Photoplethysmography, Foundation Models, Hypertension and Diabetes

George Searle [1] **Stefan van Duijvenboden** [2] **Andrew Tinker** [3] **Patricia Munroe** [3] **Pier Lambiase** [1] **Michele Orini** [1 4]

Abstract

Cardiovascular disease remains the leading cause of death globally, with hypertension and diabetes as two key risk factors. These conditions are frequently underdiagnosed because current diagnostic methods often require in-clinic or invasive procedures, which delay detection until symptoms arise - often too late for optimal intervention. In this work, we focus on photoplethysmography (PPG), a non-invasive signal that can be passively collected using widely available consumer devices such as smartwatches and smartphones. This makes PPG particularly well-suited for remote, continuous health monitoring. We leverage foundation models (PaPaGeI and TabPFN) to extract features from single-heartbeat PPG signals to detect hypertension and diabetes. Using data from 215,000 subjects in the UK Biobank, we demonstrate that these models significantly outperform current state-of-the-art approaches for PPG-based disease detection.

1. Introduction

Cardiovascular disease (CVD) is already the leading cause of death globally, with the majority of these deaths occurring in low and medium income countries (Benziger et al., 2016). The prevalence of CVD is rapidly increasing as populations age and lifestyle factors such as poor diet and lack of exercise become more common, leading to a need for novel preventative strategies to reduce the burden on global healthcare services. Diabetes and hypertension are two major risk factors associated with cardiovascular disease, and the prevalence of these conditions is also rising rapidly, with many cases going undiagnosed (Richardson et al., 2024)(Ogurtsova et al., 2022). Diagnosis using in-clinic tests such as finger-prick tests and blood pressure measurements are powerful tools, but they are limited by the need for patients to seek them out - a potentially low priority task for an asymptomatic individual. Widescale screening for these conditions may help to catch cases earlier, leading to improve outcomes for patients and a reduced burden on healthcare systems. Over the last decade, consumer devices such as smartwatches and smartphones have become ubiquitous, and they are now capable of collecting a wide range of physiological signals, which, when combined with advances in machine learning, have the potential to enable remote, continuous screening for these conditions (Masoumian Hosseini et al., 2023).

Photoplethysmography (PPG) is a non-invasive optical technique that has been universally adopted in healthcare settings for measuring heart rate and SpO_2 for almost a century (Allen, 2007). It is well understood, simple, cheap, and available in almost all consumer smart devices. It can be recorded anywhere, with no medical expertise required. A growing body of research into the use of PPG in cardiovascular healthcare has shown its utility in measuring heart rate variability, arterial stiffness, and other physiological parameters (Charlton et al., 2022). Research into the use of PPG for detecting hypertension and diabetes is both promising and growing in popularity, yet most studies have relatively small sample sizes and require extended measurements(Zanelli et al., 2022)(Elgendi et al., 2019). Here, we use PPG signals covering a *single* heartbeat alongside age, sex and BMI to detect hypertension and diabetes in a large cohort of 215,000 subjects from the UK Biobank. Until now, research has largely focused on morphological feature engineering with classical machine learning techniques such as XGBoost and SVMs, with deep learning methods showing limited discriminatory improvement at the cost of interpretability (Yan et al., 2023)(Oliveira et al., 2023)(Charlton & Marozas, 2022). Recent developments in tabular and time-series foundation models have shown promise in significantly outperforming traditional tree-based methods, and here we study their application in healthcare screening and compare them with more traditional morphological feature based analysis. We use two foundation models, PaPaGeI(Pillai et al., 2025) and TabPFN(Hollmann et al., 2025)(Hollmann et al., 2023)

[1]University College London, London, UK [2]Nuffield Department of Population Health, University of Oxford, Oxford, UK [3]William Harvey Research Institute, Queen Mary University of London, UK [4]School of Biomedical Engineering & Imaging Sciences, King's College London, London, UK. Correspondence to: George Searle <george.searle.22@ucl.ac.uk>.

Proceedings of the 42nd International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025. Copyright 2025 by the author(s).

to extract features from the PPG signals and then classify subjects into those with prevalent hypertension and diabetes and those without.

2. Methods

2.1. Data

Participants from the UK Biobank with a first visit between 2006-2010 were included in this study as part of UK Biobank application number 8256. The UK Biobank study has approval from the North West Multi-Centre Research Ethics Committee (MREC) and all participants provided written informed consent (Sudlow et al., 2015). Baseline characteristics and outcomes were derived from hospital episode statistics. Incidence of hypertension and diabetes were 54.7% and 6.0% respectively. The PPG signals were recorded for each participant using a Pulse Trace device, with each signal being the average of six full heart cycles and given as a time series of 101 samples such that the time axis is the percentage of a single heartbeat. These waveforms were then interpolated using the heart rate to a sampling rate of 250Hz using linear interpolation.
Age, sex, and BMI for each participant were also taken from the Biobank.

2.2. Feature Extraction

2.2.1. Morphological Features

Morphological features were extracted from the PPG signals using the *PyPPG*(Goda et al., 2023) package. This is an open source Python package for extracting common, standardised morphological features from PPG signals based on fiducial points, areas and time intervals. 74 features were extracted from each signal, and these were then reduced to 20 using a combination of correlation and VIF-based feature selection to avoid any effects of multicollinearity. Outliers were removed by taking the 1st and 99th percentiles of each feature, and the data was then standardised using z-scores.

2.2.2. PaPaGeI

PaPaGeI is the first open foundation model specifically designed for PPG signals. It was pre-trained on over 57,000 hours of data, comprising 20 million PPG segments from publicly available datasets. It introduces a novel representation learning approach that leverages domain knowledge of PPG signal morphology across individuals, enabling rich representations of PPG signals. For compatibility with PaPaGeI, our single-waveform signals were padded to meet the 10 second length requirement. We then used the PaPaGeI-S architecture as a feature extractor, deriving 512 embeddings from each signal for each participant. These embeddings were then passed through PCA to reduce the dimensionality of the set to cover 99% of the variance, with these components then used as features for the downstream classification tasks.

2.3. Classification

2.3.1. Dataset Splitting

The dataset was split into training and test sets using a stratified 80/20 split in order to ensure that the proportion of positive and negative cases was preserved in both sets.

2.3.2. TabPFN

TabPFN is a promising tabular foundation model designed for small to medium-sized datasets, limited to 10,000 samples and 500 features (Angelaki et al., 2025). Given that our dataset is much larger than this, we adopted a strategy of bagging and soft voting. We drew K=20 random samples from the training set, each with a size of 10,000 samples, and conditioned one TabPFN instance on each bootstrap. Each model produced class-1 probabilities for every test-set instance, and the final ensemble prediction was made by taking the element-wise average of the 20 probability vectors. Discrimination was assessed by computing the ROC-AUC (in the case of hypertension) and the AUCPR (in the case of diabetes) once over the pooled ensemble probabilities. Uncertainty was estimated via 1000-fold non-parametric bootstrapping of the test set to give 95% confidence intervals.

2.3.3. XGBoost

XGBoost is a widely used tree-based ensemble method that has been shown to perform well on tabular data. We used the XGBoost classifier with weighted classes to account for the class imbalance in the dataset. The model was trained on the morphological features and the PaPaGeI embeddings, with hyperparameters optimised using a random search over a validation set. The best model was then evaluated on the test set, and discrimination was assessed using ROC-AUC for hypertension and AUCPR for diabetes. Uncertainty was estimated using 1000-fold non-parametric bootstrapping of the test set to give 95% confidence intervals.

3. Results

3.1. Hypertension

Hypertension detection results are shown in Table 1 and Figure 1. The results show that using PaPaGeI embeddings over morphological features does not significantly improve performance, but changing the classifier from XGBoost to TabPFN does have a significant impact.

Features	XGB	TabPFN
M0 (Age/Sex/BMI)	0.74 (0.73–0.75)	0.76 (0.75–0.77)
M1 (PyPPG)	0.72 (0.71–0.73)	0.74 (0.73–0.75)
M2 (M1+M0)	0.81 (0.80–0.82)	0.85 (0.83–0.86)*
M3 (PaPaGeI)	0.73 (0.72–0.74)	0.75 (0.74–0.76)
M4 (M3+M0)	0.82 (0.81–0.83)	0.86 (0.85–0.87)*

Table 1. ROC-AUC comparison for Hypertension; * indicates significant difference between XGB and TabPFN.

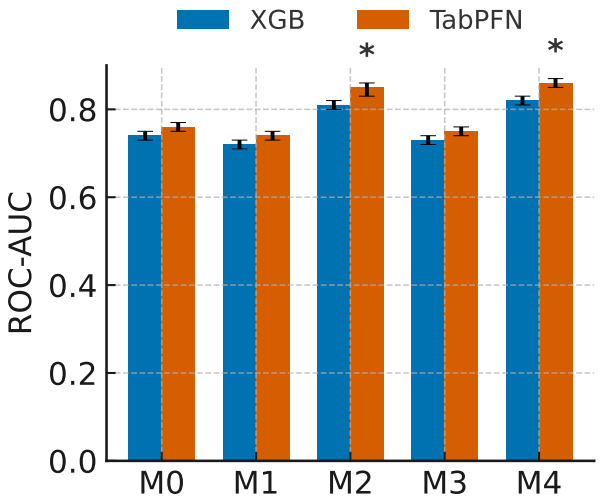

Figure 1. ROC-AUC (95% CIs) for XGB vs. TabPFN across feature sets M0–M4. * indicates significant difference between XGB and TabPFN.

3.2. Diabetes

For diabetes detection, we quote AUCPR due to the imbalance. The results are shown in Table 2 and Figure 2. Again, PaPaGeI embeddings do not significantly improve performance over morphological features, but TabPFN does significantly outperform XGBoost.

Features	XGB	TabPFN
M0 (Age/Sex/BMI)	0.17 (0.16–0.18)	0.17 (0.16–0.18)
M1 (PyPPG)	0.12 (0.11–0.13)	0.14 (0.12–0.15)
M2 (M1+M0)	0.21 (0.19–0.23)	0.26 (0.24–0.28)*
M3 (PaPaGeI)	0.11 (0.10–0.12)	0.15 (0.14–0.16)*
M4 (M3+M0)	0.22 (0.21–0.23)	0.27 (0.26–0.28)*

Table 2. AUC–PR comparison; * indicates significant difference between XGB and TabPFN.

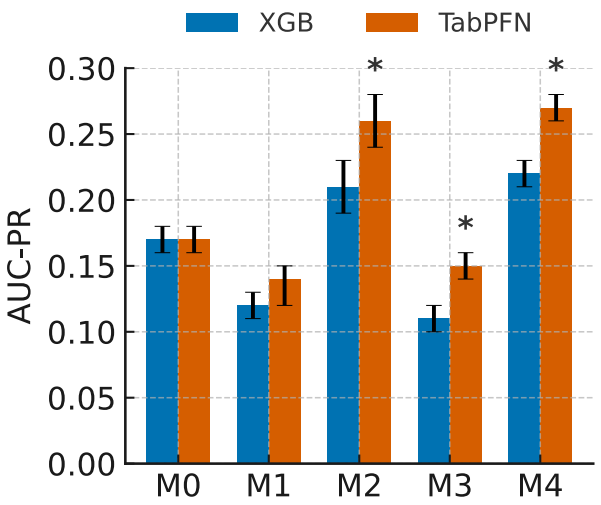

Figure 2. AUC–PR (95% CIs) for XGB vs. TabPFN across feature sets M0–M4. * indicates significant difference between XGB and TabPFN.

4. Discussion

This study explored whether single-beat photoplethysmography (PPG) can detect hypertension and diabetes in a large, population-representative UK Biobank cohort ($N = 215{,}000$), and whether foundation-model techniques offer an advantage over conventional approaches. Interpretable morphological features alone provided strong discrimination, while substituting *PaPaGeI* embeddings yielded no measurable gain. The principal strength of this work is its scale: the UK Biobank spans the entire United Kingdom and mirrors national demographics, supporting external validity. At the same time, interpretability remains critical. The continued effectiveness of hand-crafted PPG morphology, combined with TabPFN's black-box nature, means future deployments should pair TabPFN with explanatory techniques such as feature-deletion analysis or surrogate models.

Several limitations warrant caution. First, our PPG signals cover only a single heartbeat and are resampled to 101 equidistant points, hampering direct comparison with studies that analyse longer or variable-length recordings. Second, PPG accuracy can degrade on darker skin tones (Bermond et al., 2023); we did not stratify by pigmentation, so our results may not generalise to all populations.

In summary, while foundation-model embeddings offered no clear benefit, the TabPFN classifier delivered a modest yet meaningful performance lift at acceptable computational cost. These findings motivate further investigation of TabPFN-based pipelines for large-scale cardiovascular screening, provided that future work addresses interpretabil-

ity and demographic fairness.

5. Conclusion

In this work, we have shown that signal exists in single-heartbeat PPG signals that can be used to detect hypertension and diabetes in a large cohort of 215,000 subjects from the UK Biobank. We have shown that using foundation models such as TabPFN can significantly improve performance over more traditional tree-based methods, but that using PaPaGeI embeddings as features does not significantly improve upon standard morphological features. This suggests that the morphological features extracted from the PPG signals are already capturing the relevant information for these tasks, and that the foundation models are not able to extract any additional signal. This work demonstrates the potential of using PPG signals for remote, continuous health monitoring and screening for hypertension and diabetes, which could lead to earlier detection and improved outcomes for patients.

References

Allen, J. Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, 28(3):R1–39, March 2007. ISSN 0967-3334. doi: 10.1088/0967-3334/28/3/R01.

Angelaki, E., Barmparis, G. D., Fragkiadakis, K., Maragkoudakis, S., Zacharis, E., Plevritaki, A., Kampanieris, E., Kalomoirakis, P., Kassotakis, S., Kochiadakis, G., Tsironis, G. P., and Marketou, M. E. Diagnostic performance of single-lead electrocardiograms for arterial hypertension diagnosis: A machine learning approach. *Journal of Human Hypertension*, 39(1):58–65, January 2025. ISSN 1476-5527. doi: 10.1038/s41371-024-00969-4.

Benziger, C. P., Roth, G. A., and Moran, A. E. The Global Burden of Disease Study and the Preventable Burden of NCD. *Global Heart*, 11(4):393, December 2016. ISSN 2211-8179, 2211-8160. doi: 10.1016/j.gheart.2016.10.024.

Bermond, M., Davies, H. J., Occhipinti, E., Nassibi, A., and Mandic, D. P. Reducing racial bias in SpO2 estimation: The effects of skin pigmentation. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, 2023:1–5, July 2023. ISSN 2694-0604. doi: 10.1109/EMBC40787.2023.10341069.

Charlton, P. H. and Marozas, V. 12 - Wearable photoplethysmography devices. In Allen, J. and Kyriacou, P. (eds.), *Photoplethysmography*, pp. 401–439. Academic Press, January 2022. ISBN 978-0-12-823374-0. doi: 10.1016/B978-0-12-823374-0.00011-6.

Charlton, P. H., Kyriacou, P. A., Mant, J., Marozas, V., Chowienczyk, P., and Alastruey, J. Wearable Photoplethysmography for Cardiovascular Monitoring. *Proceedings of the IEEE*, 110(3):355–381, March 2022. ISSN 1558-2256. doi: 10.1109/JPROC.2022.3149785.

Elgendi, M., Fletcher, R., Liang, Y., Howard, N., Lovell, N. H., Abbott, D., Lim, K., and Ward, R. The use of photoplethysmography for assessing hypertension. *npj Digital Medicine*, 2(1):1–11, June 2019. ISSN 2398-6352. doi: 10.1038/s41746-019-0136-7.

Goda, M. A., Charlton, P. H., and Behar, J. A. pyPPG: A Python toolbox for comprehensive photoplethysmography signal analysis, September 2023.

Hollmann, N., Müller, S., Eggensperger, K., and Hutter, F. TabPFN: A Transformer That Solves Small Tabular Classification Problems in a Second, September 2023.

Hollmann, N., Müller, S., Purucker, L., Krishnakumar, A., Körfer, M., Hoo, S. B., Schirrmeister, R. T., and Hutter, F. Accurate predictions on small data with a tabular foundation model. *Nature*, 637(8045):319–326, January 2025. ISSN 1476-4687. doi: 10.1038/s41586-024-08328-6.

Masoumian Hosseini, M., Masoumian Hosseini, S. T., Qayumi, K., Hosseinzadeh, S., and Sajadi Tabar, S. S. Smartwatches in healthcare medicine: Assistance and monitoring; a scoping review. *BMC Medical Informatics and Decision Making*, 23:248, November 2023. ISSN 1472-6947. doi: 10.1186/s12911-023-02350-w.

Ogurtsova, K., Guariguata, L., Barengo, N. C., Ruiz, P. L.-D., Sacre, J. W., Karuranga, S., Sun, H., Boyko, E. J., and Magliano, D. J. IDF diabetes Atlas: Global estimates of undiagnosed diabetes in adults for 2021. *Diabetes Research and Clinical Practice*, 183:109118, January 2022. ISSN 0168-8227. doi: 10.1016/j.diabres.2021.109118.

Oliveira, F. A. C., Dias, F. M., Toledo, M. A. F., Cardenas, D. A. C., Almeida, D. A., Ribeiro, E., Krieger, J. E., and Gutierrez, M. A. Machine Learning-Based Diabetes Detection Using Photoplethysmography Signal Features, August 2023.

Pillai, A., Spathis, D., Kawsar, F., and Malekzadeh, M. PaPaGei: Open Foundation Models for Optical Physiological Signals, February 2025.

Richardson, L. C., Vaughan, A. S., Wright, J. S., and Coronado, F. Examining the Hypertension Control Cascade in Adults With Uncontrolled Hypertension in the US. *JAMA network open*, 7(9):e2431997, September 2024. ISSN 2574-3805. doi: 10.1001/jamanetworkopen.2024.31997.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., and Collins, R. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Medicine*, 12(3):e1001779, March 2015. ISSN 1549-1277. doi: 10.1371/journal.pmed.1001779.

Yan, L., Wei, M., Hu, S., and Sheng, B. Photoplethysmography Driven Hypertension Identification: A Pilot Study. *Sensors (Basel, Switzerland)*, 23(6), March 2023. doi: 10.3390/s23063359.

Zanelli, S., Ammi, M., Hallab, M., and El Yacoubi, M. A. Diabetes Detection and Management through Photoplethysmographic and Electrocardiographic Signals Analysis: A Systematic Review. *Sensors (Basel, Switzerland)*, 22(13):4890, June 2022. ISSN 1424-8220. doi: 10.3390/s22134890.