meaningful visual representations from radiology reports [17].

In this study, a pretrained Vision Transformer (ViT-B/32) was used as the image encoder to extract embeddings from each CXR [22]. The pretrained weights were used without additional fine-tuning. Each input image was resized to  $256 \times 256$  pixels and processed by the encoder to generate a 512-dimensional feature vector representing the radiographic characteristics of the image. These embeddings were subsequently used as input features to construct a logistic regression model, referred to as the CXR-risk score, for anemia prediction.

To benchmark model performance, we additionally trained an XGBoost classifier using the same dataset. The XGBoost model was applied to (1) clinical variables alone, (2) CXR-derived embeddings alone, and (3) a combination of both clinical variables and CXR-derived features to evaluate the potential benefit of integrating imaging representations with patient characteristics.

Embedding extraction was implemented in Python (version 3.10.10) using the Torch package (version 2.0.1), while XGBoost models were developed using standard machine learning frameworks.

**Outcome and statistical analysis**

The details were described in Supplemental material.

**Results**

**Baseline characteristics**

Patient characteristics in the development, tuning, internal validation and external validation sets are shown in Table 1. All the *p* values for the characteristics were less than 0.001, except for the anemia rate. This discrepancy may be attributed to differences in population composition between the internal and external validation sets. The mean age ranged from 50.9 to 52.2 years.

**Performance of the framework to predict anemia**

The Receiver Operator Characteristic (ROC) Curve and the related variances, which are presented in Fig. 2, are used to calculate the Area Under Curve (AUC) and identify an optimal cutoff point for predicting anemia.

In the internal validation set, the AUC curve was 0.845, with a sensitivity of 68.5%, specificity of 84.7%, positive predictive value of 30.0%, and negative predictive value of 96.5%. In the external validation set, the AUC curve was 0.852, with a sensitivity of 71.5%, specificity of 83.2%, positive predictive value of 29.5%, and negative predictive value of 96.7%.

**AUC values for anemia detection across various subgroups**

We analyzed the AUC values for anemia detection across various subgroups in both internal and external validation datasets, presented in Fig. 3. Factors considered included age, gender, diagnostic method (e.g., computed radiography, digital radiography), clinical settings (e.g., ER, IPD, OPD), and underlying clinical conditions. Overall, performance varied across these dimensions, with notable differences between the internal and external validation results.

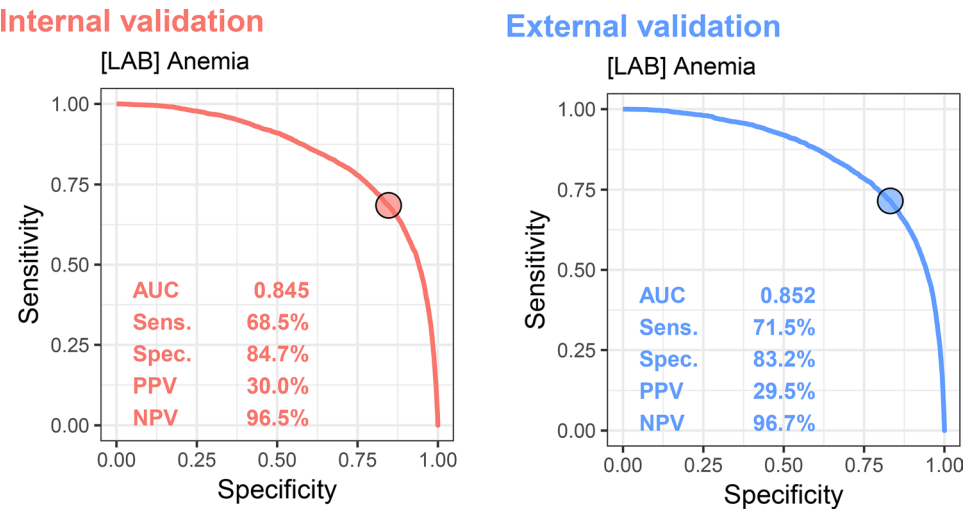


**Fig. 2** Receiver operating characteristic (ROC) curves for predicting anemia. The Area under the curve (AUC) values are 0.845 for internal validation and 0.852 for external validation. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) are indicated for both sets. The figure illustrates the performance of the model in identifying an optimal cutoff point for anemia diagnosis across internal and external validation datasets

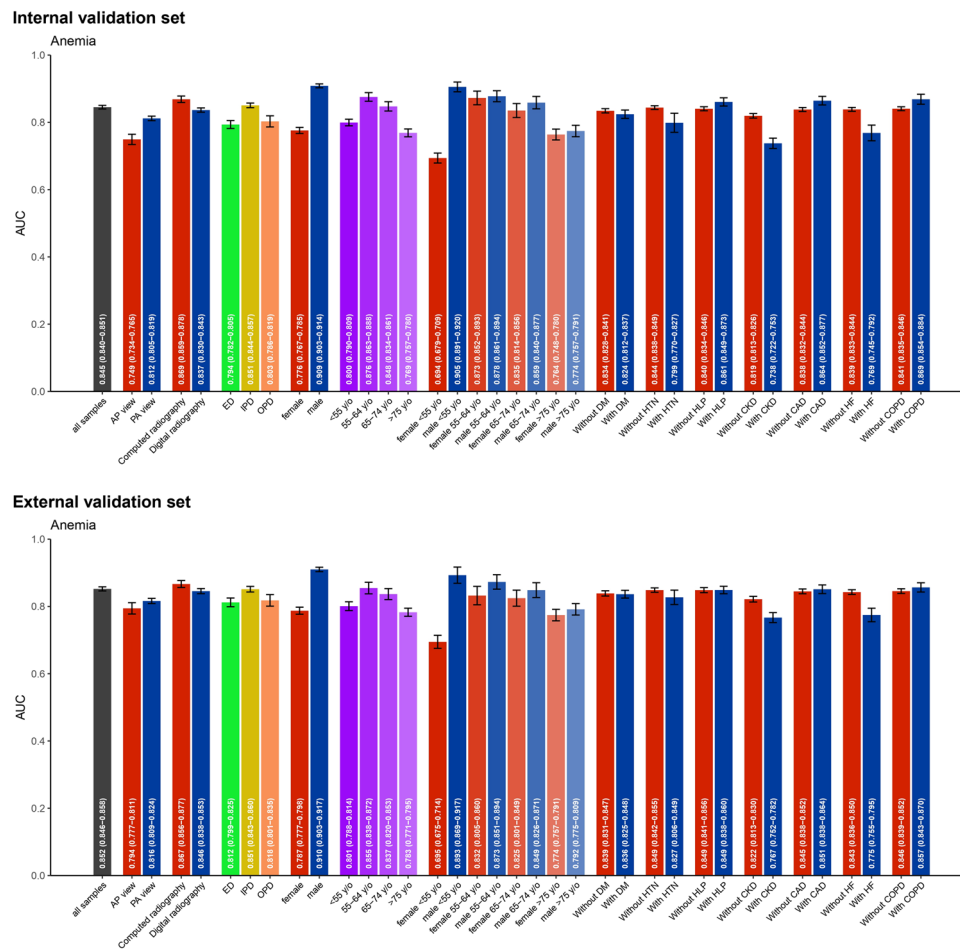


**Fig. 3** AUC values for anemia detection across various subgroups in both internal and external validation sets. This figure illustrates the performance of the model in different patient populations, including variations by age, gender, diagnostic method (e.g., computed radiography, digital radiography), clinical settings (e.g., emergency department, inpatient, outpatient), with and without clinical underlying conditions, such as diabetes (DM), hypertension (HTN), hyperlipidemia (HLP), chronic kidney disease (CKD), coronary artery disease (CAD), heart failure (HF), and chronic obstructive pulmonary disease (COPD)

In the internal validation, computed radiography (AUC 0.869, 95% CI: 0.859–0.878) and PA view (AUC 0.812, 95% CI: 0.805–0.819) showed better predictive performance. The inpatient department (AUC 0.851, 95% CI: 0.844–0.857) performed better than the outpatient and emergency departments. Regarding age and gender, males (AUC 0.909, 95% CI: 0.903–0.914) and individuals aged 55 to 64 years (AUC 0.876, 95% CI: 0.863–0.888) demonstrated superior predictive ability, although a significant decline was observed in younger females under 55 years. For underlying clinical conditions, patients with hypertension (AUC 0.794, 95% CI: 0.782–0.805), chronic kidney disease (AUC 0.764, 95% CI: 0.748–0.780), and heart failure (AUC 0.767, 95% CI: 0.752–0.782) exhibited lower predictive performance compared to those without these conditions.

In the external validation, similar patterns emerged. Computed radiography (AUC 0.867, 95% CI: 0.856–0.877) again showed high performance, but there was no statistical difference between AP and PA views (AUC 0.816, 95% CI: 0.809–0.824). The IPD (AUC 0.855, 95% CI: 0.838–0.872) outperformed the OPD and ER. In terms of demographic factors, males (AUC 0.910, 95% CI: 0.903–0.917) and individuals aged 55 to 64 years

(AUC 0.876, 95% CI: 0.863 – 0.888) showed better performance, but detection ability was reduced in younger females. Additionally, patients with chronic kidney disease (AUC 0.774, 95% CI: 0.757 – 0.791) and heart failure (AUC 0.775, 95% CI: 0.755 – 0.795) demonstrated lower predictive accuracy compared to those without these conditions.

Related important features with anemia

In Fig. 4A, the subgroup with a history of chronic kidney disease had the highest relative importance in predicting anemia. Other important features included age (64.4% relative importance compared to the chronic kidney disease subgroup), PA or AP view (49.6%), nasogastric tube (38.3%), and costophrenic angle blunting (32.4%).

Comparison with other prediction models

In Fig. 4B, the predictive performance of the deep learning model was benchmarked against clinical baselines. In the internal validation set, while a model relying solely on patient characteristics achieved an AUC of 0.808 (95% CI: 0.802–0.814), the stand-alone DLM demonstrated superior discriminative capacity with an AUC of 0.845 (95% CI: 0.840–0.851). Crucially, integrating CXR-derived features with clinical data yielded a substantial incremental benefit, with the combined model achieving the highest AUC of 0.878 (95% CI: 0.873–0.882). A consistent trend was observed in the within-system external validation set; although patient characteristics alone provided a baseline AUC of 0.828 (95% CI: 0.822–0.835), the standalone DLM achieved 0.852 (95% CI:

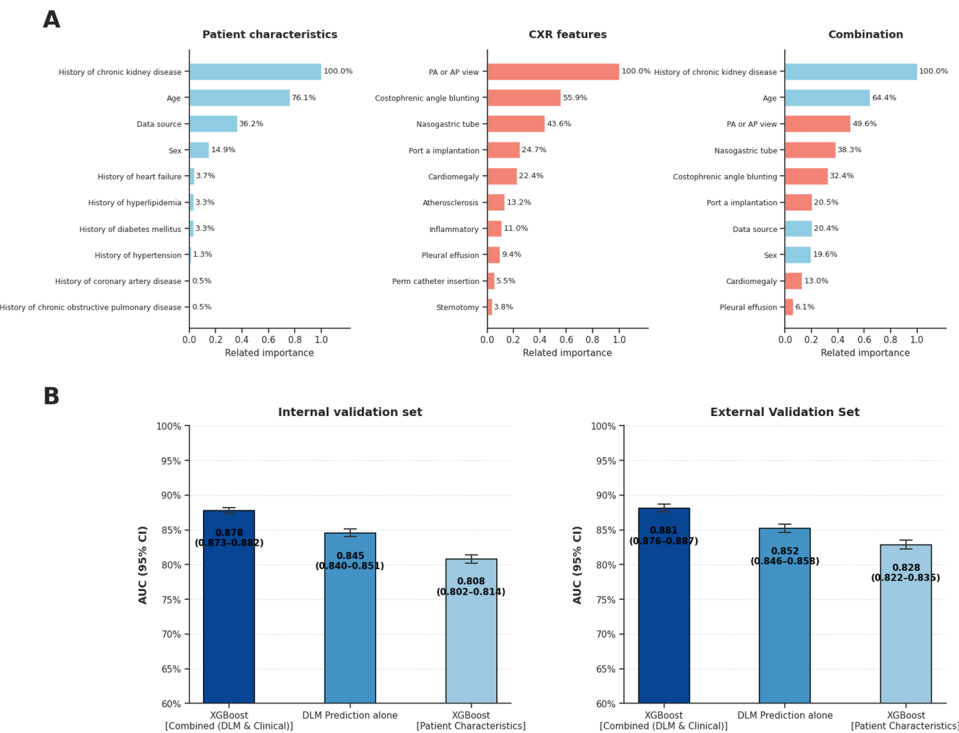


**Fig. 4** The related important about patient characteristics and CXR features in predicting anemia, along with AUC comparisons between several models for internal and external validation sets. **(A)** The relative importance of various patient characteristics (e.g., history of chronic kidney disease, age, history of heart failure) and chest X-ray (CXR) features (e.g., PA or AP view, costophrenic angle blunting, nasogastric tube) are displayed. A combination of these factors was also evaluated for their impact on the prediction model. **(B)** Area under the curve (AUC) values with 95% confidence intervals (CI) are presented for different models in both the internal and external validation sets

0.846–0.858), and their integration further elevated the predictive performance to 0.881 (95% CI: 0.876–0.887). This significant performance leap upon incorporating imaging data highlights that the DLM captures critical, independent physiopathological indicators of anemia that are not encapsulated by routine clinical demographics alone.

**Performance across clinical thresholds and disease-free subgroups**

To address clinical adaptability, we evaluated the model at different thresholds (Table 2A). Using a High-risk cut-off, the model achieved a specificity of 0.93–0.94 and improved the PPV to 44%, providing a more conservative option for clinical use. Furthermore, we conducted a subgroup analysis by excluding patients with major comorbidities to evaluate potential confounding effects (Table 2B). The model demonstrated consistent performance; particularly in the external set, the model achieved an AUC of 0.81 and an external sensitivity of 0.75, though the PPV fell to 0.12–0.16. Nonetheless, the sensitivity remained stable, suggesting the model identifies radiographic patterns associated with anemia rather than secondary signs of cardiopulmonary failure.

**Clinical utility and decision curve analysis**

To evaluate the practical clinical utility of the FRAMEWORK, we performed a Decision Curve Analysis (DCA) to compare the net benefit of model-guided screening against “Treat All” (universal blood testing) and “Treat None” strategies (Fig. 5). The FRAMEWORK demonstrated a consistently higher net benefit across a broad range of threshold probabilities, specifically from 10% to 60%. This indicates that within these clinically relevant risk thresholds, using the FRAMEWORK to trigger follow-up hemoglobin testing provides superior clinical value while minimizing unnecessary diagnostic procedures.

**Table 2** Subgroup analysis and model robustness across various cut-off points and comorbidity statuses

| A. Separate by different cut point |            |             |      |      |      |      |      |
|------------------------------------|------------|-------------|------|------|------|------|------|
| Patient                            | cut point  | N           | AUC  | Sen  | Spe  | PPV  | NPV  |
| internal-set                       | Mid risk   | 67623       | 0.84 | 0.68 | 0.85 | 0.30 | 0.97 |
| external-set                       | Mid risk   | 47879       | 0.85 | 0.71 | 0.83 | 0.29 | 0.97 |
| internal-set                       | High risk  | 67623       | 0.84 | 0.5  | 0.94 | 0.44 | 0.95 |
| external-set                       | High risk  | 47879       | 0.85 | 0.53 | 0.93 | 0.44 | 0.95 |
| B. Separate by disease             |            |             |      |      |      |      |      |
| Patient                            | cut point  | Case number | AUC  | Sen  | Spe  | PPV  | NPV  |
| internal-set                       | No Disease | 51565       | 0.81 | 0.68 | 0.78 | 0.16 | 0.97 |
| external-set                       | No Disease | 29847       | 0.81 | 0.75 | 0.70 | 0.12 | 0.98 |
| internal-set                       | Disease    | 16058       | 0.85 | 0.80 | 0.76 | 0.42 | 0.94 |
| external-set                       | Disease    | 18032       | 0.85 | 0.74 | 0.81 | 0.41 | 0.94 |

**(A) Model performance across different cut-off thresholds.** The Area Under the Curve (AUC) values and performance metrics are presented for specific subgroup analyses. To evaluate the sensitivity-specificity trade-off, the model was assessed at different clinical thresholds, specifically categorized into middle-risk and higher-risk groups to facilitate clinical decision-making

**(B) Model performance in disease-present vs. disease-free subgroups.** To address potential confounding effects from structural abnormalities, a specialized subgroup analysis was conducted by excluding patients with major chronic and cardiopulmonary comorbidities. This comparison demonstrates the model’s ability to identify anemia-related radiographic patterns independently of secondary signs associated with chronic conditions. The disease-free subgroup strictly excludes patients with major chronic and cardiopulmonary comorbidities. Specifically, individuals diagnosed with DM, HTN, HLP, CKD, CAD, HF, or COPD based on standard ICD codes were systematically excluded (as detailed in the Supplementary Material)

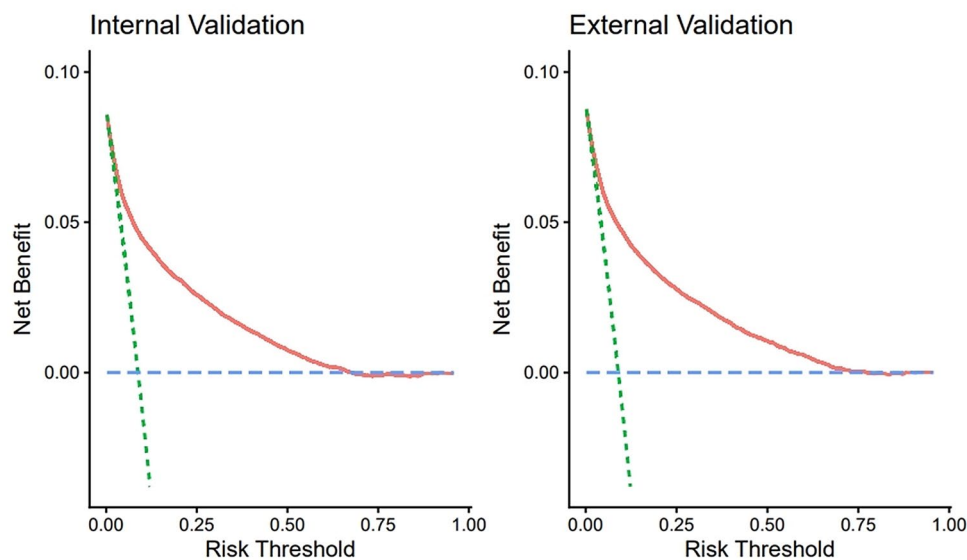


**Fig. 5** Clinical decision curve analysis. The red line (DLM) shows a higher net benefit than the strategies of “Treat all” (green) and “Treat none” (blue). Within the risk threshold range of 10% to 60%, the DLM provides superior clinical utility for anemia screening

#### Calibration performance

The global calibration of the model was further evaluated and optimized across the independent testing cohorts. Following post-hoc recalibration, the prediction error drastically minimized, yielding a standard Brier score of 0.06306 and a Scaled Brier score of 0.21283 for the internal validation set (baseline prevalence: 8.8%). Crucially, this high-quality calibration sustained robustly on the independent external screening cohort, demonstrating a standard Brier score of 0.06326 and a Scaled Brier score of 0.22516 (baseline prevalence: 9.0%). This corresponds to a substantial 21.28% and 22.52% reduction in overall predictive uncertainty, respectively, with empirical calibration curves for both cohorts tracking tightly along the ideal reference line (Supplementary Figure 1). Detailed methodological descriptions are provided in the Supplementary material.

#### Discussion

In this retrospective cohort study, we developed a CXR-based deep feature driven framework to detect anemia. The ROC curve revealed AUC values of 0.845 for internal validation and 0.852 for external validation, with both sets displaying high negative predictive values and moderate sensitivity and specificity in predicting anemia. Subgroup analysis showed that in the internal validation, computed radiography and the inpatient setting performed the best, with males and individuals aged 55–64 achieving the highest predictive accuracy. In contrast, younger females and patients with hypertension, chronic kidney disease, and heart failure showed lower predictive performance, a trend also observed in the external validation set. Further analysis identified several key features associated with anemia prediction, with a history of chronic kidney disease having the highest relative importance. Other notable factors included age, PA or AP view, nasogastric tube placement, and costophrenic angle blunting. Comparisons of various models demonstrated that the DLM combined with patient data had the highest AUC in the internal validation set, while other models, including those using all available

information or the DLM alone, also performed well. A similar trend was seen in the external validation, further underscoring the model's predictive strength.

Our study demonstrated that the framework has excellent ability in predicting anemia. According to previous studies and our own findings, anemia can induce noticeable changes in chest imaging [9, 23–25]. From a physiological standpoint, when hemoglobin levels are greater than 10 g/dL, non-hemodynamic mechanisms, such as elevated 2,3-diphosphoglycerate (2,3-DPG), enable a rightward shift in the oxygen dissociation curve, which typically compensates for hemoglobin deficits [26]. However, in non-resting conditions or when hemoglobin falls below 10 g/dL, cardiac output begins compensating for tissue hypoxia [6]. Several mechanisms are responsible for this compensatory response. First, afterload decreases due to reduced systemic vascular resistance, potentially caused by hypoxic vasodilation induced by hypoxia-generated metabolites and flow-mediated vasodilation driven by increased blood flow. This vasodilation effect is mediated by endothelial cells and endothelium-derived relaxing factors [27]. Secondly, preload increases due to enhanced venous return [6]. Finally, and most importantly, anemia increases preload through the Frank-Starling mechanism and augments inotropic response via sympathetic activation, leading to changes in left ventricular hemodynamic performance [7, 28]. However, these changes are difficult for human interpretation on CXR alone. Nevertheless, previous research on unenhanced chest CT has shown that the aortic ring sign is sensitive (82.5%) and specific (87%) in detecting anemia, while the interventricular septum sign is less sensitive (32%) but highly specific (99.2%). These results suggest that anemia can be detected using unenhanced chest CT scans [9]. Moreover, The objective analysis data revealed a strong correlation between blood pool attenuation at different anatomical sites and serum hemoglobin levels [23]. Thus, it can be concluded that anemia indeed causes cardiovascular changes detectable on chest imaging. Several studies also support this conclusion [24, 25]. To further elucidate the model's decision-making process, we employed feature visualizations, which identified several anatomically consistent focus areas. Our results showed significant model attention on the hilar, heart and pulmonary artery regions [Fig. 6]. This aligns with the physiological hyperdynamic state induced by anemia, where increased cardiac output and pulmonary blood flow redistribution may alter vascular prominence and radiodensity in the hilar area.

Several factors may influence the accuracy of detecting anemia. In our subanalysis, better image quality was associated with higher AUC values, as seen in the PA view and computed radiography. Additionally, the overall predictive accuracy tends to decrease with age. However, females under 55 years old were an exception, likely due to regular blood loss from menstruation—a common cause of anemia in women that involves a distinct pathological mechanism compared to other chronic causes [29]. Our finding also highlights that chronic structural and cardiovascular alterations are significant contributors to the model's performance. Younger individuals typically present with fewer of these chronic structural alterations on imaging, which may limit the DLM's ability to extract secondary predictive cues, thereby leading to decreased performance compared to older populations with more pronounced cardiopulmonary findings. Furthermore, subgroups heart failure-related cardiac remodeling, which affects left ventricular systolic and diastolic function [30], potentially obscuring the aortic ring and other possible

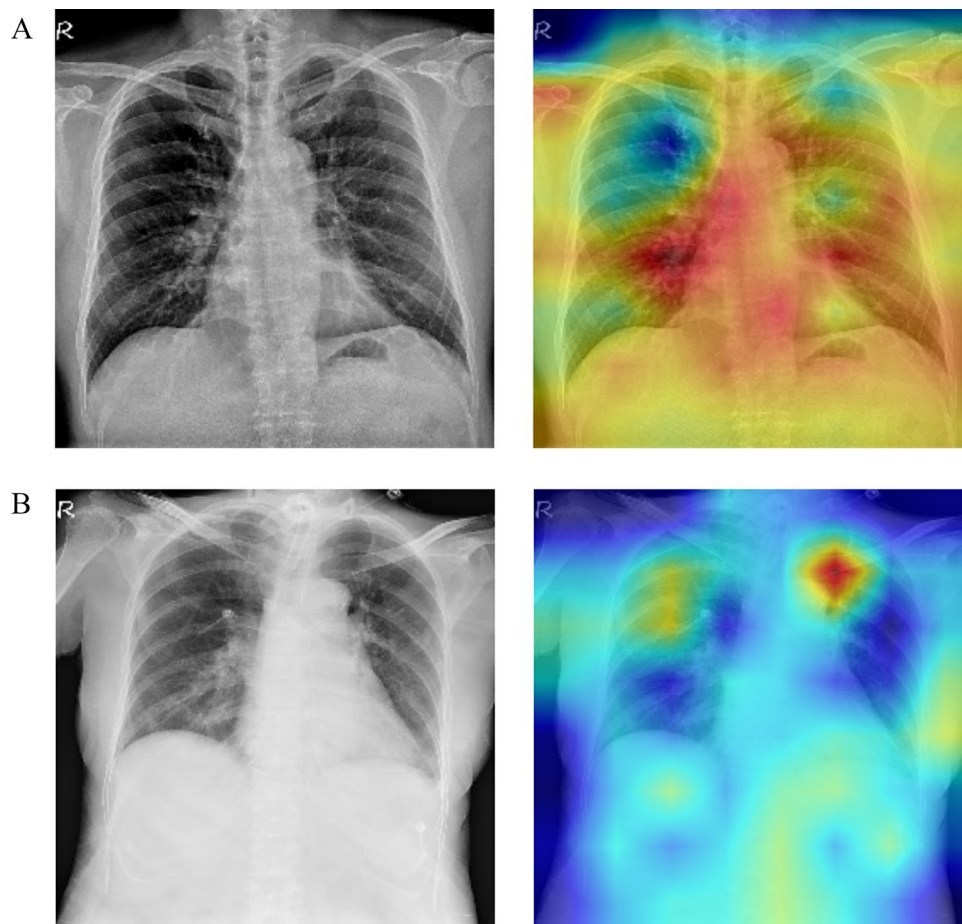


**Fig. 6** Feature visualization of the model for anemia detection. Heart, Hilar and pulmonary artery regions. Prominent activation is observed in the heart, hilar and pulmonary arteries, reflecting changes in blood flow redistribution and radiodensity caused by an anemia-induced hyperdynamic state. The visual observations are exploratory and intended solely for hypothesis generation. The details of feature visualization are showed on Supplementary methods

anemia sign on imaging. Nonetheless, the overall AUC remained relatively stable across the different subgroups.

Several key predictive features were identified. According to previous data, anemia is associated with various clinical scenarios such as hemorrhage, gastrointestinal bleeding, infections, and inflammation [4]. In our study, a history of CKD showed the highest predictive importance. Several factors contributed to this outcome. First, although anemia is associated with many factors such as age, underlying disease history, various interventions, and imaging characteristics (e.g., consolidation change, pneumonia, pulmonary edema), patients with a history of CKD (37.4%) had a 6.8-fold higher prevalence of anemia compared to those without CKD (5.5%) [Supplementary table 1]. The prevalence of anemia in CKD reported in the United States—stage 3: 17.4%, stage 4: 50.3%, stage 5: 53.4%—corresponds with our findings [31]. Several mechanisms contribute to anemia in CKD, including reduced erythropoietin (EPO) production, iron deficiency, chronic inflammation, reduced red blood cell lifespan due to uremia, and blood loss from dialysis [32]. The second most important factor is age. It is well-established that anemia can be caused by various physiological factors, with age being a significant contributor [4]. The third key factor is the AP view, which is often associated with patients too weak to

complete a PA view. These factors are strongly associated with anemia detection and contribute to the FRAMEWORK's predictive capabilities. To evaluate potential acquisition bias from positioning, we performed a stratified analysis at Fig. 3. A slight disparity in discriminative capacity between PA and AP views was observed only in the internal validation set, whereas this performance gap was fully mitigated and stable in the external validation set. This consistency demonstrates that the model genuinely relies on features beyond simple positional proxies. Furthermore, to assess the potential impact of device-related features, we performed a sensitivity analysis by excluding cases with visible medical devices such as nasogastric tubes and Port-A catheters. After removal, model performance decreased only slightly (0.845 to 0.813 (−3.2%) in the internal validation set and from 0.852 to 0.817 (−3.5%) in the external validation set), suggesting that the model's predictions were not primarily driven by the presence of medical devices.

Establishing direct causality between CXR features and anemia is challenging, as comorbidities like heart failure or CKD can independently drive radiographic changes. To address this, we performed a subgroup analysis excluding patients with major cardiopulmonary diseases. Notably, the model's performance—particularly sensitivity in the external set—remained stable. This suggests that while comorbidities influence the imaging landscape, the DLM effectively identifies radiographic patterns associated with anemia independent of overt structural pathology.

To the best of our knowledge, this is the first study to use a CXR-based deep learning-enabled model to predict anemia. Previous studies have explored the use of CXR-based models for various clinical predictions, but none have focused on anemia detection. [11] Several DLMs have attempted to detect anemia through the use of EKG, facial images, and images of the palpebral conjunctiva [10, 33, 34]. However, these methods face several limitations, such as the difficulty of obtaining images compared to CXR, lack of a standardized clinical protocol, the need for additional patient consent, and the absence of a large database and comprehensive external validation. Our results align with studies that demonstrate the efficacy of integrating patient data with imaging techniques to improve predictive accuracy. We also analyzed the performance of our FRAMEWORK, and the results showed that the model outperformed predictions based solely on routine clinical data, especially when combined with patient data. This represents a significant advancement in predictive modeling for anemia compared to traditional methods.

The CXR-based DLM is designed primarily as an opportunistic screening and clinical decision support tool within routine clinical workflows. Because CXR examinations are frequently performed in emergency departments, outpatient clinics, and during hospitalization, the model can leverage these already-acquired images to estimate the probability of anemia without requiring additional tests. In practice, the system can automatically analyze CXR images and generate a risk score for anemia. When the predicted probability exceeds a predefined threshold, the system may prompt clinicians to consider hemoglobin testing or closer clinical evaluation. In this way, the DLM functions as a screening aid rather than a diagnostic replacement, helping identify patients who may benefit from further laboratory assessment.

One of the key advantages of this approach is that it utilizes existing imaging data without introducing additional procedures, allowing clinicians to obtain preliminary information about anemia risk without increasing workload. This opportunistic use of routinely acquired CXR images may help reduce missed or delayed identification of

clinically significant anemia in patients who have not recently undergone laboratory testing.

However, the operational consequences of this strategy require careful evaluation in healthier subpopulations. As shown in our subgroup analysis excluding major comorbidities (Table 2B), the model maintained acceptable discrimination (AUC 0.81; external sensitivity 0.75), but the PPV fell to 0.12–0.16, meaning approximately 85% of positive flags are false positives. Consequently, the high negative predictive values (0.97–0.98) strongly support the tool's utility within a 'rule-out' framework to safely exclude non-anemic patients. Conversely, clinicians using it as a 'rule-in' tool must remain cognizant of this low PPV and the subsequent burden of unnecessary confirmatory laboratory testing.

Although primarily intended for hospital workflows, the model holds potential for secondary applications in resource-limited settings where laboratory access is restricted. However, prospective, multinational validation is required before such clinical deployment can be recommended.

### Limitation

Several limitations were found in our work. First, our research is hospital-based and includes only data from the ER, IPD, and OPD, which may limit the applicability of the model to healthy individuals. However, to address this, we conducted a robust subgroup analysis by excluding patients with major cardiopulmonary comorbidities. Even after these exclusions, we maintained a substantial sample size ( $n = 51,565$  for internal set and  $n = 29,847$  for external set), and the model's performance remained consistently high with a stable AUC. This suggests that the DLM remains effective even in relatively healthy individuals without overt structural lung or heart disease, supporting its potential for broader clinical use. Secondly, the inherent interpretability of the DLM remains a challenge [35]. Unlike typical radiologic findings, anemia does not present with direct, identifiable visual signs on a chest radiograph. Consequently, the model's predictive capability likely relies on recognizing secondary imaging features or structural variations associated with underlying chronic conditions, which are inherently correlated with anemia. This entanglement makes it exceptionally difficult to definitively disentangle whether the DLM is detecting anemia-specific pathophysiology or merely capturing proxies of chronic systemic diseases. To probe this, we employed feature visualizations, which offered exploratory visual clues by highlighting recurring regions of attention like the hilar vessels, though these heatmap observations remain strictly descriptive. To further systematically unravel this black box, we conducted feature importance and device-exclusion sensitivity analyses to objectively evaluate the contribution of known clinical variables. Given that the primary utility of this model is opportunistic screening rather than diagnosis, prioritizing robust predictive performance and generalizability across diverse settings remains paramount. These comprehensive validation efforts suggest that despite its partial opacity, the DLM provides substantial clinical value by effectively identifying high-risk patients who require subsequent diagnostic blood tests. Third, the DLM was developed using a binary classification approach for hemoglobin levels, without predicting between mild and severe anemia. However, in clinical practice, distinguishing between mild and severe anemia is important. This limitation arises because CXR cannot reflect real-time hemoglobin levels, especially in acute cases like

GI bleeding. Blood tests are prioritized for real-time assessment, while CXR is considered only after blood transfusion and stabilization. Timing discrepancies between CXR and blood tests further complicate the model's performance in acute settings. We plan to pursue this direction in future research. Besides, the net-benefit advantage demonstrated within the 10–60% threshold range is conditional on the current cohort prevalence (~9%) and would expectedly shrink in a lower-prevalence community screening setting. Future prospective studies are warranted to validate its clinical utility across populations. Finally, in our study, there were baseline differences between the internal and external validation sets, particularly in the distribution of imaging characteristics [Supplementary table 2]. We hypothesize that this may be related to the hospital characteristics. For example, in the external validation set, most patients were from the OPD, and the inpatient ward only had respiratory care and nursing home patients without a general ward. Thus, our model may not be directly applicable to foreign data or datasets from countries with different disease distributions, such as general healthy communities, and further validation with more diverse databases is required. However, the higher AUC in the external validation suggests that overfitting is not an issue, demonstrating that our FRAMEWORK maintains a consistent level of accuracy across different datasets with varying distributions.

## Conclusion

Our study presents a novel clinical tool that leverages deep learning to predict moderate-to-severe anemia ( $Hb \leq 10 \text{ g/dL}$ ) from chest images. This approach can facilitate quicker decision-making and additional interventions in clinical settings. With further research, we aim to refine the model, making it more continuous and closely aligned with real-time clinical conditions, thereby demonstrating its robustness and utility. Ultimately, this model has the potential to enhance patient care by enabling faster anemia detection and supporting better clinical outcomes.

## Supplementary information

The online version contains supplementary material available at <https://doi.org/10.1186/s13040-026-00588-4>.

Supplementary Material 1

## Acknowledgements

The authors would like to thank all colleagues and doctors who contributed to this study. Thanks for Ministry of National Defense-Medical Affairs Bureau, and National Science and Technology Council of the Republic of China, who supported this research.

## Author contributions

Hao-Chun Liao designed the research study and wrote the initial draft of the manuscript. Shih-Hua Lin contributed to data interpretation. Dung-Jang Tsai performed the data analysis. Shan-Yueh Chang assisted in data interpretation. Chin Lin contributed to study design, data analysis, and interpretation. Wen-Hui Fang critically revised the manuscript for important intellectual content. All authors read and approved the final version of the manuscript.

## Funding

The authors also thank for Wen-Hui Fang, who contributed to funding project preparation. This study was supported by funding from the Ministry of National Defense-Medical Affairs Bureau (MND-MAB-C02-115007 to W.H. Fang), and National Science and Technology Council of the Republic of China, Taiwan (NSTC 115-2314-B-016-037-MY3 to W.H. Fang), NSTC 115-2321-B-016-003 (Awarded to Shih-Hua Lin).

## Data availability

The clinical imaging and electronic health record data underlying this study were collected at Tri-Service General Hospital under IRB approval C202305019 and contain protected health information. Due to institutional policy and the Personal Data Protection Act (Taiwan), the individual-level data cannot be made publicly available. De-identified data may be shared upon reasonable request and subject to IRB approval and a data-use agreement; requests can be

directed to the corresponding author, who will coordinate with the Tri-Service General Hospital IRB. All aggregate results supporting the findings are available in the main text and Supplementary Information.

## Declarations

### Ethics approval and consent to participate

Ethics approval and consent to participate. In accordance with the spirit of the Helsinki Declaration and relevant domestic laws and regulations, the study was approved by the Tri-Service General Hospital Institutional Review Board (TSGH IRB) (IRB NO. C202305019). Written informed consent was waived by the TSGH IRB due to the retrospective design, minimal risk, and analysis of de-identified data. Generative AI tools (ChatGPT, OpenAI) were used to assist in improving language clarity and organization under direct author supervision. All content, data interpretation, and conclusions were conceived and verified by the authors.

### Consent for publication

Written informed consent was obtained from all participants at the time of clinical care, which included explicit notification that de-identified clinical data and images may be used for research and publication purposes. All data used in this study were fully de-identified, and no information enabling individual identification is included in the manuscript.

### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>Department of Internal Medicine, Tri-Service General Hospital, and School of Medicine, National Defense Medical University, Taipei, Taiwan, Republic of China

<sup>2</sup>Department of Family and Community Medicine, Tri-Service General Hospital, and School of Medicine, National Defense Medical University, Taipei, Taiwan, Republic of China

<sup>3</sup>Medical Technology Education Center, School of Medicine, National Defense Medical University, Taipei, Taiwan, Republic of China

<sup>4</sup>Department of Artificial Intelligence and Internet of Things, Tri-Service General Hospital, National Defense Medical University, Taipei, Taiwan, Republic of China

<sup>5</sup>Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Tri-Service General Hospital, and School of Medicine, National Defense Medical University, Taipei, Taiwan, Republic of China

<sup>6</sup>Division of Nephrology, Department of Internal Medicine, Tri-Service General Hospital, Taipei, Taiwan, Republic of China

<sup>7</sup>Department of Education and Research, Taipei City Hospital, Taipei, Taiwan, Republic of China

Received: 6 January 2026 / Accepted: 15 July 2026

Published online: 23 July 2026

## References

1. Garcia-Casal MN, Dary O, Jefferds ME, Pasricha S-R. Diagnosing anemia: challenges selecting methods, addressing underlying causes, and implementing actions at the public health level. *Ann New Y Acad Sci*. 2023;1524(1):37–50. <https://doi.org/10.1111/nyas.14996>.
2. Organization WH. Guideline on haemoglobin cutoffs to define anaemia in individuals and populations. n.d.; Available from: <https://www.who.int/publications/i/item/9789240088542>.
3. Patel KV. Epidemiology of anemia in older adults. *Semin Hematol*. 2008;45(4):210–17. <https://doi.org/10.1053/j.seminhema.2008.06.006>.
4. Turner J, PM BMA [Updated 2023 Aug 8]. in: StatPearls [Internet]. 2024 Jan. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK499994/>. Treasure Island (FL): StatPearls Publishing.
5. Tyburski EA, Gillespie SE, Stoy WA, Mannino RG, Weiss AJ, Siu AF, et al. Disposable platform provides visual and color-based point-of-care anemia self-testing. *J Clin Invest*. 2014;124(10):4387–94. <https://doi.org/10.1172/JCI76666>.
6. Metivier F, Marchais SJ, Guerin AP, Pannier B, London GM. Pathophysiology of anaemia: focus on the heart and blood vessels. *Nephrol Dial Transpl*. 2000;15(suppl\_3):14–18. <https://doi.org/10.1093/oxfordjournals.ndt.a027970>.
7. Müller R, Steffen HM, Brunner R, Saric J, Pollok M, Baldamus CA, et al. Changes in the alpha adrenergic system and increase in blood pressure with recombinant human erythropoietin (rHuepo) therapy for renal anemia. *Clin Invest Med*. 1991;14(6):614–22.
8. Quekel LGBA, Kessels AGH, Goei R, van Engelshoven JMA. Detection of lung cancer on the chest radiograph: a study on observer performance. *Eur J Radiol*. 2001;39(2):111–16. [https://doi.org/10.1016/S0720-048X\(01\)00301-1](https://doi.org/10.1016/S0720-048X(01)00301-1).
9. Abbasi B, Seyed Hosseini M, Moodi Ghalibaf A, Akhavan R, Emadzadeh M, Bolvardi E. Evaluating anemia on non-contrast thoracic computed tomography. *Sci Rep*. 2022;12(1):21380. <https://doi.org/10.1038/s41598-022-24265-8>.
10. Kwon J-M, Cho Y, Jeon K-H, Cho S, Kim K-H, Baek SD, et al. A deep learning algorithm to detect anaemia with ECGs: a retrospective, multicentre study. *Lancet Digit Health*. 2020;2(7):e358–67. [https://doi.org/10.1016/S2589-7500\(20\)30108-4](https://doi.org/10.1016/S2589-7500(20)30108-4).
11. Çalli E, Sogancioglu E, van Ginneken B, van Leeuwen KG, Murphy K. Deep learning for chest X-ray analysis: a survey. *Med Image Anal*. 2021;72:102125. <https://doi.org/10.1016/j.media.2021.102125>.
12. Liao HC, Lin C, Wang C-H, Fang W-H. The deep learning algorithm estimates chest radiograph-based sex and age as independent risk factors for future cardiovascular outcomes. *Digit Health*. 2023;9:20552076231191055. <https://doi.org/10.1177/20552076231191055>.
13. Raghunath S, Ulloa Cerna AE, Jing L, VanMaanen DP, Stough J, Hartzel DN, et al. Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. *Nat Med*. 2020;26(6):886–91. <https://doi.org/10.1038/s41591-020-0870-z>.

14. 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(5):428–36. <https://doi.org/10.1001/jamacardio.2019.0640>.
15. Natanson C, Applefeld WN, Klein HG. Hemoglobin-based transfusion strategies for cardiovascular and other diseases: restrictive, liberal, or neither? *Blood*. 2024;144(20):2075–82. <https://doi.org/10.1182/blood.2024025927>.
16. Chen KC, Kuo M, Lee C-H, Liao H-C, Tsai D-J, Lin S-A, et al. A pretraining approach for small-sample training employing radiographs (PASTER): a multimodal transformer trained by chest radiography and free-text reports. *J Med Syst*. 2025;49(1):120. <https://doi.org/10.1007/s10916-025-02263-3>.
17. Radford A, et al., Marina M, Tong Z, editors. Learning transferable visual models from natural language supervision. Proceedings of the 38th International Conference on Machine Learning. 2021. 8748–63, PMLR: Proceedings of Machine Learning Research.
18. Zhang Y, et al. Contrastive learning of medical visual representations from paired images and text. Machine learning for healthcare conference. 2022.PMLR.
19. Zhang S, Xu Y, Usuyama N, Xu H, Bagga J, Tinn R, et al. A multimodal biomedical foundation model trained from fifteen million image–text pairs. *Nejm Ai*. 2025;2(1):Aloa2400640. <https://doi.org/10.1056/Aloa2400640>.
20. Wang Z, et al. Medclip: contrastive learning from unpaired medical images and text. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. 2022.
21. Tiu E, Talus E, Patel P, Langlotz CP, Ng AY, Rajpurkar P. Expert-level detection of pathologies from unannotated chest X-ray images via self-supervised learning. *Nat Biomed Eng*. 2022;6(12):1399–406. <https://doi.org/10.1038/s41551-022-00936-9>.
22. Chetoui M, Akhloufi MA. Explainable vision transformers and radiomics for COVID-19 detection in chest X-rays. *JCM*. 2022;11(11):3013. <https://doi.org/10.3390/jcm11113013>.
23. Kamel EM, Rizzo E, Duchosal MA, Duran R, Goncalves-Matoso V, Schnyder P, et al. Radiological profile of anemia on unenhanced MDCT of the thorax. *Eur Radiol*. 2008;18(9):1863–68. <https://doi.org/10.1007/s00330-008-0950-9>.
24. Zopfs D, Rinneburger M, Pinto dos Santos D, Reimer RP, Laukamp KR, Maintz D, et al. Evaluating anemia using contrast-enhanced spectral detector CT of the chest in a large cohort of 522 patients. *Eur Radiol*. 2021;31(6):4350–57. <https://doi.org/10.1007/s00330-020-07497-y>.
25. Iuga AI, Pennig L, Caldeira LL, Maintz D, Hickethier T, Doerner J. Prediction of anemia on enhanced computed tomography of the thorax using virtual non-contrast reconstructions. *Med (Baltim)*. 2021;100(48):e28014. <https://doi.org/10.1097/MD.00000000000028014>.
26. Oski FA, Marshall BE, Cohen PJ, Sugerman HJ, Miller LD. Exercise with anemia. *Ann Intern Med*. 1971;74(1):44–46. <https://doi.org/10.7326/0003-4819-74-1-44>.
27. Anand IS, Chandrashekar Y, Wander GS, Chawla LS. Endothelium-derived relaxing factor is important in mediating the high output state in chronic severe anemia. *J Am Coll Cardiol*. 1995;25(6):1402–07. [https://doi.org/10.1016/0735-1097\(95\)00007-Q](https://doi.org/10.1016/0735-1097(95)00007-Q).
28. Beard JL, Tobin BW, Smith SM. Effects of iron repletion and correction of anemia on norepinephrine turnover and thyroid metabolism in iron deficiency. *Exp Biol Med*. 1990;193(4):306–12. <https://doi.org/10.3181/00379727-193-43040>.
29. Shah SA, Soomro U, Ali O, Tariq Y, Waleed MS, Guntipalli P, et al. The prevalence of anemia in working women. *Cureus*. 2023;15(8):e44104. <https://doi.org/10.7759/cureus.44104>.
30. Azevedo PS, Polegato BF, Minicucci MF, Paiva SAR, Zornoff LAM. Cardiac remodeling: concepts, clinical impact, pathophysiological mechanisms and pharmacologic treatment. *Arq Bras Cardiol*. 2016;106(1):62–69. <https://doi.org/10.5935/abc.20160005>.
31. Stauffer ME, Fan T. Prevalence of anemia in chronic kidney disease in the United States. *PLoS One*. 2014;9(1):e84943. <https://doi.org/10.1371/journal.pone.0084943>.
32. Executive summary. *Am J Kidney Dis*. 2006;47:S11–15. <https://doi.org/10.1053/j.ajkd.2006.03.010>.
33. Zhang A, Lou J, Pan Z, Luo J, Zhang X, Zhang H, et al. Prediction of anemia using facial images and deep learning technology in the emergency department. *Front Public Health*. 2022;10:964385. <https://doi.org/10.3389/fpubh.2022.964385>.
34. Dimauro G, et al. Detecting clinical signs of anaemia from digital images of the palpebral conjunctiva. Vol. 7. IEEE Access; 2019. p. 113488–98.
35. Asare JW, Appiahene P, Donkoh ET. Detection of anaemia using medical images: a comparative study of machine learning algorithms – a systematic literature review. *Inf Med Unlocked*. 2023;40:101283. <https://doi.org/10.1016/j.imu.2023.101283>.

## Publisher's Note

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