Biomarker-Based Pretraining for Chagas Disease Screening in Electrocardiograms

Elias Stenhede[1,2] **Arian Ranjbar**[1,2]

[1]Medical Technology & E-Health, Akershus University Hospital, Norway [2]Faculty of Medicine, University of Oslo, Norway

Introduction

Chagas screening with ECGs is promising, but available datasets contain scarce and noisy labels, limiting traditional supervised learning. We address this issue by introducing a biomarker-based pretraining strategy:

1. Pretrain an ECG feature extractor to predict biomarker levels.
2. Fine-tune on Brazilian cohorts for Chagas detection.
3. Use an ensemble of models for robust generalization.

The idea is to leverage abundant biomarker measurements to learn transferable ECG representations.

Methods

MIMIC-IV is collected in the USA and used for pretraining. Only Brazilian datasets are used for finetuning.

- **Pretraining:** MIMIC-IV-ECG (523,275 ECGs, 102,511 patients) paired with 11 selected blood biomarkers when available.
- **Fine-tuning:** CODE15% (345,799 ECGs, 233,770 patients) and SaMi-Trop (1,631 ECGs, 1,631 patients), paired with weak and strong chagas labels.

Preprocessing

All ECGs are resampled to 400 Hz, randomly cropped into 2-second segments, and z-normalized per lead, in both training and inference.

Pretraining Task

Network trained to predict biomarker 11 values as percentile bins (100 classes per biomarker). Bin-smoothing regularization applied to stabilize learning under label sparsity. The following 11 biomarkers are used: Albumin, Calcium (Total), Creatinine, Hematocrit, Hemoglobin, INR (PT), NT-proBNP, Potassium, Red Blood Cells, Troponin T, and Urea Nitrogen. The maximum time difference between the test and ECG is 24 hours.

Fine-tuning for Chagas Disease

Network initialized from pretraining; final layer replaced with a binary classifier. 5-fold cross-validation with early stopping on the best epoch.

Model architecture

InceptionTime 1D-CNN with 2x extra initial strided convolutions and GELU activations to increase receptive field. In inference, an ensemble of 5 models; predictions averaged across ECG segments and models.

Parameter	Pretraining	Finetuning
Batch size	64	128
Optimizers	Muon + Adam	Muon + Adam
Learning rate	0.0037	0.001
Loss function	Cross-entropy	Binary cross-entropy
Model parameters	624,576	483,904

Results

Biomarker-based pretraining

The model learns to predict probability distributions of biomarker values using a single ECG as input. For some ECGs, the model outputs a relatively uniform distribution, indicating uncertainty. When evaluated with top-10 accuracy, all tests get scores between 18% and 0.21%.

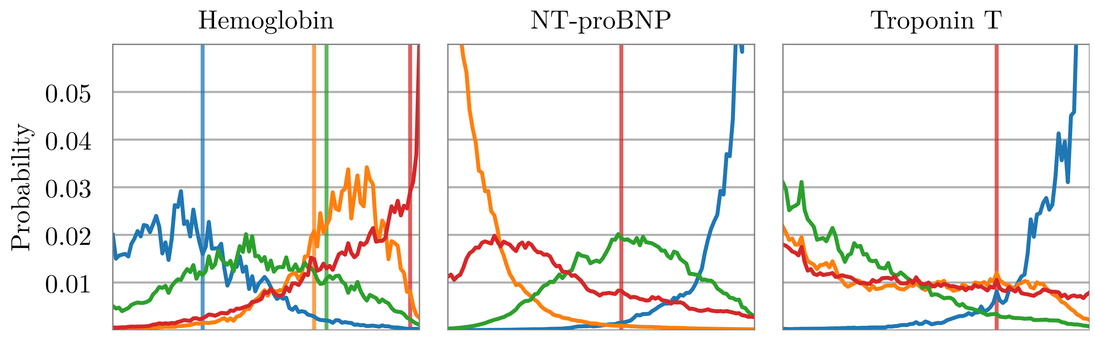

Figure 1: Predicted probability distributions over blood test percentiles for four ECGs taken from four different patients (one colour per patient). The vertical lines represent the actual values for the same patients, measured within 24 hours of ECG.

Chagas finetuning

The fine-tuned ensemble was evaluated on the hidden validation set and evaluated with the official challenge metric. The model assigned 41.2% of the Chagas-positive patients to the top 5% of predictions.

Team	Validation score
Ahus AIM	0.412

Mean cross-validated results on the development were 0.439 for the same score, indicating only a slight performance drop when evaluating the model out of domain. Increasing model size resulted in better cross-validated results but lower scores on the validation set.

Conclusions

- Biomarker-based pretraining enables ECG networks to learn clinically relevant representations without relying solely on scarce disease labels.
- The ensemble achieved strong performance, assigning 41.2% of Chagas-positive patients to the top 5% of predictions.
- This strategy offers a possible path for improving AI-ECG screening of diseases where labeled data is limited.
- More research is warranted, as it is uncertain how this method compares to other label noise mitigation methods.

Contact information and code

elias.stenhede@ahus.no

github.com/Ahus-AIM/physionet-challenge-2025