Using consumer wearables for clinical decision making

Dr Peter H. Charlton

University of Cambridge

City, University of London

Undiagnosed disease in UK:

0.4 mill **atrial fibrillation**

1.4 mill **obstructive sleep apnea**

5.5 mill **hypertension**

Public Health England, Atrial fibrillation prevalence estimates in England: Application of recent population estimates of AF in Sweden, 2015

Knauert M et al., Clinical consequences and economic costs of untreated obstructive sleep apnea syndrome, *World J Otorhinolaryngol Head Neck Surg*, 2015

Photo by Luke Chesser on Unsplash


St. Thomas' Hospital
St. Thomas' Hospital

(artistic license – this heart rate is fictional)

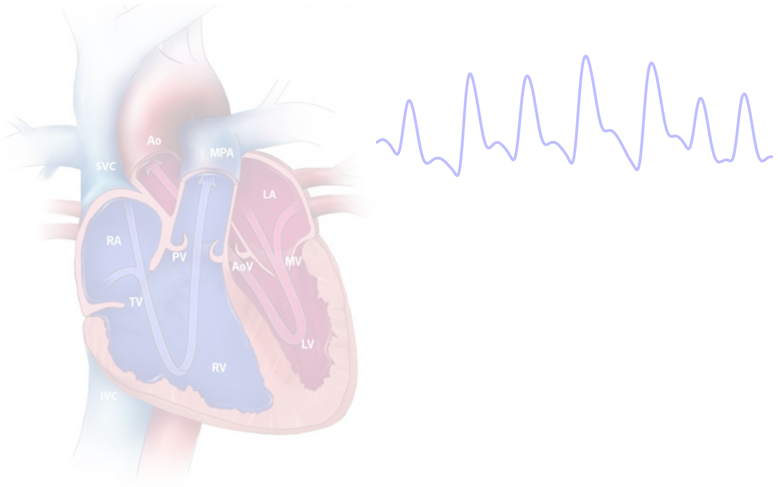

St. Thomas' Hospital
St. Thomas' Hospital


5:42
12
5:00–6:00PM
17°

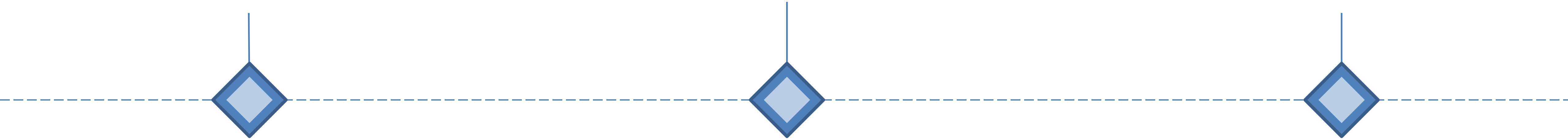
Pexels, https://pixabay.com/photos/man-heartache-chest-pain-hurt-pain-1846050/ (Pixabay Licence)

Consumer Wearables

Clinical Applications

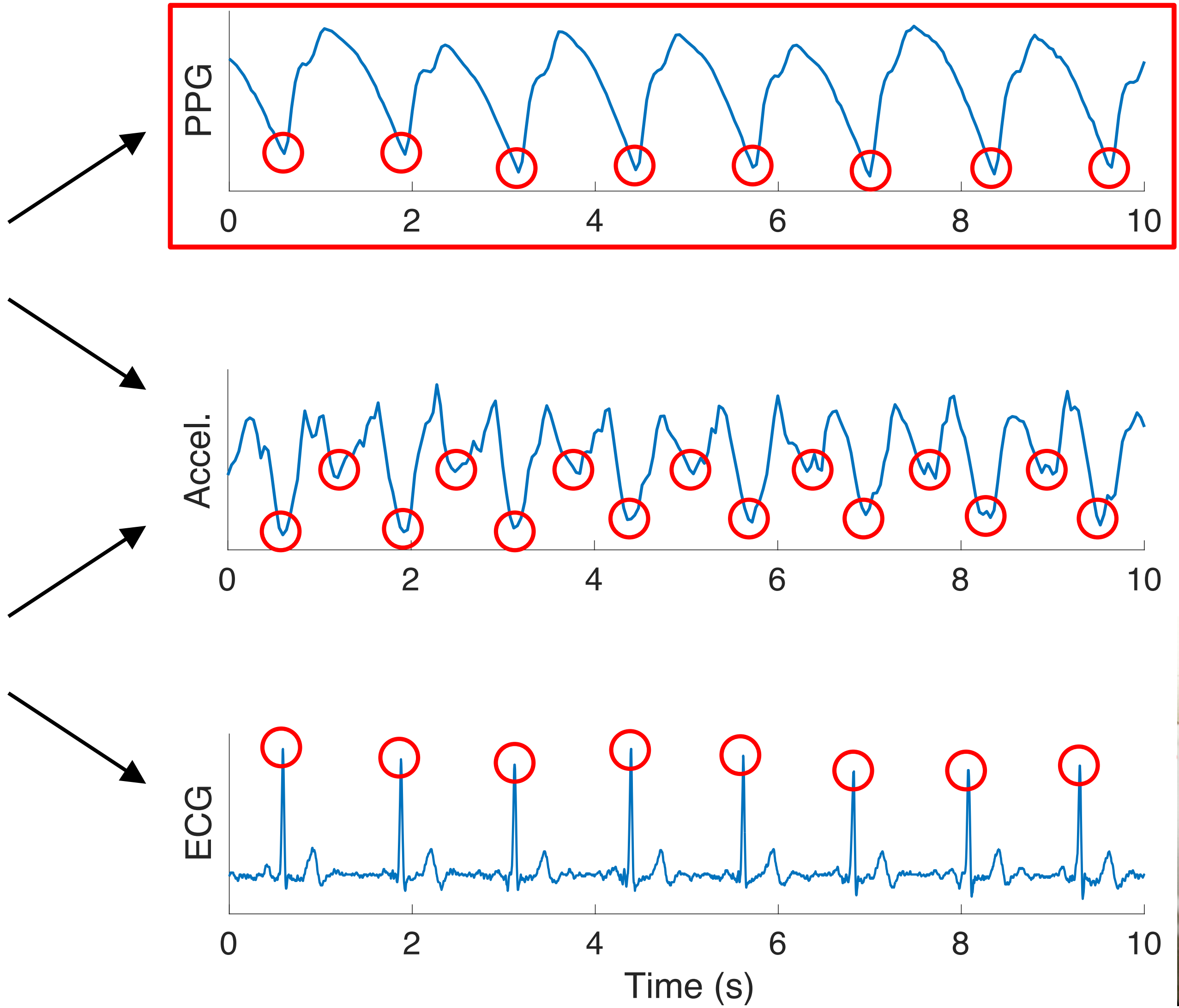


Translation



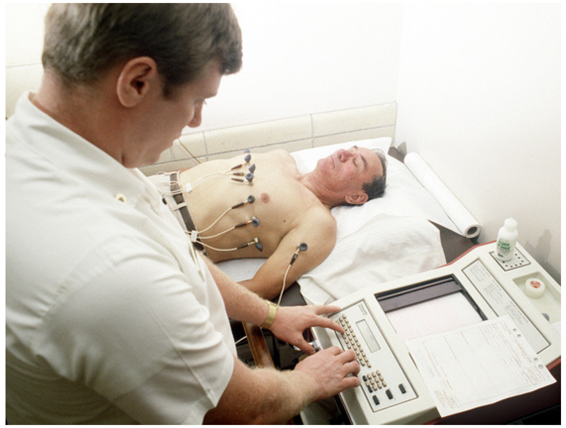
Source: Charlton et al. Individual images: P. Charlton under CC BY 4.0; cropped from image by Marco Verch (CC BY 2.0); cropped image from Passler et al. under CC BY 4.0; cropped from image by GEEK KAZU (CC BY 2.0); cropped from image by Pixels (Pixabay License); cropped from image by Luke Chesser (CC0 1.0).

A **Fitness tracker** which acquires photoplethysmography (PPG) and accelerometry (Accel.) signals



Pulse waves used to:
- estimate heart rate
- identify an irregular pulse

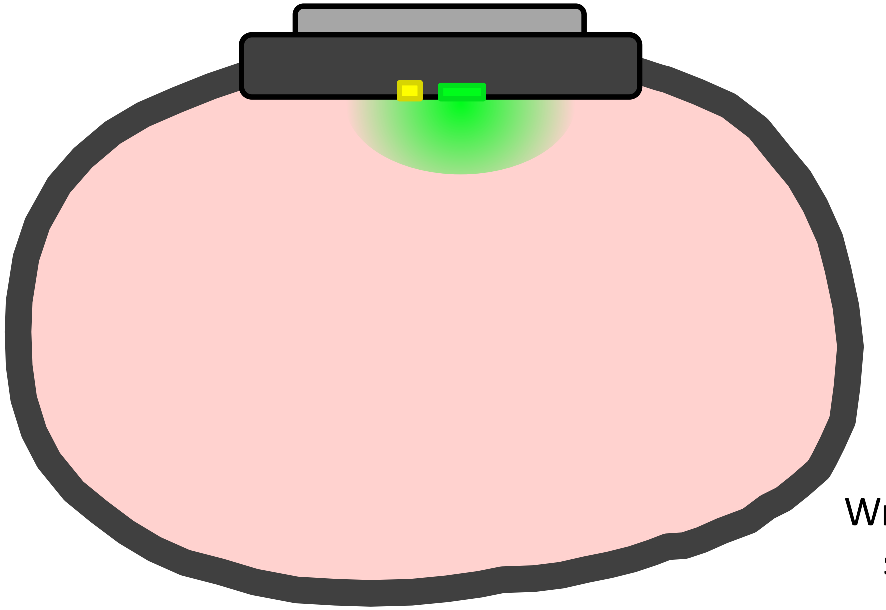
Steps used to:
- estimate step count



A **Smartwatch** which acquires electrocardiography (ECG) and accelerometry (Accel.) signals

The Photoplethysmogram

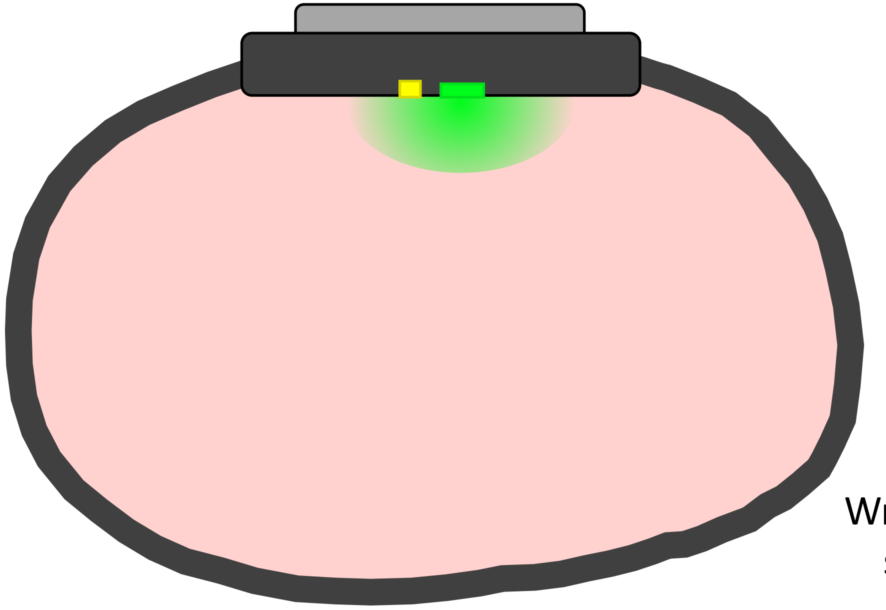
Photoplethysmogram (PPG) Sensor



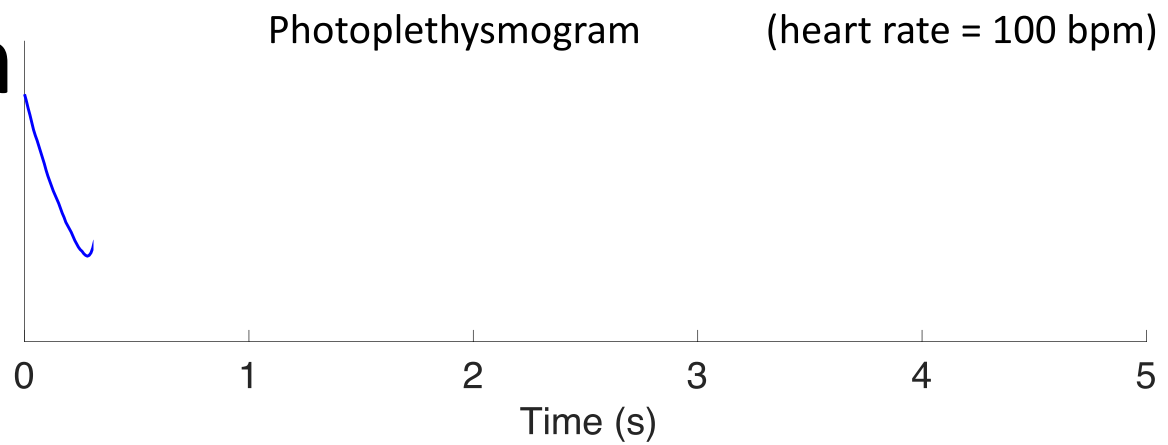
Wrist cross-section

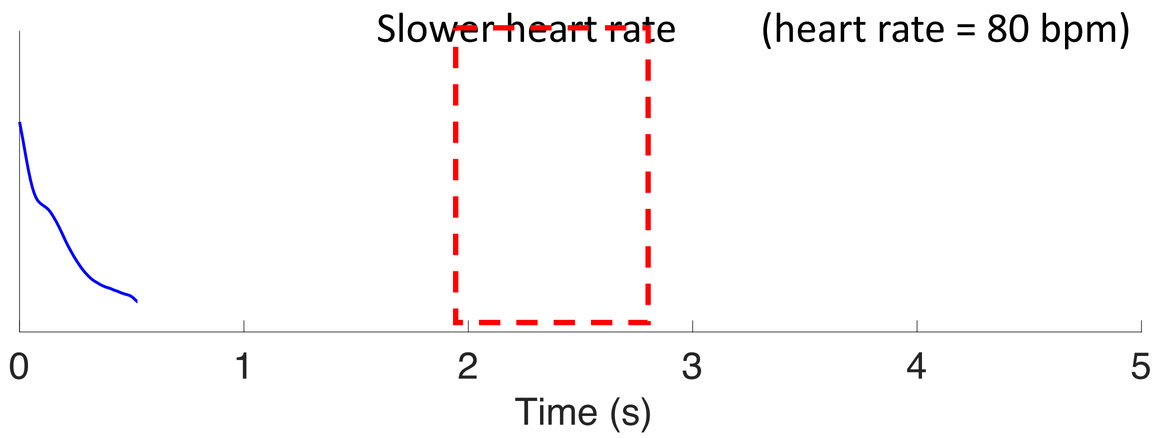
The Photoplethysmogram

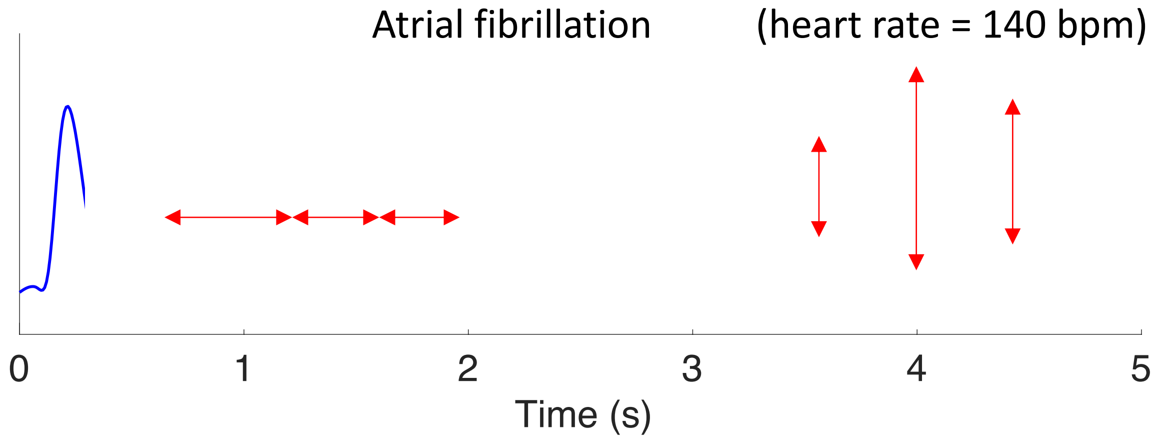
Photoplethysmogram (PPG) Sensor



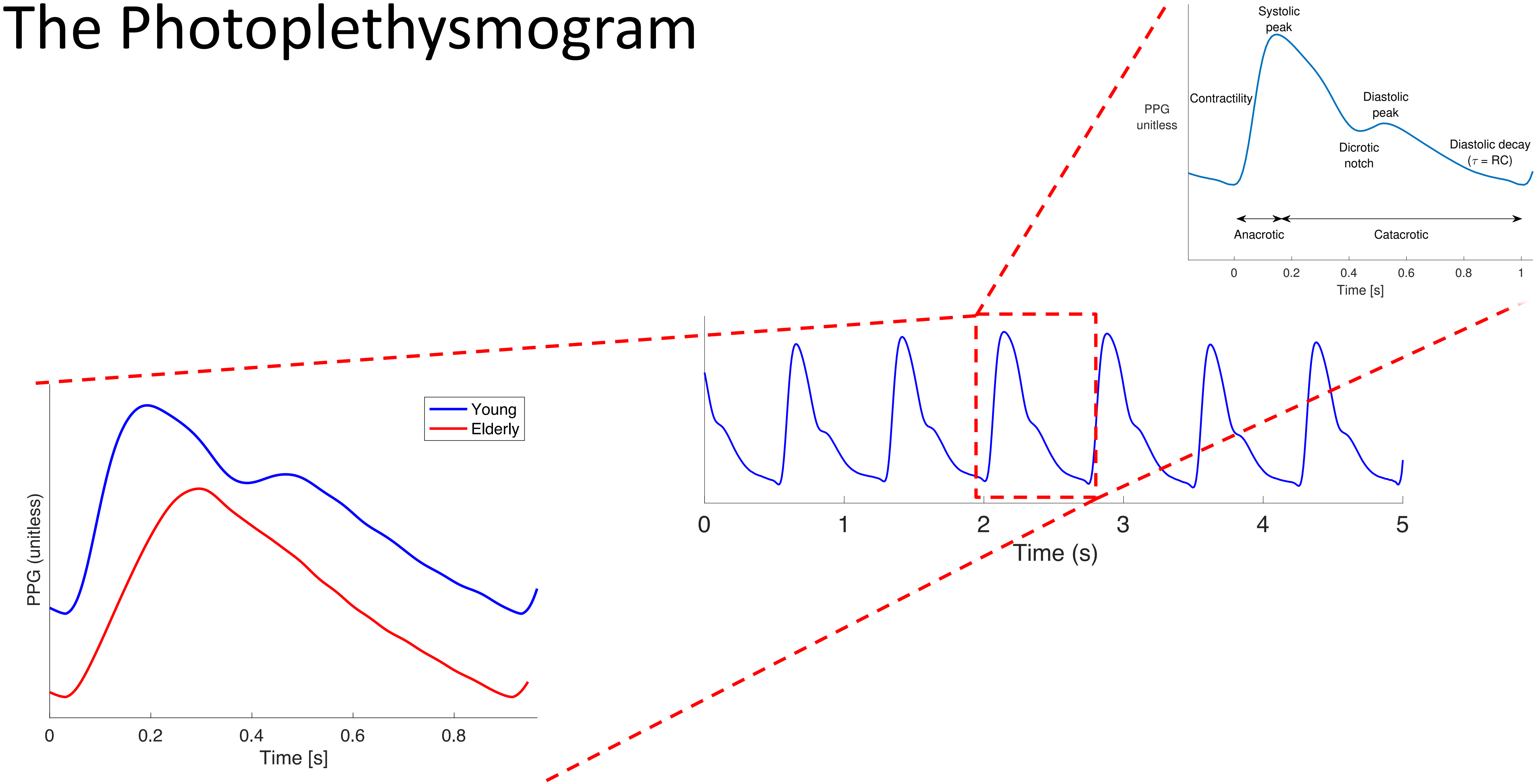
Wrist cross-section







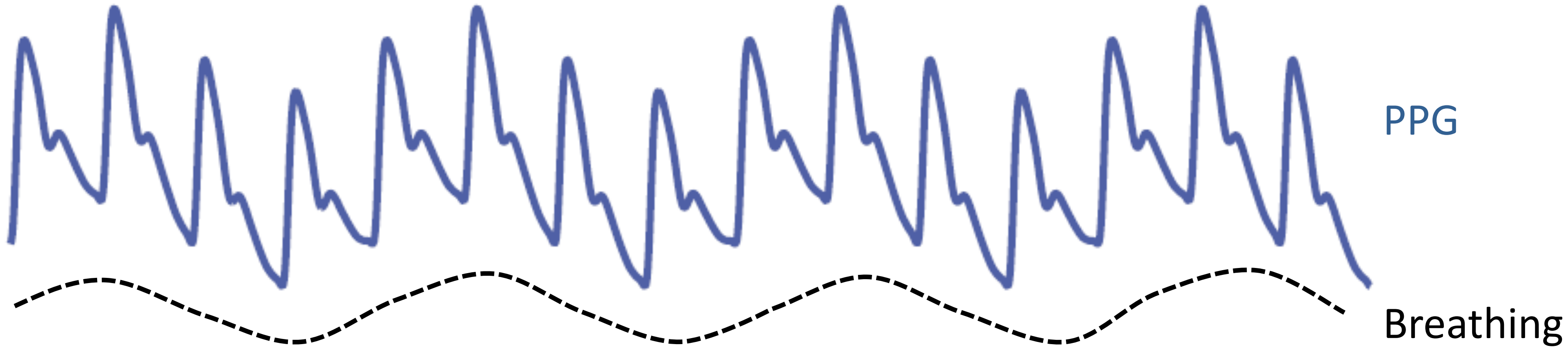
The Photoplethysmogram



Charlton PH *et al.*, Wearable photoplethysmography for cardiovascular monitoring, Proc. IEEE, 2022. https://doi.org/10.1109/JPROC.2022.3149785 (CC BY 4.0)

Peter Charlton, PhD Thesis (CC BY 4.0)

The Photoplethysmogram



Functionality

Vascular age
VO2 max
Arrhythmias
Stress level
Heart rate
Oxygen saturation
Respiratory rate
Sleep assessment
Blood pressure
Energy expenditure

Potential Applications

Infectious diseases
Fitness tracking
Biometric authentication
CV risk prediction
Health insurance
Clinical deterioration
Menstrual cycle
Vascular age
Chronic kidney disease
Mental stress
Pre-eclampsia
Sleep monitoring
Seizure detection
Obstructive sleep apnea
Orthostatic hypotension
Atrial fibrillation
Asthma
Heart failure
COPD
Sepsis

Further Reading

Introduction:

Charlton P.H., **Realising the potential of wearables**, SCOPE, 2021, https://www.ipem.ac.uk/media/5eelohm5/scope-spring-2021-locked.pdf#page=50

Review article:

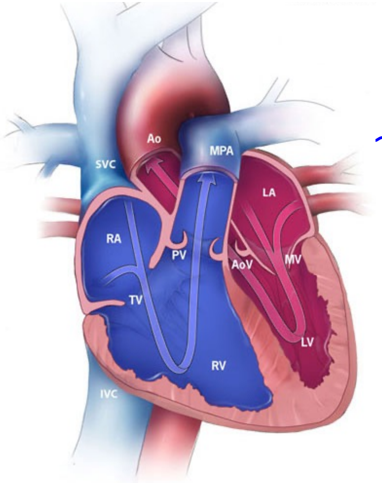
Charlton P.H., Marozas V., *et al.*, **Wearable Photoplethysmography for Cardiovascular Monitoring**, Proc. *IEEE*, 2022, https://doi.org/10.1109/JPROC.2022.3149785

Textbook chapter:

Charlton P.H. and Marozas V., **Wearable photoplethysmography devices**, *Photoplethysmography*, 2021, https://doi.org/10.1016/B978-0-12-823374-0.00011-6

Consumer Wearables

Clinical Applications



Translation

If atrial fibrillation was adequately treated in England:

2,000 Lives saved p.a.

7,000 Strokes prevented

425k Additional diagnoses

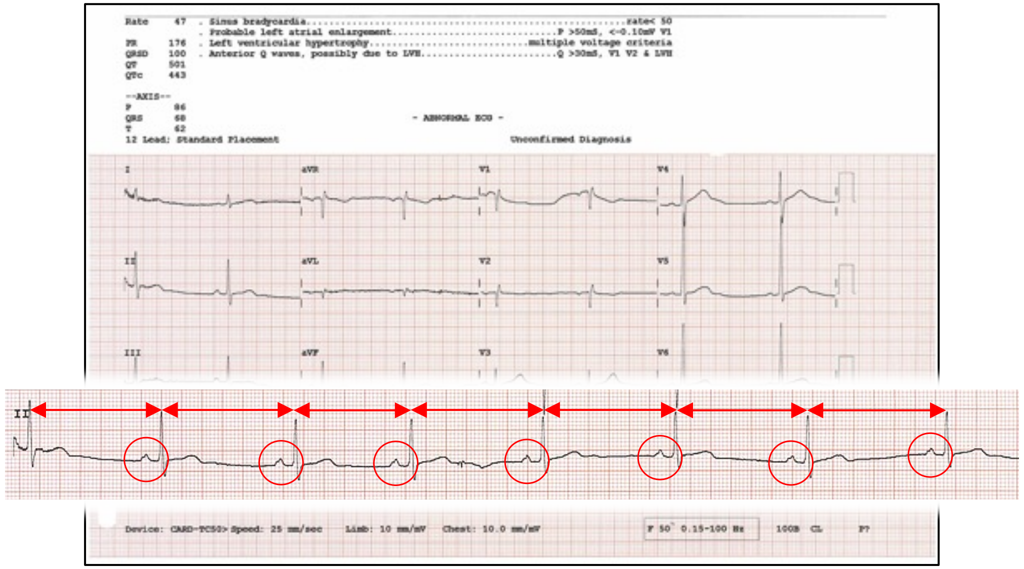
Stroke Association, "State of the Nation," 2017.
Public Health England, "Atrial fibrillation prevalence estimates in England …", 2015.

Pexels, https://pixabay.com/photos/man-heartache-chest-pain-hurt-pain-1846050/ (Pixabay Licence)

Detecting Atrial Fibrillation

Current approach

12-lead ECG:



An ECG is diagnostic of AF if it exhibits:

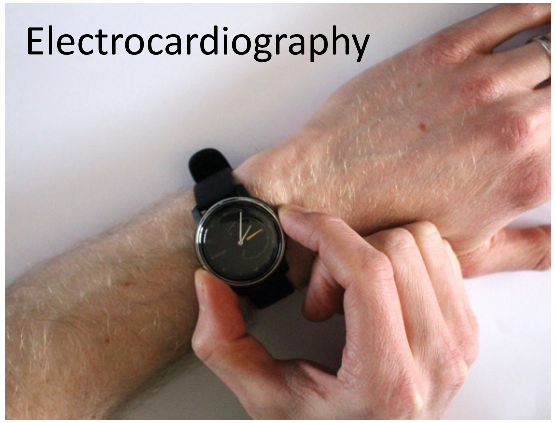
- irregular RR intervals, and
- no discernible repeating P waves

Source: G. Hindricks et al., '2020 ESC Guidelines for the diagnosis and management of atrial fibrillation …'

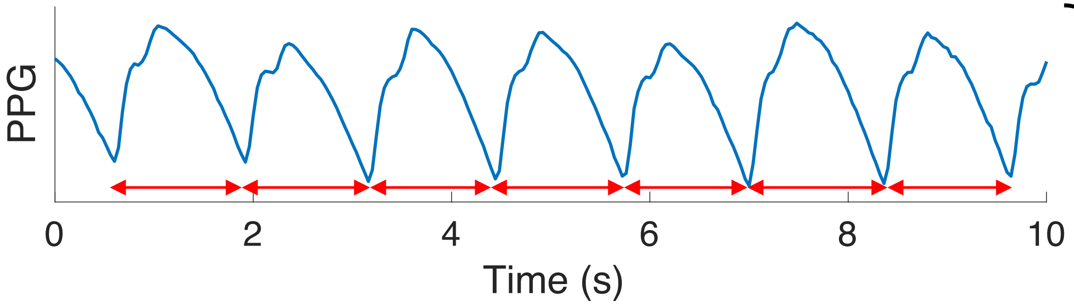
Screening

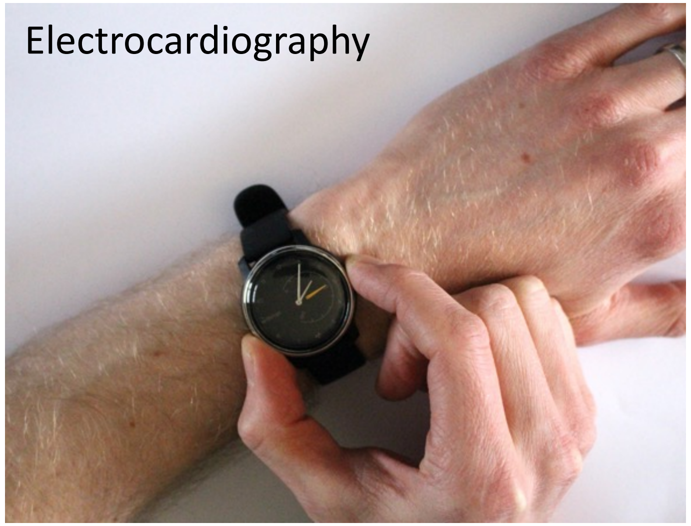
Electrocardiography

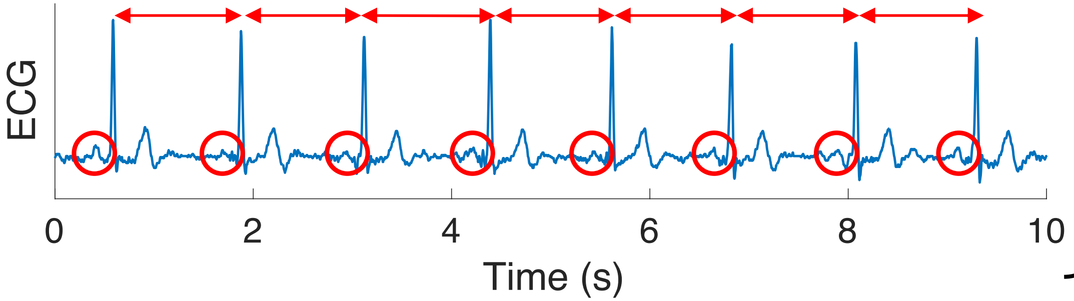
Consumer Wearables



Detecting Atrial Fibrillation



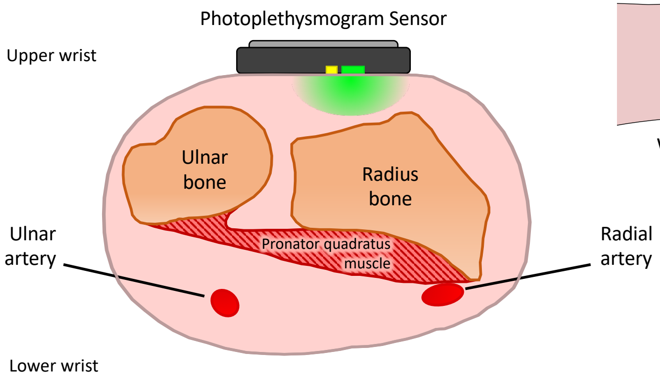


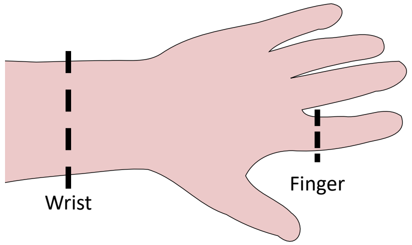


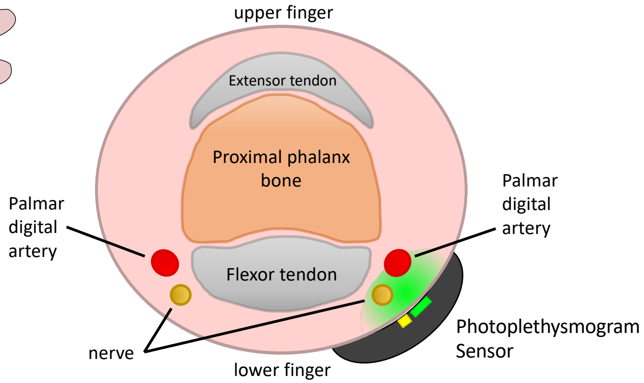
Inter-beat intervals

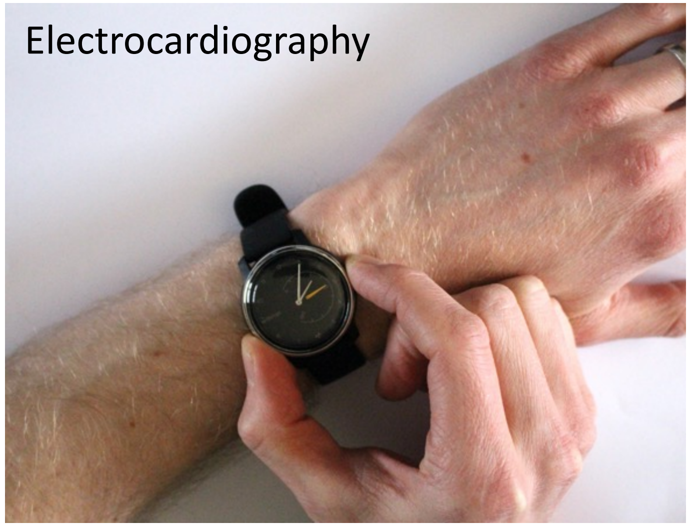
P-waves

Detecting Atrial Fibrillation









Peter Charlton, Wikimedia Commons, https://commons.wikimedia.org/wiki/File:12_lead_ECG.jpg (CC BY 4.0)

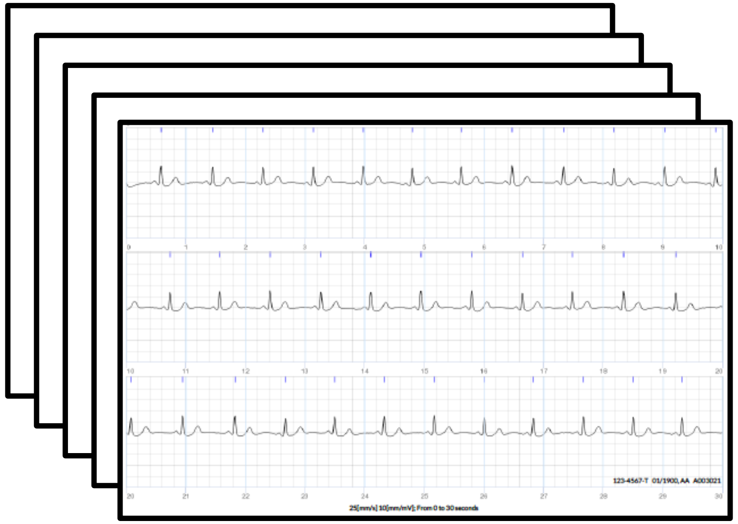
Charlton PH *et al.*, Wearable photoplethysmography for cardiovascular monitoring, Proc. IEEE, 2022. https://doi.org/10.1109/JPROC.2022.3149785 (CC BY 4.0)

Adapted from: H. Gray, Anatomy of the Human Body, W. H. Lewis, Ed., 2nd ed. New York, NY, USA: Lea and Febiger, 1918

Potential Clinical Use Cases

Population-based Screening Programme

Electrocardiography



Requirements:

- High sensitivity to AF

For example:

- Sensitivity of 98% to AF
- Low positive predictive value: 35 ECGs reviewed per diagnosis
- Detect even infrequent AF (present on at least 1 ECG)

Source: Svennberg et al. 2017, https://doi.org/10.1093/europace/euw286

Opportunistic Detection

Photoplethysmography (+ Electrocardiography)

- High positive predictive value

- High positive predictive value of 84%
- (Sensitivity not investigated)
- Only detect fairly persistent AF (present in 5 out of 6)

Source: Perez et al. 2019, https://doi.org/10.1056/NEJMoa1901183

Photo by Luke Chesser on Unsplash

Research Directions

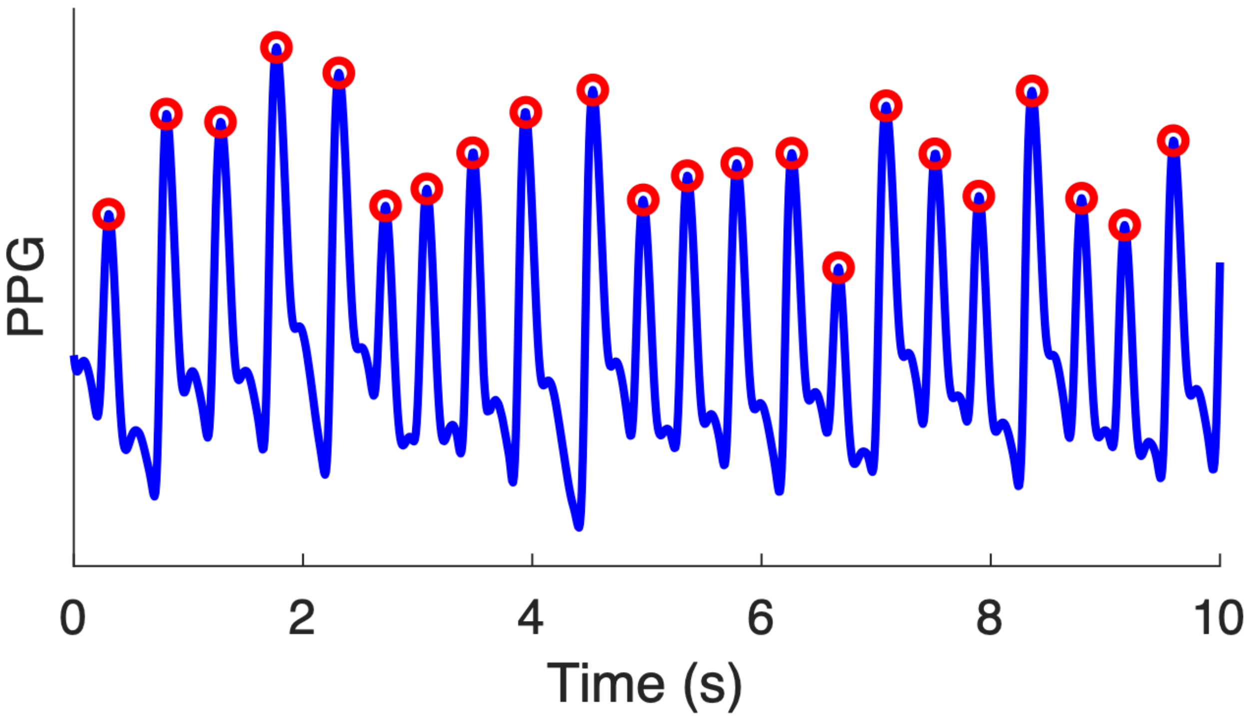
Optimising AF detection algorithms

- Beat detection
- Signal quality assessment
- Detecting AF from inter-beat intervals
- Incorporating P-wave analyses

Reducing clinical workload

- Assessing ECG reviewing workload in real-world screening
- Prioritising ECGs for manual review
- (incorporating optimized AF detection algorithms)

Benchmarking photoplethysmogram beat detectors



Several approaches have been proposed to detect beats:

- Detect **maxima or minima** in the PPG
- Compare weakly and strongly **filtered** PPGs
- Identify line segments indicating systolic **upslopes**
- Detect maximum upslopes using the first **derivative**
- Identify systolic upslopes using the first derivative
- Identify systolic upslopes using a slope sum function
- Identify pulse onsets using a **wavelet transform**
- Analyze the **local maxima scalogram**

It's not clear which performs best.

Aims:

1. Develop a framework with which to design and test PPG beat detectors
2. Assess the performance of PPG beat detectors in different use cases
3. Investigate how their performance is affected by patient demographics and physiology

Benchmarking photoplethysmogram beat detectors

Methods:

- Assemble publicly available datasets containing PPG signals and simultaneous ECG signals (for reference beats)

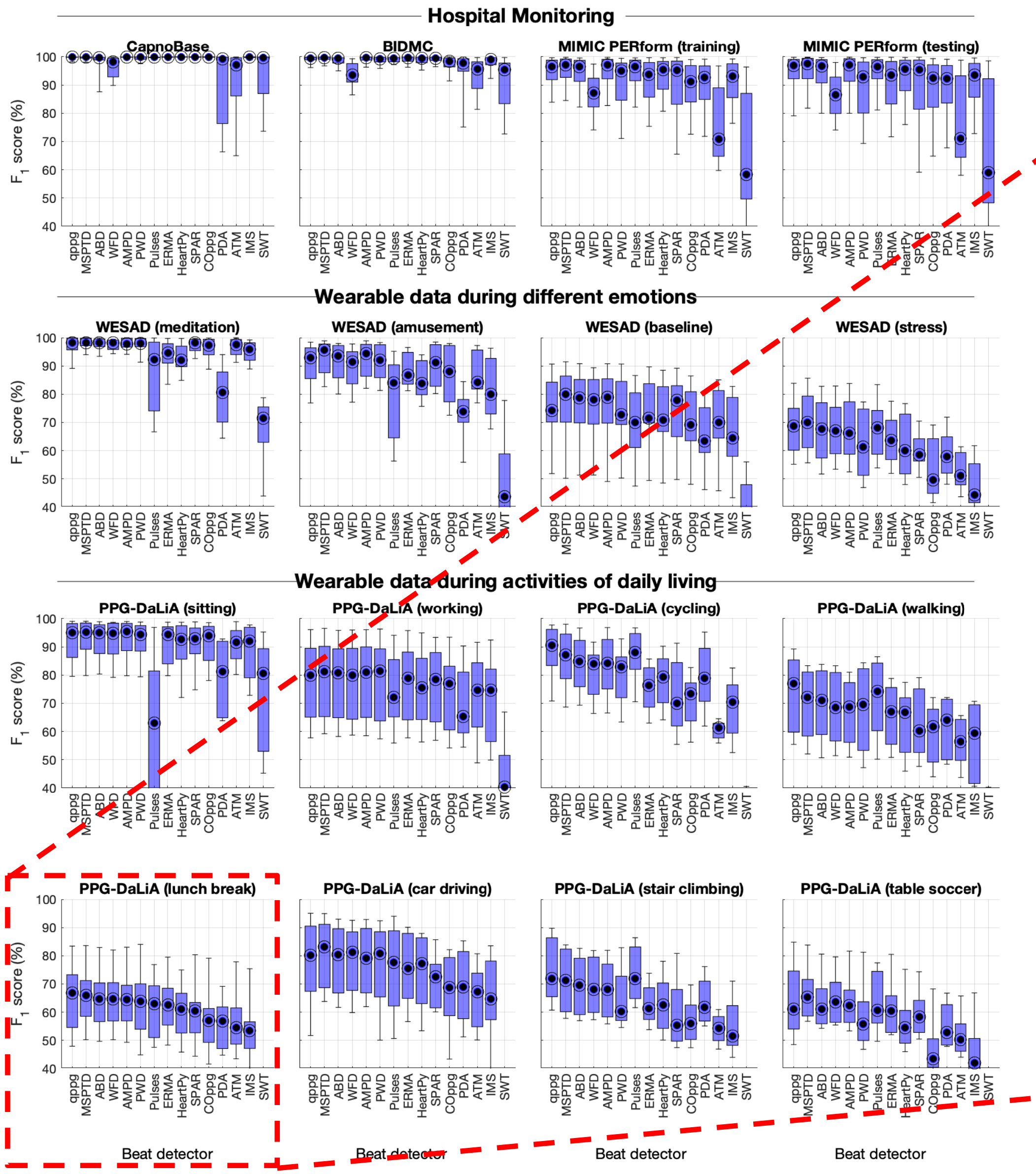
Dataset	Number of Subjects	Description
		Hospital datasets
CapnoBase	42	High-quality recordings from patients undergoing elective surgery and routine anaesthesia.
BIDMC	53	High-quality recordings from critically-ill adults during routine clinical care.
MIMIC PERform Training Dataset	200	Recordings from patients during routine clinical care, who are categorised as either adults or neonates.
MIMIC PERform Testing Dataset	200	Recordings from patients during routine clinical care, who are categorised as either adults or neonates.
MIMIC PERform AF Dataset	35	Recordings from critically-ill adults during routine clinical care, categorised as either AF (atrial fibrillation) or non-AF.
MIMIC PERform Ethnicity Dataset	200	Recordings from critically-ill adults during routine clinical care, who are categorised as either Black or White ethnicity.
		Wearable datasets
WESAD	15	Recordings from volunteers during a laboratory-based protocol designed to induce different emotions.
PPG-DaLiA	15	Recordings from volunteers during a protocol of activities of daily living.

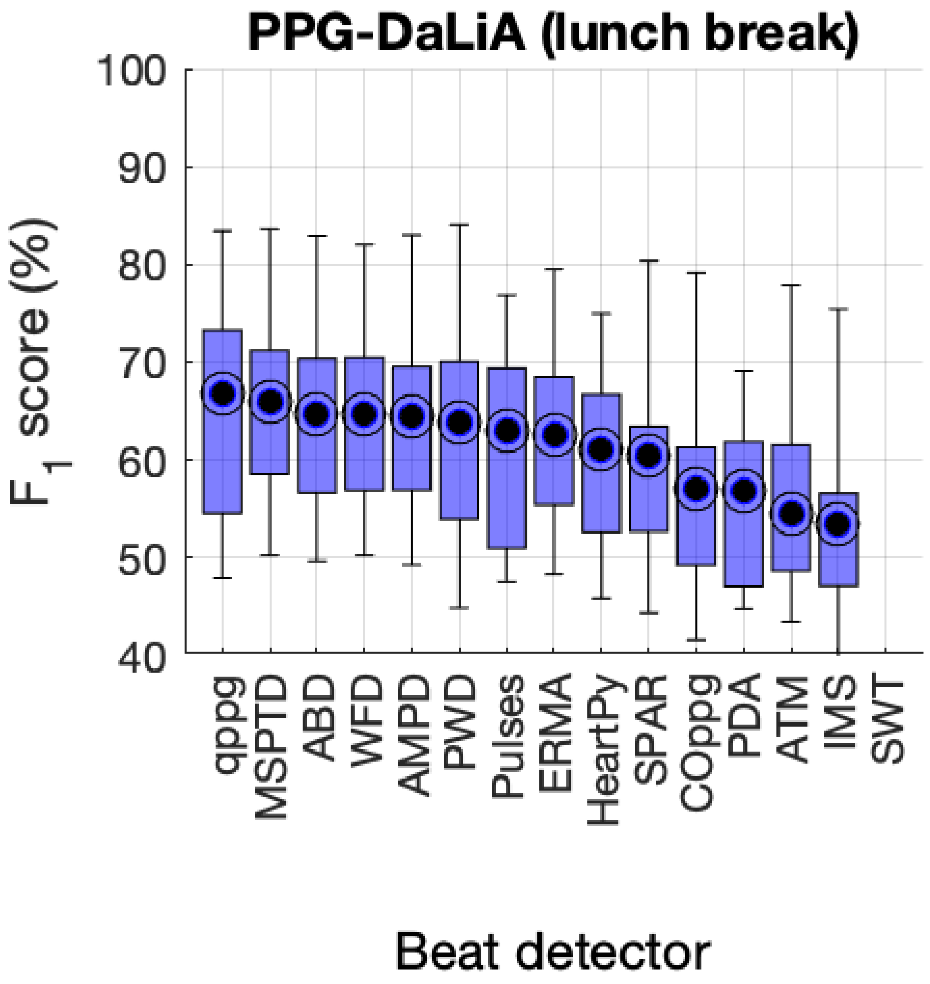
Benchmarking photoplethysmogram beat detectors

Methods:

- Assemble publicly available datasets containing PPG signals and simultaneous ECG signals (for reference beats)
- Detect beats using fifteen open-source PPG beat detection algorithms (either existing or implemented for this study):
  - Filter signals between 0.67 and 8.0 Hz
  - Segment into 20 second windows
  - Exclude windows of (clearly) low signal quality, e.g. flat line, clipped signal
- Detect reference beats from ECG using two separate beat detectors
  - Deem reference beats to be correct if the two beat detectors agreed (within +/- 150 ms)
  - Exclude windows containing disagreements
- Align PPG and ECG derived beats to account for clock synchronization and physiological time differences
- Deem PPG beats to be correct if within +/- 150 ms of ECG beats
- Quantify performance using:
  - F1-score, which is the harmonic mean of:
    - Sensitivity
    - Positive predictive value

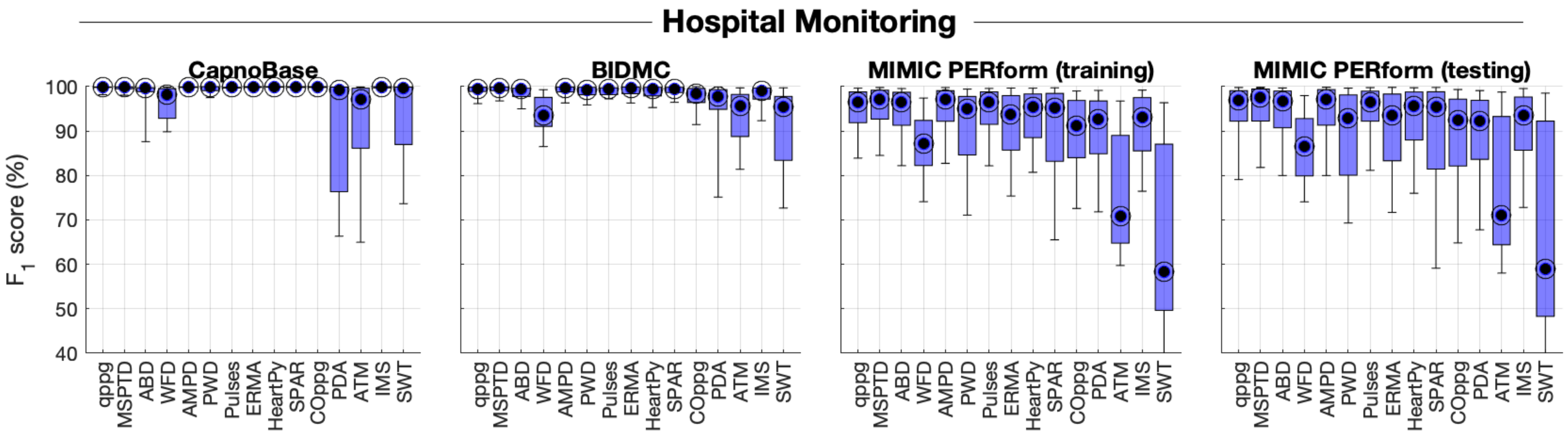
Benchmarking photoplethysmogram beat detectors





Charlton PH *et al.*, Detecting beats in the photoplethysmogram: benchmarking open-source algorithms, Phys Meas., 2022. https://doi.org/10.1088/1361-6579/ac826d

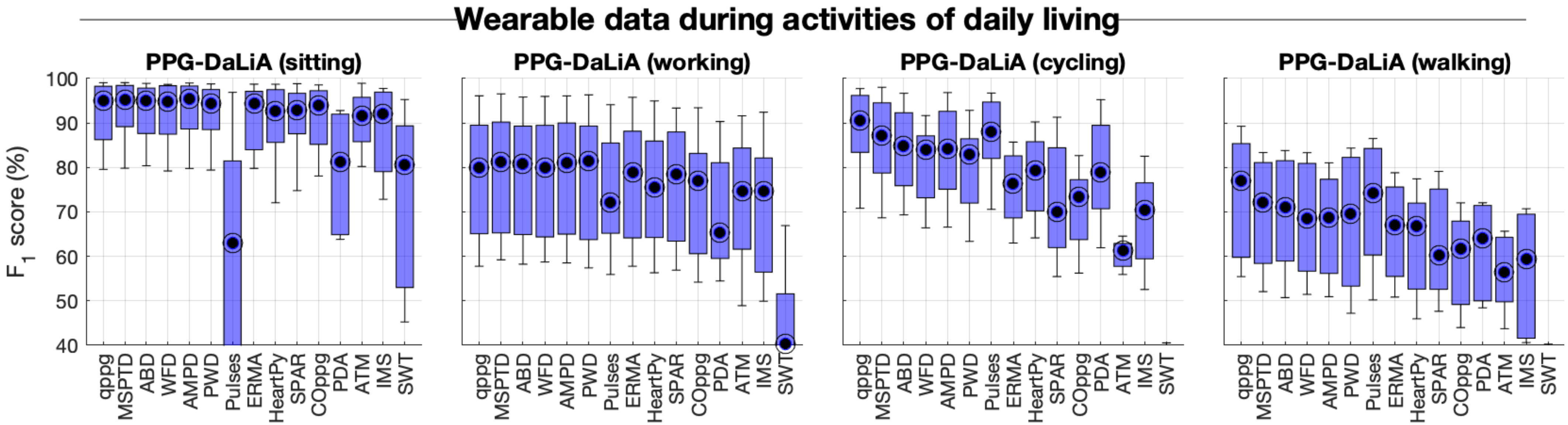
Benchmarking photoplethysmogram beat detectors



Eight beat detectors performed well at rest in the absence of movement

- F1 scores of >=90% on hospital data and wearable data

Benchmarking photoplethysmogram beat detectors



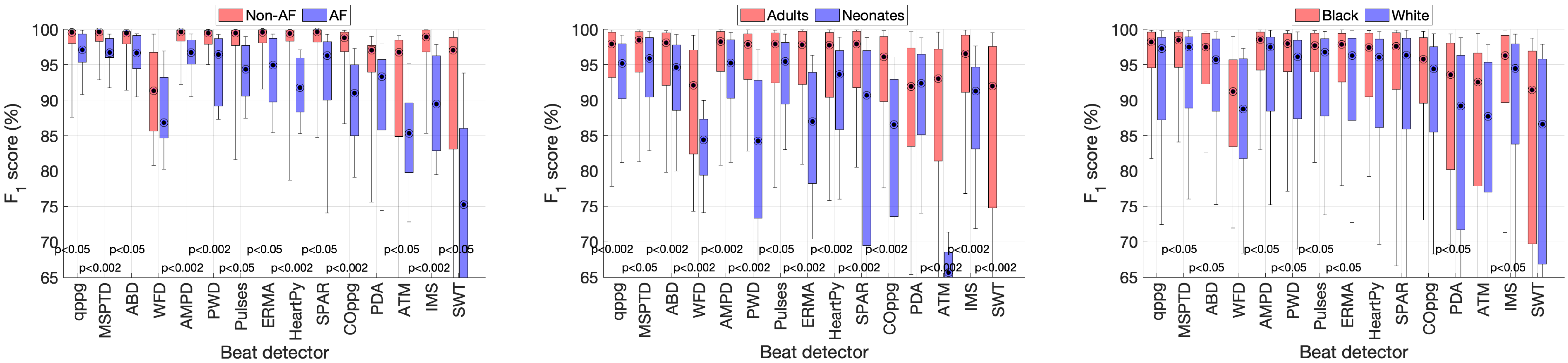
Eight beat detectors performed well at rest in the absence of movement

- F1 scores of >=90% on hospital data and wearable data

Their performance was poorer during exercise:

- F1 scores of 55%–91%

Benchmarking photoplethysmogram beat detectors



Eight beat detectors performed well at rest in the absence of movement

- F1 scores of >=90% on hospital data and wearable data

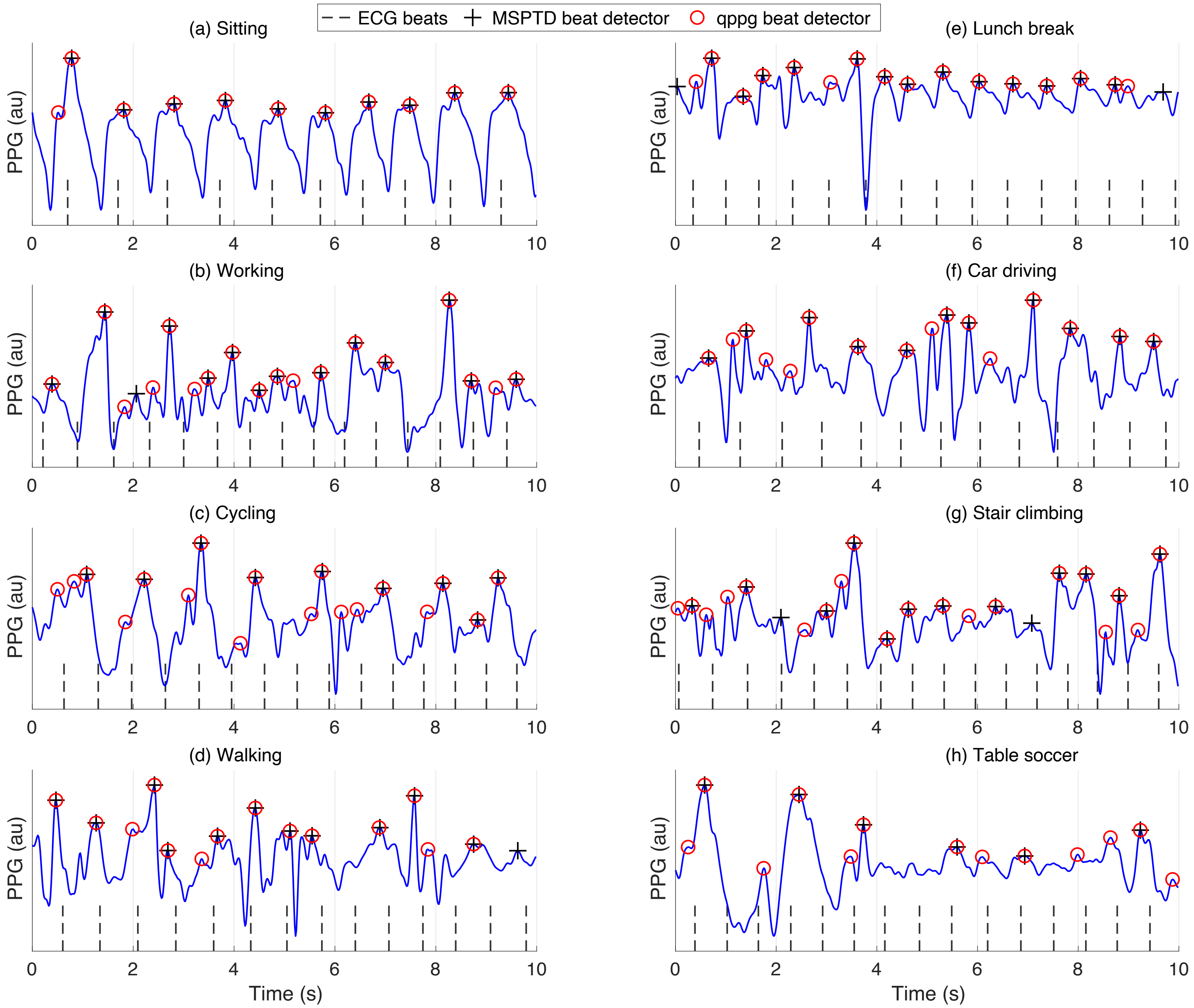
Their performance was poorer during exercise:

- F1 scores of 55%–91%

Performance was:

- Poorer in AF
- Poorer in neonates than adults
- Not associated with ethnicity (Black compared with White)

Concluded that 'MSPTD' and 'qppg' performed best, although this is somewhat subjective.



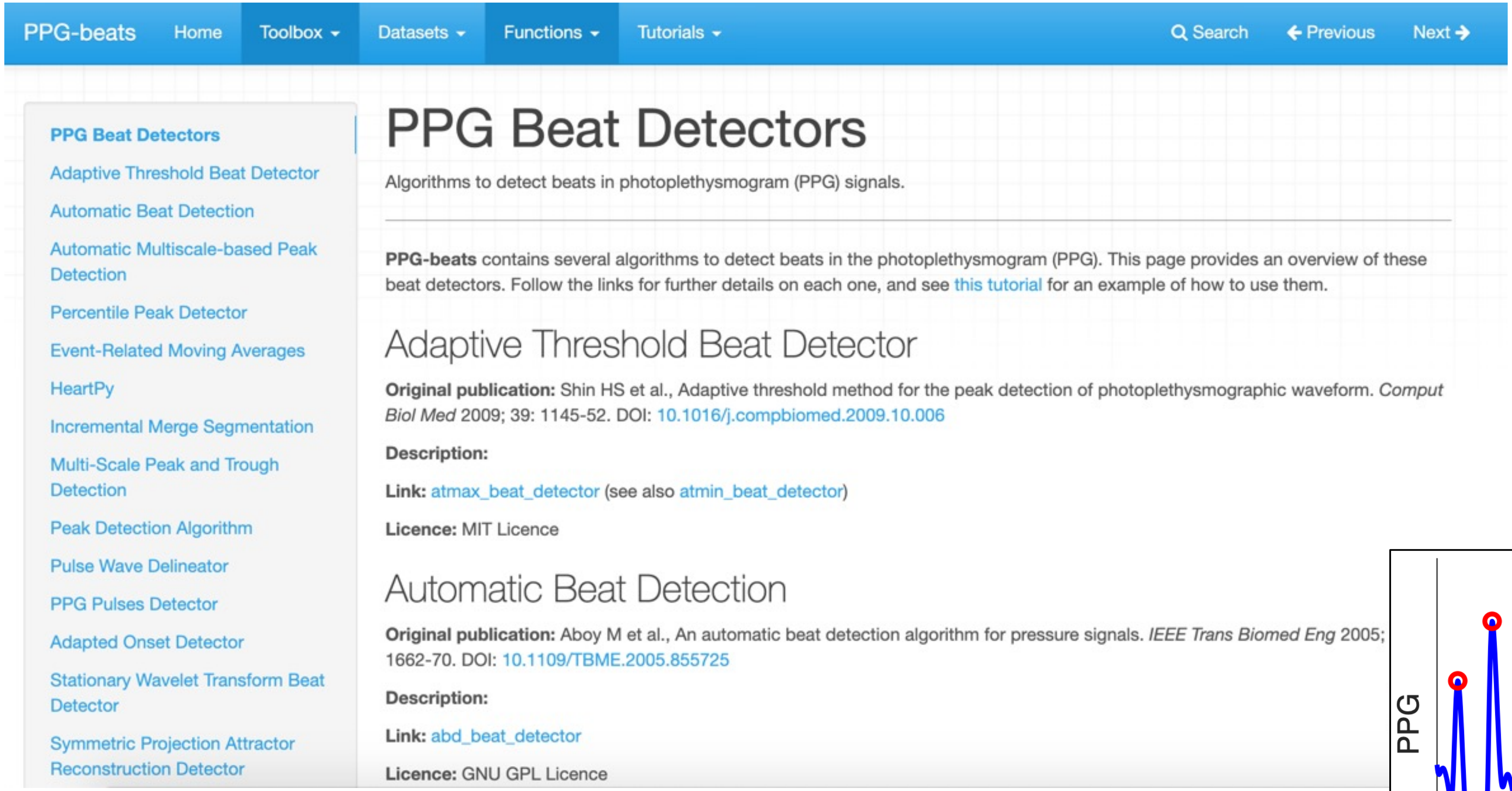
Highlights importance of:

- Using high quality signals
- Signal quality assessment

Ongoing work

Charlton PH *et al.*, Detecting beats in the photoplethysmogram: benchmarking open-source algorithms, Phys Meas., 2022. https://doi.org/10.1088/1361-6579/ac826d

Benchmarking photoplethysmogram beat detectors



PPG-beats | Home | Toolbox | Datasets | Functions | Tutorials | Search | Previous | Next

- PPG Beat Detectors
- Adaptive Threshold Beat Detector
- Automatic Beat Detection
- Automatic Multiscale-based Peak Detection
- Percentile Peak Detector
- Event-Related Moving Averages
- HeartPy
- Incremental Merge Segmentation
- Multi-Scale Peak and Trough Detection
- Peak Detection Algorithm
- Pulse Wave Delineator
- PPG Pulses Detector
- Adapted Onset Detector
- Stationary Wavelet Transform Beat Detector
- Symmetric Projection Attractor Reconstruction Detector

PPG Beat Detectors

Algorithms to detect beats in photoplethysmogram (PPG) signals.

PPG-beats contains several algorithms to detect beats in the photoplethysmogram (PPG). This page provides an overview of these beat detectors. Follow the links for further details on each one, and see this tutorial for an example of how to use them.

Adaptive Threshold Beat Detector

Original publication: Shin HS et al., Adaptive threshold method for the peak detection of photoplethysmographic waveform. *Comput Biol Med* 2009; 39: 1145-52. DOI: 10.1016/j.compbiomed.2009.10.006

Description:

Link: atmax_beat_detector (see also atmin_beat_detector)

Licence: MIT Licence

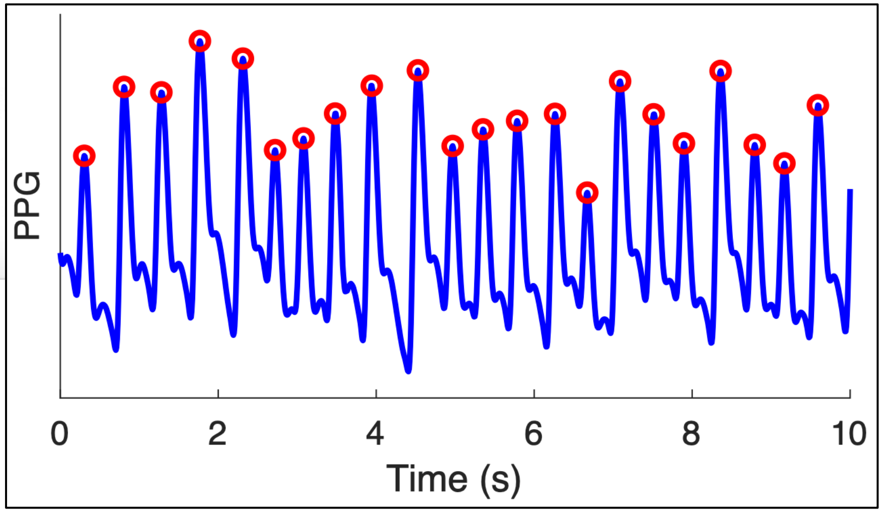
Automatic Beat Detection

Original publication: Aboy M et al., An automatic beat detection algorithm for pressure signals. *IEEE Trans Biomed Eng* 2005; 1662-70. DOI: 10.1109/TBME.2005.855725

Description:

Link: abd_beat_detector

Licence: GNU GPL Licence



https://ppg-beats.readthedocs.io/

Research Directions

Optimising AF detection algorithms

- Beat detection
- Signal quality assessment
- Detecting AF from inter-beat intervals
- Incorporating P-wave analyses

Reducing clinical workload

- Assessing ECG reviewing workload in real-world screening
- Prioritising ECGs for manual review
- (incorporating optimized AF detection algorithms)

SAFER Programme

Aim: to determine whether screening for AF is effective and cost-effective in reducing stroke and other key outcomes compared to current practice.

Feasibility Study 1

10 GP Practices, delivered face-to-face: 2,141 participants screened

SAFER trial

Screening for Atrial Fibrillation with ECG to Reduce stroke
https://www.safer.phpc.cam.ac.uk/

SAFER Programme

Aim: to determine whether screening for AF is effective and cost-effective in reducing stroke and other key outcomes compared to current practice.

Feasibility Study 1

10 GP Practices, delivered face-to-face: 2,141 participants screened

Feasibility Study 2

3 GP Practices, delivered remotely

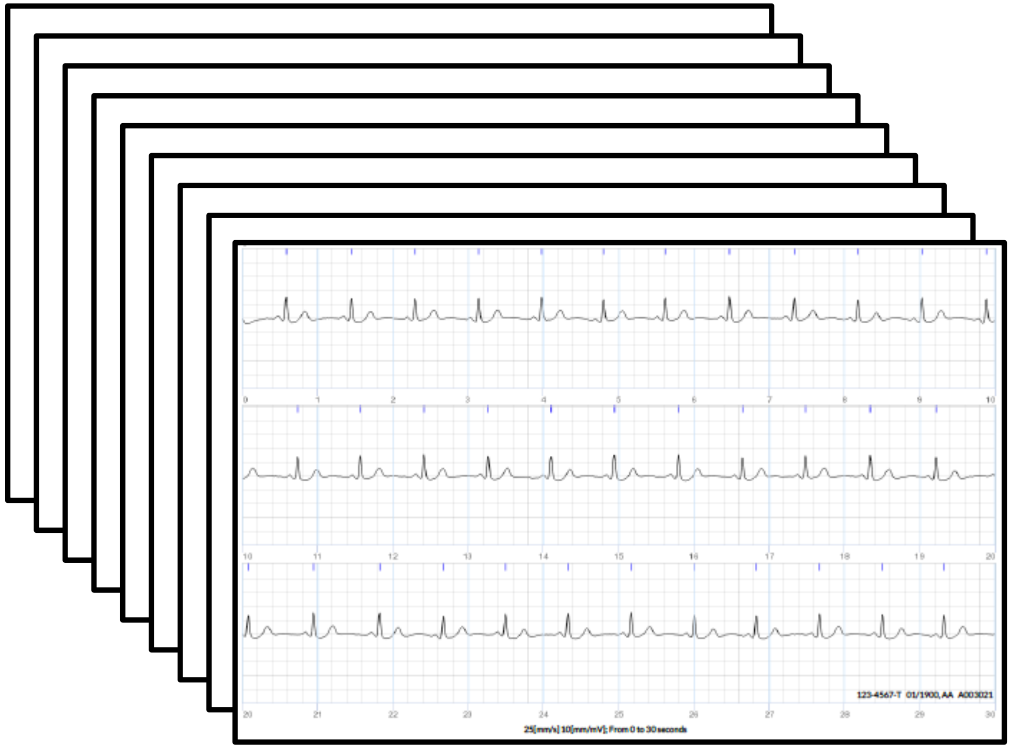
Internal Pilot Trial

36 GP Practices, delivered remotely

Trial (in progress)

100s GP Practices, delivered remotely

ECGs recorded at home



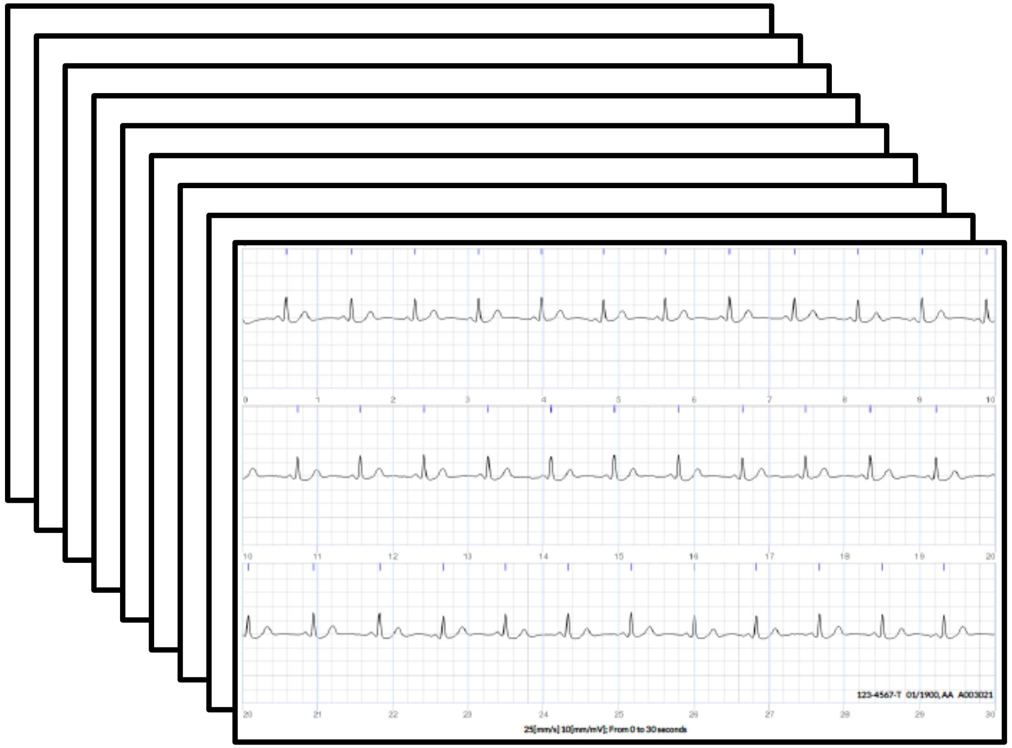
30-second ECGs:

- 4 ECGs per day
- 1-4 weeks
- 61 (53-111) ECGs per participant

Only a small minority exhibit AF
(0.4% were found to exhibit AF in this study)

SAFER trial

ECGs recorded at home



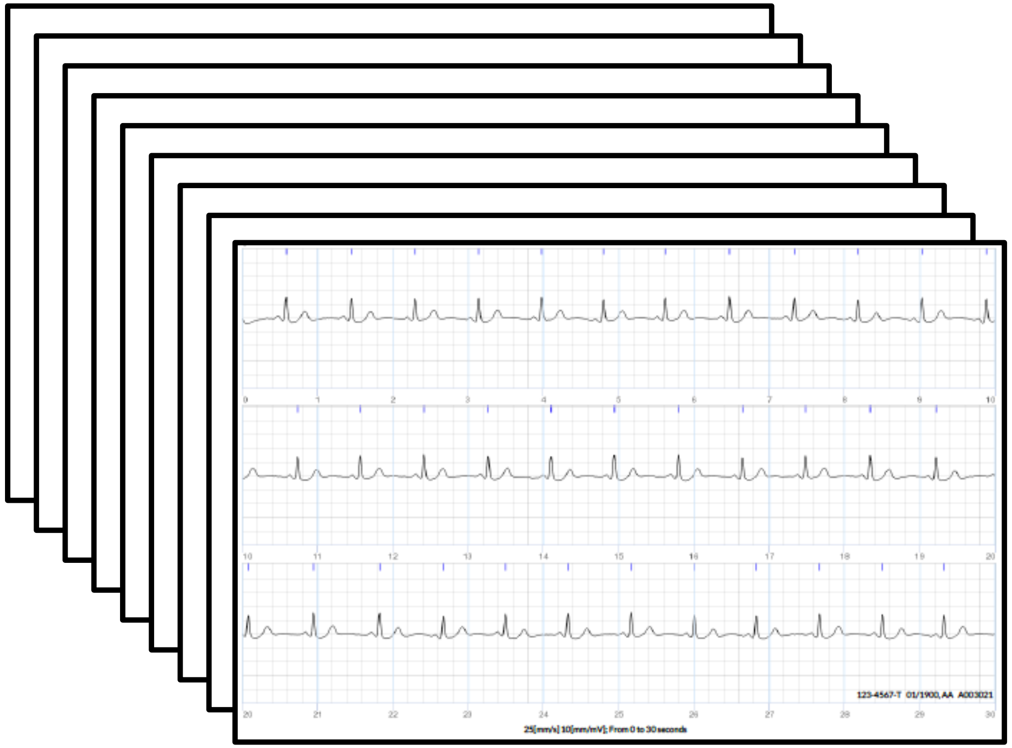
Automated analysis

ECG exhibits:

- ☒ Irregular sequence
- ☐ Fast regular heart rate
- ☐ Fast episode
- ☐ Slow regular heart rate
- ☐ Slow episode
- ☐ Bigemini
- ☐ Trigemini
- ☐ Wide QRS
- ☐ Vent. extra systoles
- ☐ Supra vent. extra systoles
- ☐ Pause
- ☐ Poor Quality

SAFER trial

ECGs recorded at home



Automated analysis

ECG exhibits:

- ☒ Irregular sequence
- ☐ Fast regular heart rate
- ☐ Fast episode
- ☐ Slow regular heart rate
- ☐ Slow episode
- ☐ Bigemini
- ☐ Trigemini
- ☐ Wide QRS
- ☐ Vent. extra systoles
- ☐ Supra vent. extra systoles
- ☐ Pause
- ☐ Poor Quality

Clinical Review

SAFER trial

ECGs recorded at home

162,515 ECGs recorded

(2,141 participants)

Automated analysis

11,975 ECGs identified for review

(1,538 participants)

Clinical Review

687 ECGs found to contain AF

(48 participants)

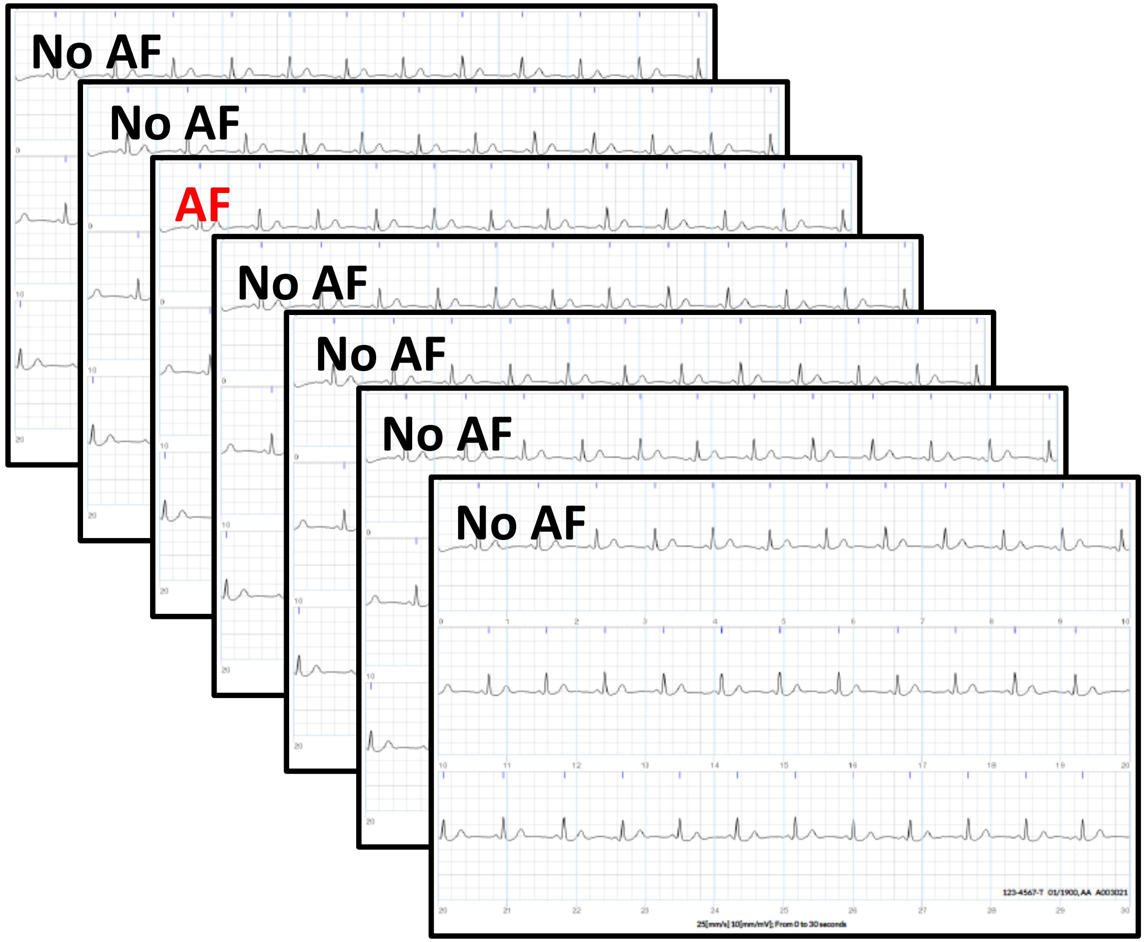
Can we reduce the number identified for review?

Can we reduce the number reviewed?

(reviewing stops after an AF ECG is found)

Prioritising ECGs

- Currently ECGs are reviewed chronologically

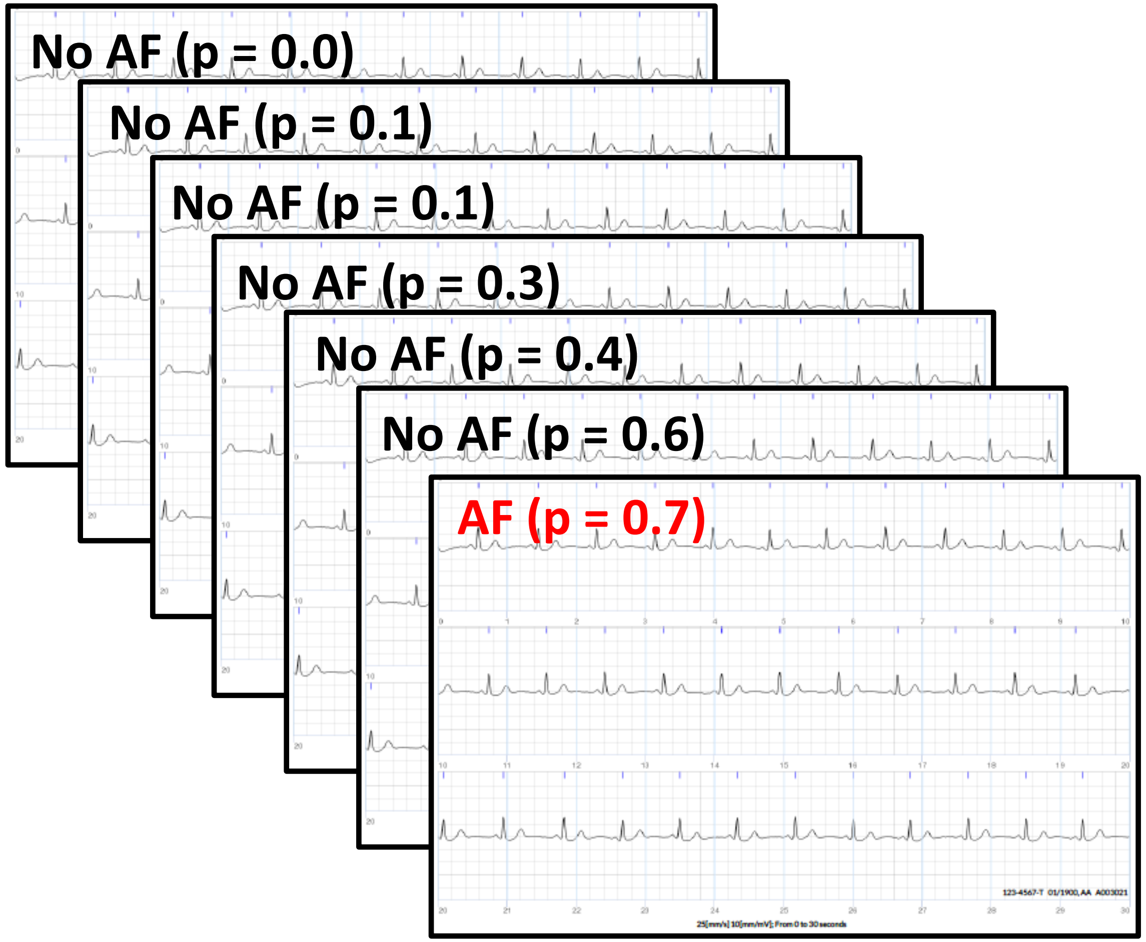


Prioritising ECGs

- ECGs are currently reviewed chronologically

Alternatively:

- Use a model to order an individual's ECGs according to likelihood of AF



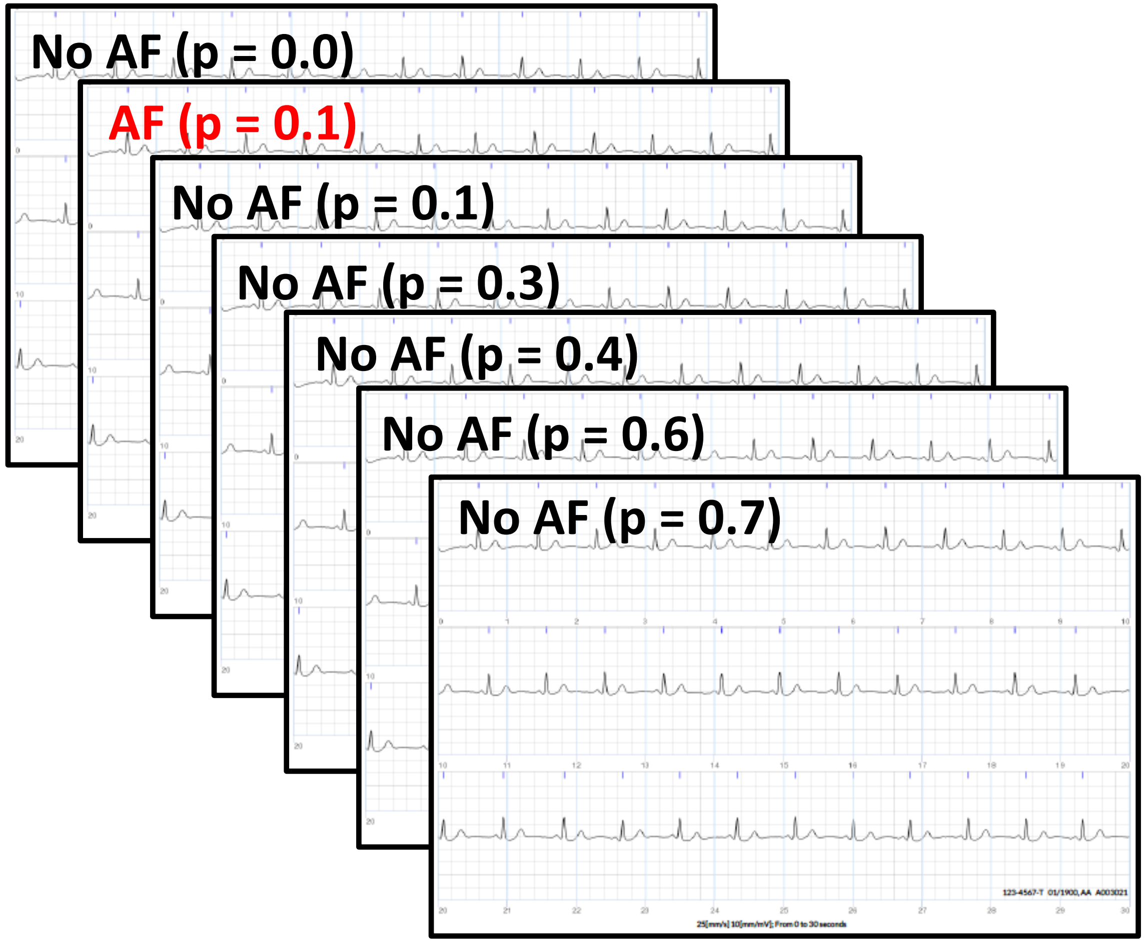
Reduces number of ECGs reviewed, whilst ensuring any participants exhibiting AF are correctly identified.

Prioritising ECGs

- ECGs are currently reviewed chronologically

Alternatively:

- Use a model to order an individual's ECGs according to likelihood of AF



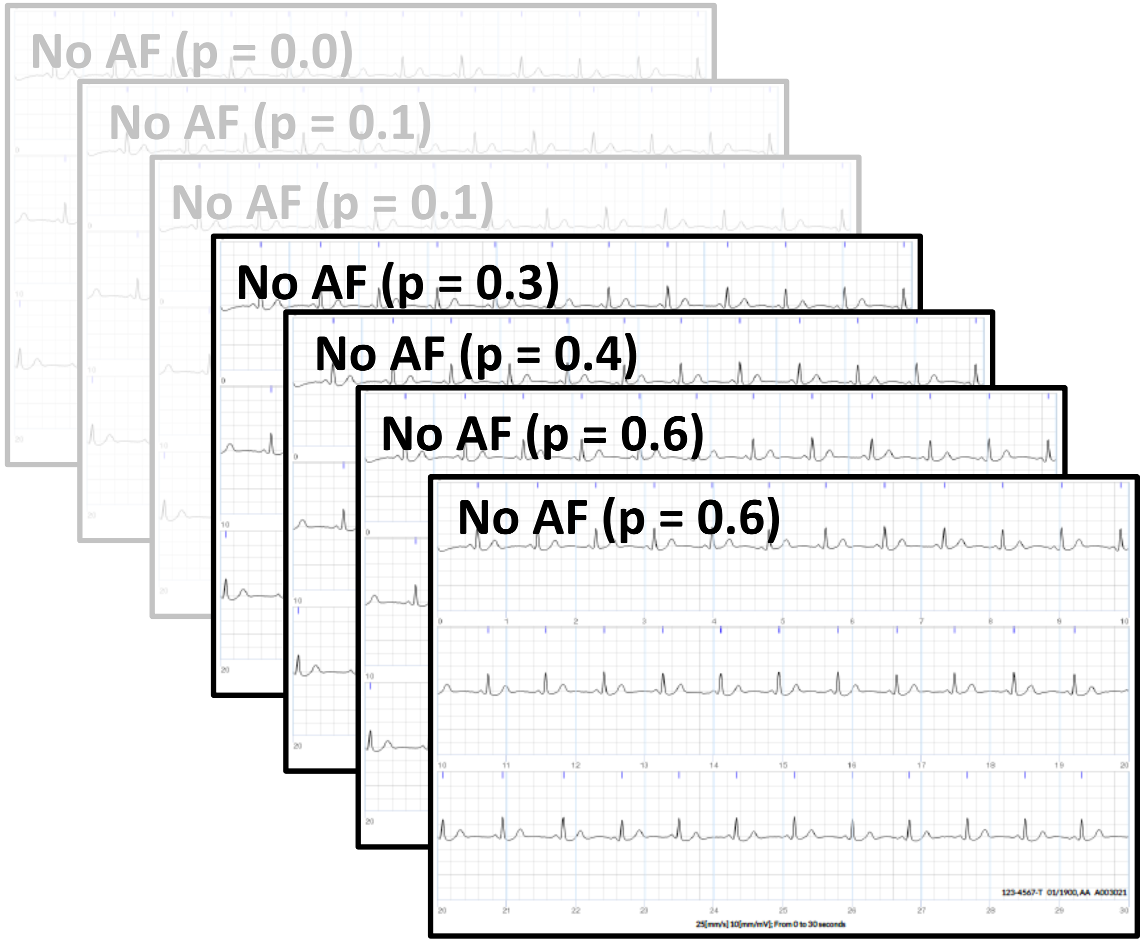
Reduces number of ECGs reviewed, whilst ensuring any participants exhibiting AF are correctly identified.

Prioritising ECGs

- ECGs are currently reviewed chronologically

Alternatively:

- Use a model to order an individual's ECGs according to likelihood of AF
- Possibly only review those ECGs with a likelihood above a threshold



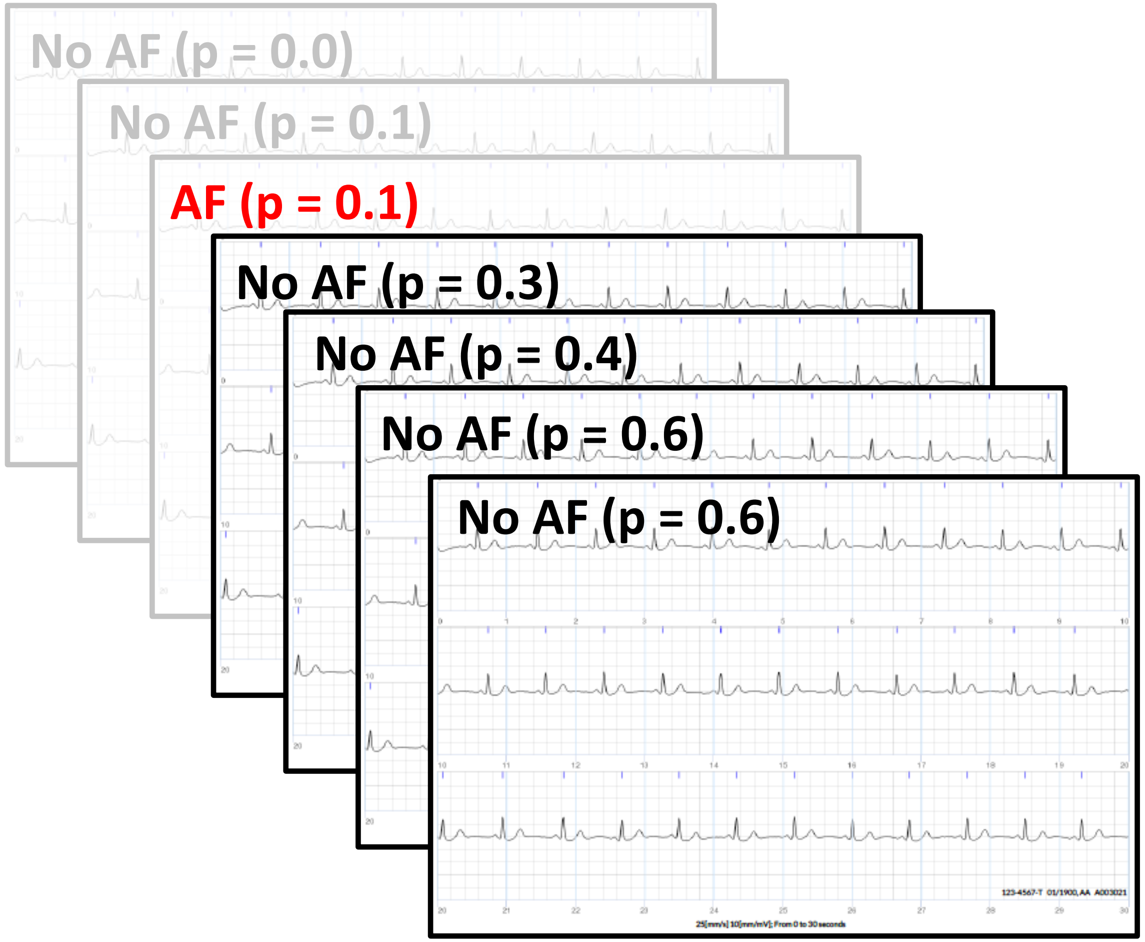
Greatly reduces number of ECGs sent for review, but could result in participants exhibiting AF being missed.

Prioritising ECGs

- ECGs are currently reviewed chronologically

Alternatively:

- Use a model to order an individual's ECGs according to likelihood of AF
- Possibly only review those ECGs with a likelihood above a threshold



Greatly reduces number of ECGs sent for review, but could result in participants exhibiting AF being missed.

Model Development

Inspecting input variables:

ECG Characteristic	Value, median (lower-upper quartiles)	
	AF	non-AF
HR (bpm)	83.0 (71.0 - 94.0)	69.0 (59.0 - 78.0)*
RRmean (ms)	722.9 (638.3 - 845.1)	869.6 (769.2 – 1016.9)*
RRstd (ms)	120.7 (87.5 - 169.5)	131.2 (90.2 - 187.6)
RRvar (%)	16.5 (12.8 - 22.3)	14.3 (10.0 – 20.4)*

RRvar = (RRstd / RRmean) x 100 %

* - significant difference between AF and non-AF

Model Development

Training and validation:

- Multiple logistic regression
- Stepwise
- used 1,428 ECGs (687 manually labelled as AF, 741 non-AF)
- 5-fold cross-validation
  - RRvar and RRmean included in 4 out of 5 models
  - HR in 3 models
  - RRstd in 1 model
- Median AUROC: 71.5%

Model Evaluation

Model configuration	Number of AF diagnoses (%)	Number of reviews (per AF diagnosis)
No model	48 (100)	10,293 (214)

Reviewing ECGs chronologically

Model Evaluation

Model configuration	Number of AF diagnoses (%)	Number of reviews (per AF diagnosis)
No model	48 (100)	10,293 (214)
Model + no threshold	48 (100)	9,950 (207)

Number of ECGs reviewed in AF participants reduced by 74% from 463 to 120, without reducing AF diagnoses. Overall number of reviews reduced by 3.3%.

Model Evaluation

Model configuration	Number of AF diagnoses (%)	Number of reviews (per AF diagnosis)
No model	48 (100)	10,293 (214)
Model + no threshold	48 (100)	9,950 (207)
Model + 25% threshold	48 (100)	7,455 (155)

Number of ECGs reviewed reduced by 28%, without reducing number of AF diagnoses.

Model Evaluation

Model configuration	Number of AF diagnoses (%)	Number of reviews (per AF diagnosis)
No model	48 (100)	10,293 (214)
Model + no threshold	48 (100)	9,950 (207)
Model + 25% threshold	48 (100)	7,455 (155)
Model + 50% threshold	46 (96)	4,885 (106)
Model + 75% threshold	40 (83)	2,375 (59)

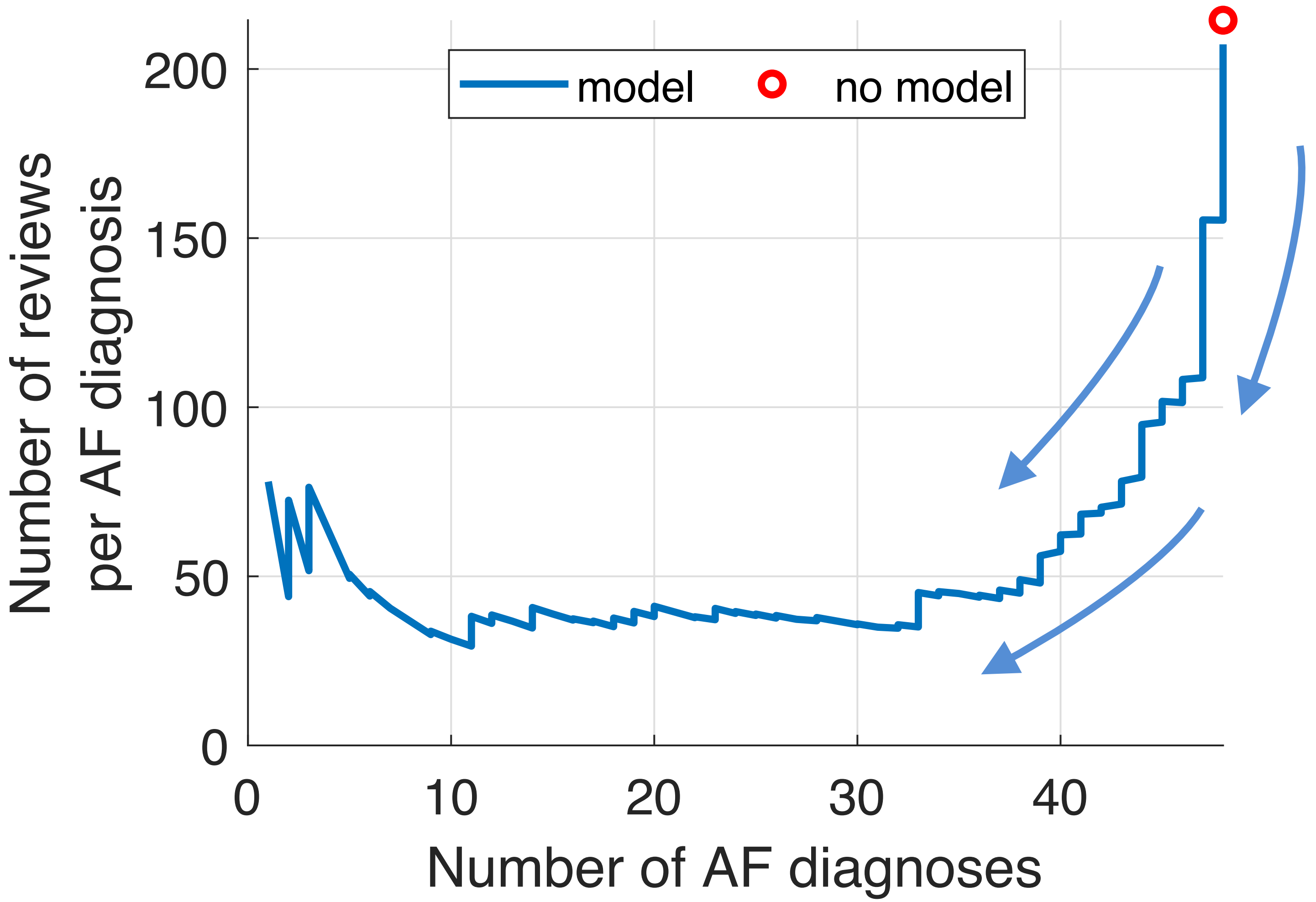
Number of ECGs reviewed reduced by up to 77%, with reduction AF diagnoses of up to 17%.

Discussion

- Using a model to order ECGs only slightly reduced the reviewing workload since most participants don't have AF, and didn't result in missed diagnoses.
- Using a threshold to exclude ECGs can greatly reduce the reviewing workload, but results in some missed diagnoses.

Discussion

- How do we decide what is acceptable?



Research Directions

Optimising AF detection algorithms

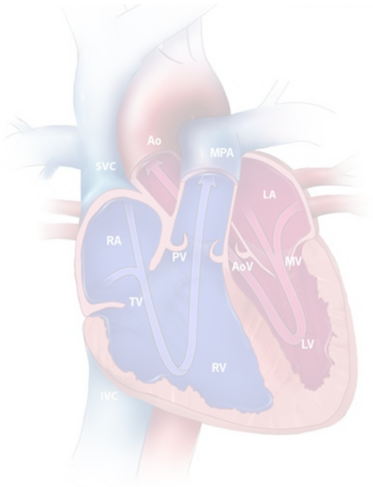
- Beat detection
- Signal quality assessment
- Detecting AF from inter-beat intervals
- Incorporating P-wave analyses

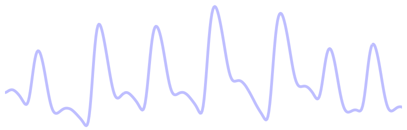
Reducing clinical workload

- Assessing ECG reviewing workload in real-world screening
- Prioritising ECGs for manual review
- (incorporating optimized AF detection algorithms)

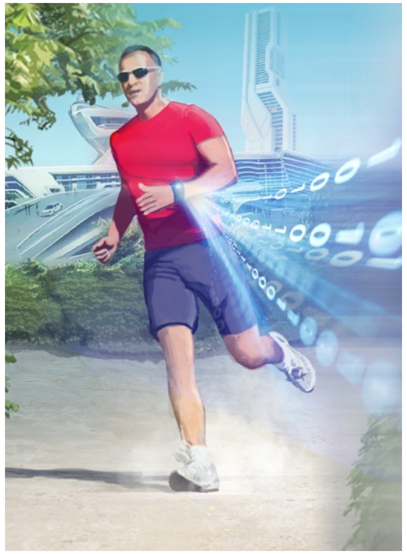
Consumer Wearables

Clinical Applications





Translation



Observations based on the literature

- Much research in this field focuses on developing signal processing algorithms
- Less research focuses on translation
- There is much opportunity for, and benefit to, translational research

Based on:
- Charlton PH *et al.,* 'Breathing rate estimation from the electrocardiogram and photoplethysmogram: a review': https://doi.org/10.1109/RBME.2017.2763681
- Charlton PH et al., 'Assessing hemodynamics from the photoplethysmogram to gain insights into vascular age: a review from VascAgeNet': https://doi.org/10.1152/ajpheart.00392.2021

Selecting a clinical use case for research

Consider:

- Would a wearable add value to current practice? *e.g.*
    - Physiological assessment where it would not otherwise be performed
    - Frequent monitoring where measurements would otherwise be intermittent
- Would the results be actionable? *e.g.*
    - Prompt further assessment (relatively safe)
    - Diagnosis (and treatment) (higher risk)
- Could it be integrated into a clinical pathway?
- Is it cost-effective?
- Many use cases focus on "prevention of avoidable illness and its exacerbations" (NHS Long Term Plan)

Potential clinical pathways

- Screening
- Patient-led measurements to prompt clinical assessment *e.g.*
  - Bradycardia as sign of possible heart block: https://doi.org/10.1016/j.jaccas.2019.11.087
  - Smartwatch ECG capturing ventricular tachycardia: https://doi.org/10.1016%2Fj.hrcr.2020.08.003
- Population-level surveillance
- Self-directed health monitoring

Validation processes

- Much work has been done, and is ongoing, on validation of cuffless blood pressure devices, e.g.
  - IEEE Standard for Wearable, Cuffless Blood Pressure Measuring Devices
  - New standard in development
  - Mukkamala *et al.*, 'Evaluation of the Accuracy of Cuffless Blood Pressure Measurement Devices: Challenges and Proposals': https://doi.org/10.1161/HYPERTENSIONAHA.121.17747
- The Interlive Network has produced validation protocols for other parameters:
  - step counts
  - heart rate
  - energy expenditure
  - maximal oxygen consumption
- A personal interest: Can individual parts of a system be validated in isolation?
  - Explored somewhat in: Charlton PH *et al.*, 'Establishing best practices in photoplethysmography signal acquisition and processing': https://doi.org/10.1088/1361-6579/ac6cc4

Developing devices suitable for clinical decision making

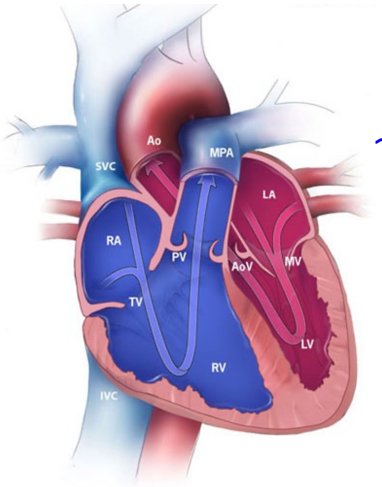
Personal considerations:

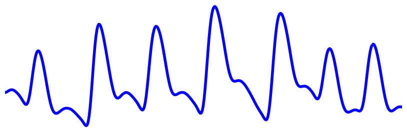
- What are the respective roles of academia and industry? *e.g. in academia:*
  - Algorithm development
  - Algorithm validation
  - (?) Algorithm source code
  - (?) Teaching, e.g. Biomedical Signal Processing Jupyter book
- What are the most pressing research directions?
  - 'The 2023 Wearable Photoplethysmography Roadmap' should be published shortly. It's written by 51 researchers from academia, industry and clinical practice, and presents a comprehensive overview of pressing research directions within the field of wearable photoplethysmography.

Charlton P.H. *et al.*, **Establishing best practices in photoplethysmography signal acquisition and processing**, *Phys* Meas, 2022: https://doi.org/10.1088/1361-6579/ac6cc4

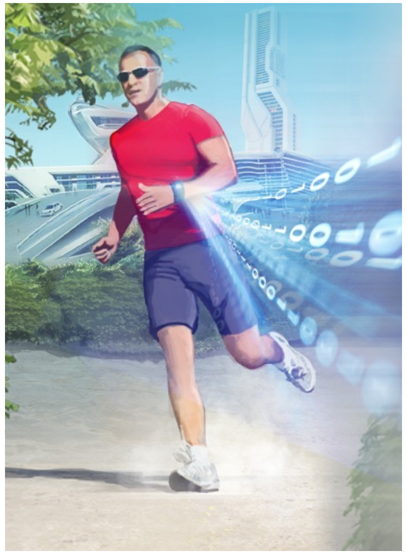
Consumer Wearables

Clinical Applications





Translation



With thanks to…

Jonathan Mant

Panicos Kyriacou

Jordi Alastruey-Arimon

The SAFER Research Team

Guy’s and St Thomas’ NHS Foundation Trust

King’s College London

University of Cambridge

EPSRC

British Heart Foundation

The functionality of wearable devices is expanding, and they have a wide range of potential clinical applications.

There is much opportunity for research into if, and how, to best use consumer wearables for clinical decision making

If used for clinical decision making, then consumer wearables should be like a climbing rope:

- highly reliable
- used for their specific, intended purpose

Sonia Sevilla, https://commons.wikimedia.org/wiki/File:Kamar_Zard_Buzhan_-_Nishapur_1.jpg (CC0 1.0)

Using consumer wearables for clinical decision making

Dr Peter H. Charlton

https://peterhcharlton.github.io

pc657@cam.ac.uk

For further reading see:

Charlton P.H. *et al.*, **Wearable Photoplethysmography for Cardiovascular Monitoring**, Proc. *IEEE*, 2022, https://doi.org/10.1109/JPROC.2022.3149785