Classification of cardiovascular disease from electrocardiogram and photoplethysmogram signal analysis

Principal Investigators

Dr Ha Thi Hai Duong,

Dr Nguyen Viet

Dr Hoang Minh Tu Van

Investigators

Dr Phanh Nguyen Quoc Khanh

Dr Lam Minh Yen

Dr Nguyen Thi Kim Anh

Dr Ly ?

Dr Louise Thwaites, Oxford University Clinical Research Unit

Dr Tingting Zhu, University of Oxford

Sponsor The University of Oxford

Contents

Background

Both cardiovascular disease (CVD) and infectious diseases are leading public health problems in Vietnam, affecting millions of people. Rates of CVD within the Vietnamese population are increasing whilst the infectious diseases burden remains high. For patients with infectious diseases and also pre-existing CVD, outcomes are worse. Providing early and accurate diagnosis of CVD would enable improved treatment and better outcomes. HIV is a particular example where CVD is more frequent and with worse outcomes [1–3].

Currently diagnosing cardiac impairment requires expensive equipment and highly trained staff to perform tests such as echocardiography. Cardiac impairment, however, is also associated with characteristic changes in heart rate and the circulation such as changes in heart rate variability and pulse wave contour. These may be detected using commonly available non-invasive methods such as electrocardiogram (ECG) and photoplethysmogram (PPG). Heart rate variability is the beat-to-beat variation in heart rate that occurs due to autonomic nervous system modulation and has been shown to be altered in cardiovascular disease [4]. Pulse wave contour is an indication of cardiac output and again is altered in cardiac impairment. In our previous work, we have demonstrated that machine learning (ML) methods applied to ECG and PPG can detect these features in patients with tetanus and hand foot and mouth disease [5–7]. These features can be extracted and analysed either 'conventionally', i.e. using recognized features of heart rate variability (see appendix) or with more complex machine learning methodology. Conventional analysis has the advantage that recognized features can be compared with published literature and may provide some indication of disease pathophysiology whereas machine learning analysis is able to incorporate many more complex features not necessarily apparent in conventional analysis. Results from previous work has shown that machine learning methodology was able to provide more accurate discrimination between mild and severe diseases than a conventional statistical model using recognized features [7].

Furthermore, work from our group has shown that machine learning analysis of ECG waveforms can be used to detect the onset of myocardial infarction with sensitivity and specificity of 94.6 [8]. Similarly, AI algorithms applied to ECG's labelled 'normal sinus rhythm' by cardiologists have been shown to predict subsequent atrial fibrillation with 80% sensitivity and specificity [9].

In this project we will investigate the relationship between heart rate variability and CVD in patients with HIV. We will also extend our work, to provide proof of principal that non-invasive sensors could detect CVD in HIV. Potentially this finding would be relevant to other infectious diseases, including the current COVID-19 pandemic. Data will be collected using low-cost wearable sensors, as we aim to develop sustainable and scalable tools suitable for resource-restricted settings. We will assess their diagnostic accuracy benchmarked against clinical standard and evaluate clinical value and acceptability from patient and staff perspectives.

Aims

To carry out a pilot study to measure heart rate variability parameters in people living with HIV to develop machine learning algorithms to detect and classify CVD risk in people living with HIV.

Design

Case control pilot study

Setting

Hospital for Tropical Diseases

Inclusion/Exclusion Criteria

Inclusion criteria:

1. ≥ 18 years old
2. HIV positive receiving outpatient treatment

Exclusion Criteria

1. Failure to give informed consent
2. Contraindications to use of ECG, PPG

Study Procedures

Cases and controls will be selected based on screening for CVD. Patients will be assessed using recognized risk scores (see below) for CVD risk and clinical history of CVD.

Based on initial screening, patients will be invited to enroll in the study aiming for approximately 20 participants with high risk or known CVD and 20 low risk patients, matched for age and sex.

Basic clinical and demographic data will be recorded for each participant, including disease specific details, such as time since diagnosis and treatment.

All participating patients will undergo vital sign measurement using PPG and ECG in addition to echocardiography. ECG and PPG data recorded using standard monitors, or wearables: E-Patch ECG monitor (Delta Electronics, Denmark) and Smartcare pulse oximeter PPG (Smartcare UK) or Shimmer ECG device. These devices have been chosen as they allow direct recording and storage of data and have been used in our previous tetanus, dengue and sepsis studies at Hospital of Tropical Diseases Ho Chi Minh City. All devices are non-invasive and simple to use. Recordings will be of approximately 15-20 minutes duration. If alternative wearable devices are to be used, the local ethics committee will be informed of this.

An experienced cardiologist will review ECG and echocardiograms.

Cardiovascular risk will be quantified using the Framingham risk score with DAD modification. The Framingham risk score is the recommended CVD risk score in Vietnam, and in people living with HIV, a specific DAD modification should be applied which has been specifically validated in HIV infected patients and incorporates some HIV specific variables.

Electronic ECG and PPG data will be downloaded and stored in a secure server for analysis. Paper ECG records will also be stored. Analysis will include extraction of standard heart rate variability parameters as well as novel feature extraction (***see Appendix***). These features will then be used to develop ML algorithms to classify CVD and CVD risk.

Acceptability of wearable devices will be assessed through a questionnaire to participants and clinic staff (***see Appendix***)

Patients' returning for a follow up visit (expected approximately monthly) will have ECG and PPG, repeated.

Sample Size

This study aims to obtain preliminary proof of principal data. A sample size of up to 60 patients has been calculated, enrolling between 20 and 40 patients with high risk of CVD (aiming for this group to include up to 20 patients with abnormal ECG or echocardiogram) and 20 patients at low risk of CVD and with normal ECG and echocardiogram.

The sample size has been estimated to allow evaluation and feasibility of screening and data collection methods and an indication of likely difference in heart rate variability and waveform parameters between groups and based on previous estimates of incidence of CVD in HIV patients with high CVD risk scores [2].

Outcome measures**:**

The primary outcome will be standard time and frequency domain HRV parameters

The secondary outcomes will be will be the accuracy of classification of cardiovascular disease risk and pre-defined ECG and echocardiographic abnormalities accuracy using ML algorithms.

An additional exploratory analysis will be the changes in ECG and PPG characteristics over time.

Participants will be classified as 'no CVD' and 'CVD' based on the results of ECG and echocardiography. Normal ECG will be defined according to recognized critieria and demonstrate no evidence of conduction abnormality, rhythm abnormality or morphology abnormality as reported by a specialist cardiologist. Echocardiographic criteria will include no evidence of structural heart disease or functional impairment, defined using American Society of Echocardiography criteria [10]. High and low risk groups will be classified using the Framingham classification and DAD calculation.

Analysis

Analysis will be performed in collaboration with the Bioengineering Department, University of Oxford and biostatistics department at OUCRU. Time-domain and frequency domain variables will be extracted from ECG and PPG signals and single and multimodal models created.

Standard HRV variables will be described in all patients and subgroups of high and low cardiovascular risk and presence or absence of ECG/echocardiographic abnormality. Changes in inter-individual HRV parameters over time will also be described.

Simple and multimodal variable machine learning algorithms will be developed to classify patients according to cardiovascular risk and ECG/echocardiographic abnormality.

Ethics and consent

Informed consent will be taken by the attending doctors, all of whom will receive specific training in the study and Good Clinical Practice and will be authorised to take consent by the trial principal investigator. These doctors will also assess whether or not the patient has mental capacity to provide informed consent. If the doctor judges that the patient does not have this capacity, they will obtain informed consent from the patient's representative (usually a relative). It will be made completely and unambiguously clear that the patient (or their representative) is free to refuse to participate in all or any aspect of the trial, at any time and for any reason, without incurring any penalty or affecting their treatment.

The informed consent form will be presented to the participants or representatives detailing no less than: the exact nature of the study; the implications and constraints of the protocol; the known side effects, risks involved and alternatives to taking part. Those who refuse consent will be treated as per the best available standard of care and will not have any study related procedures performed.

The patient or their representative must personally sign and date two of the latest approved version of the informed consent form. The study staff will also sign and date the two copies. The patient/ representative will receive one copy.

If the patient/representative is illiterate, a witness who is not a member of the study staff will be present during the informed consent discussion. The informed consent form will be read to the patient/representative in the presence of the witness. If the patient/representative agrees to participate, the form will be signed and dated by the witness.

Risks and Benefits

This study poses no additional risk to the patient - both the Smartcare pulse oximeter and E-Patch/Shimmer sensors have been used in clinical studies in Vietnam without any significant discomfort or adverse events. Patients are may benefit by taking part as they will undergo more detailed screening for CVD and may have CVD diagnosed and treated earlier. Patients in whom CVD is detected will be referred to standard services following hospital procedures.

Participant Confidentiality

The trial staff will ensure that the participants' anonymity is maintained. Participants will be identified only by initials and a participants identification number on any electronic database. All documents will be stored securely and only accessible by trial staff and authorised personnel.

Data Use and Storage

Data collected will be used for the purpose of this study as stated in the protocol and stored for future use in studies not yet conceived within Viet Nam or abroad. Any proposed plans to use data other than for those investigations detailed in this protocol will be submitted to the relevant ethics committees prior to any testing.

The participants will be identified only by a study specific participant number and/or code in any database. The name and any other identifying detail will NOT be included in any study data electronic file. Data will be stored securely and transferred in accordance with data protection principals.

Finance and Insurance

The conduct of this study is funded by the Wellcome Trust and sponsored by the University of Oxford.

This research will be appropriately covered through the University of Oxford's legal liability insurances

Publication Policy

The primary outcome data will be analysed and reported in a publication. The authors (and their respective positions in the author list) will be agreed prior to the start of the study in accordance with the guidelines of the International Committee of Medical Journal Editors.

References

1. Ballocca F, D'Ascenzo F, Gili S, Grosso Marra W, Gaita F. Cardiovascular disease in patients with HIV. *Trends Cardiovasc Med*. 2017;27:558-563. doi:10.5694/MJA16.00650.

2. Friis-Møller N, Thiébaut R, Reiss P, Weber R, Monforte ADA, De Wit S, El-Sadr W, Fontas E, Worm S, Kirk O, Phillips A, Sabin CA, Lundgren JD, Law MG. Predicting the risk of cardiovascular disease in HIV-infected patients: The Data collection on Adverse Effects of Anti-HIV Drugs Study. *Eur J Cardiovasc Prev Rehabil*. 2010;17(5):491-501. doi:10.1097/HJR.0b013e328336a150.

3. D'Ascenzo F, Cerrato E, Biondi-Zoccai G, Moretti C, Omedè P, Sciuto F, Bollati M, Modena MG, Gaita F, Sheiban I. Acute coronary syndromes in human immunodeficiency virus patients: A meta-analysis investigating adverse event rates and the role of antiretroviral therapy. *Eur Heart J*. 2012;33(7):875-880. doi:10.1093/eurheartj/ehr456.

4. Schmidt H, Müller-Werdan U, Hoffmann T, Francis DP, Piepoli MF, Rauchhaus M, Prondzinsky R, Loppnow H, Buerke M, Hoyer D, Werdan K. Autonomic dysfunction predicts mortality in patients with multiple organ dysfunction syndrome of different age groups. *Crit Care Med*. 2005;33(9):1994-2002. doi:10.1097/01.CCM.0000178181.91250.99.

5. Duong H, Tadesse G, Nhat P, Hao N, Duong D, Kien T, Prince J, LV T, Loan H, Pugh C, Chau N, Yen L, Zhu T, Clifton DA, Thwaites CL. Heart Rate Variability as an Indicator of Autonomic Nervous System Dysfunction in Tetanus.

6. Tadesse GA, Zhu T. Personalised Diagnostic Tools for Infectious Diseases. 2017:2017.

7. Tadesse GA, Zhu T, Le Nguyen Thanh N, Hung NT, Duong HTH, Khanh TH, van Quang P, Tran DD, Yen LM, van Doorn R, van Hao N, Prince J, Javed H, Kiyasseh D, van Tan L, Thwaites L, Clifton DA. Severity detection tool for patients with infectious disease. *Healthc Technol Lett*. 2020;7(2):45-50. doi:10.1049/htl.2019.0030.

8. Zhu T. Deep end-to-end classification in heart attack. *AI Med Under Rev*. 2020.

9. Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, Carter RE, Yao X, Rabinstein AA, Erickson BJ, Kapa S, Friedman PA. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet*. 2019;394(10201):861-867. doi:10.1016/S0140-6736(19)31721-0.

10. Lang RM, Badano LP, Mor-avi V, Afilalo J, Armstrong A, Ernande L, Flachskampf FA, Foster E, Goldstein SA, Kuznetsova T, Lancellotti P, Muraru D. Recommendations for Cardiac Chamber Quantification by Echocardiography in Adults : An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr*. 2015;28(1):1-39.e14. doi:10.1016/j.echo.2014.10.003.

11. Nguyen Phuc Thu T, Hernandez AI, Costet N, Patural H, Pichot V, Carrault G et al. Improving methodology in heart rate variability analysis for the premature infants: Impact of the time length. *PLoS One*. 2019;14(8):e0220692. https:// doi.org/10.1371/journal.pone.02.

Appendix

1. Physiological interpretation of standard heart rate variability indices

HRV indices	Interpretation
Time domain	
HR (beats/minute)	Heart rate
Standard deviation of NN intervals - SDNN (ms)	Global measure of R-R interval variability; reflects an estimate of short-term beat-to-beat variability over the entire recording period, giving the overall autonomic modulation regardless of sympathetic or parasympathetic arm.
Square root of the mean squared differences of successive NN intervals - RMSSD (ms)	Short-term beat-to-beat variability and is associated with parasympathetic activity [11]
Frequency domain	
Total power - Ptot (ms^2)	Variance of R-R intervals over the temporal segment
Very low frequency power - VLF (ms^2)	Power in very low frequency range (≤0.04 Hz). Activity is abolished with atropine and therefore likely to represent parasympathetic activity
Low frequency power-LF (ms^2)	Power in low frequency range (0.04 Hz-0.15 Hz) Modulated by the baroreflex and representing sympathetic and parasympathetic activity
LF nu (%)	Low frequency power in normalized units; LF/(Total power –very low frequency)*100
High frequency power-HF (ms^2)	Power in the high frequency range (0.15-0.4 Hz) representing parasympathetic (vagal activity), although can be affected by very high or low respiratory rates
HF nu (%)	HF power in normalized units; HF/(Total power –very low frequency)*100
LF/HF	Ratio LF/HF

2. Example of questionnaire to participants

1. Was the experience wearing the ECG monitor comfortable [Likeart scale 1-5]

2. Was the experience wearing the pulse oximeter comfortable [Likeart scale 1-5]

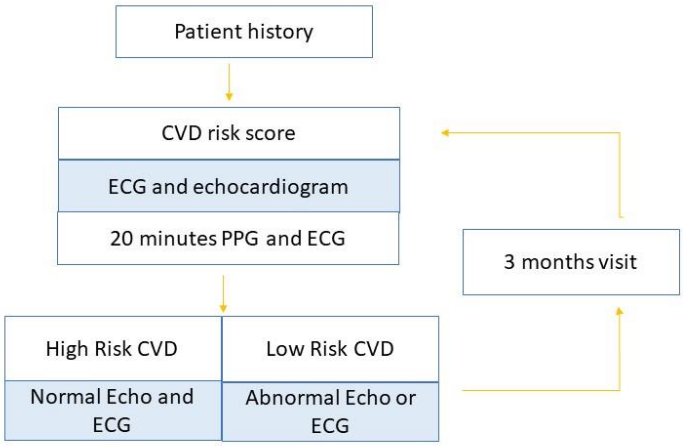
3. Did you find the exercise test difficult? [Likeart scale 1-5]

3. Study timeline

Estimated study timeline is below

	01/211	06/21	12/21
Data collection			
Analysis			
Report/ write up			

4. Flow chart


Patient history
CVD risk score
ECG and echocardiogram
20 minutes PPG and ECG
3 months visit
High Risk CVD
Low Risk CVD
Normal Echo and ECG
Abnormal Echo or ECG