Physiological Measurement

The influence of cardiac arrhythmias on the detection of heartbeats in the photoplethysmogram: benchmarking open-source algorithms

IOP Publishing Physiol. Meas. 45 (2024) 025005 https://doi.org/10.1088/1361-6579/ad2216

Physiological Measurement

IPEM Institute of Physics and Engineering in Medicine

OPEN ACCESS

RECEIVED
23 August 2023

REVISED
6 December 2023

ACCEPTED FOR PUBLICATION
24 January 2024

PUBLISHED
14 February 2024

PAPER

The influence of cardiac arrhythmias on the detection of heartbeats in the photoplethysmogram: benchmarking open-source algorithms

Loïc Jeanningros[1,2], Mathieu Le Bloa[3], Cheryl Teres[3], Claudia Herrera Siklody[3], Alessandra Porretta[3], Patrizio Pascale[3], Adrian Luca[3], Jorge Solana Muñoz[3], Giulia Domenichini[3], Théo A Meister[4], Rodrigo Soria Maldonado[4], Hildegard Tanner[4], Jean-Marc Vesin[2], Jean-Philippe Thiran[2], Mathieu Lemay[1], Emrush Rexhaj[4], Etienne Pruvot[3] and Fabian Braun[1]

[1] Swiss Center for Electronics and Microtechnology, Neuchâtel, Switzerland
[2] Swiss Federal Institute of Technology Lausanne, Lausanne, Switzerland
[3] Service of Cardiology, Lausanne University Hospital, Lausanne, Switzerland
[4] Department of Cardiology and Biomedical Research, University Hospital Bern, University of Bern, Bern, Switzerland

E-mail: loic.jeanningros@csem.ch

Keywords: photoplethysmography, beat detection, heartbeat, cardiac arrhythmias, atrial fibrillation, bigeminy, ventricular tachycardia

Abstract

Objective. Cardiac arrhythmias are a leading cause of mortality worldwide. Wearable devices based on photoplethysmography give the opportunity to screen large populations, hence allowing for an earlier detection of pathological rhythms that might reduce the risks of complications and medical costs. While most of beat detection algorithms have been evaluated on normal sinus rhythm or atrial fibrillation recordings, the performance of these algorithms in patients with other cardiac arrhythmias, such as ventricular tachycardia or bigeminy, remain unknown to date. *Approach.* The *PPG-beats* open-source framework, developed by Charlton and colleagues, evaluates the performance of the beat detectors named *QPPG*, *MSPTD* and *ABD* among others. We applied the *PPG-beats* framework on two newly acquired datasets, one containing seven different types of cardiac arrhythmia in hospital settings, and another dataset including two cardiac arrhythmias in ambulatory settings. *Main Results.* In a clinical setting, the *QPPG* beat detector performed best on atrial fibrillation (with a median F_1 score of 94.4%), atrial flutter (95.2%), atrial tachycardia (87.0%), sinus rhythm (97.7%), ventricular tachycardia (83.9%) and was ranked 2nd for bigeminy (75.7%) behind *ABD* detector (76.1%). In an ambulatory setting, the *MSPTD* beat detector performed best on normal sinus rhythm (94.6%), and the *QPPG* detector on atrial fibrillation (91.6%) and bigeminy (80.0%). *Significance.* Overall, the PPG beat detectors *QPPG*, *MSPTD* and *ABD* consistently achieved higher performances than other detectors. However, the detection of beats from wrist-PPG signals is compromised in presence of bigeminy or ventricular tachycardia.

1. Introduction

Cardiac arrhythmias (CAs) have a prevalence of 3.2%–6.6% in the elderly European and US populations (aged 65–73 years) (Khurshid *et al* 2018) and are associated with high morbidity and mortality (Tsao *et al* 2023). Indeed, ventricular arrhythmias are a major cause of sudden cardiac deaths, which are estimated to 10%–20% of all deaths in Europe (Zeppenfeld *et al* 2022). Due to the asymptomatic and intermittent nature of certain CAs in their early stages (Rho and Page 2005, Gorenek Chair *et al* Gorenek (chair) 2017), they are often diagnosed late, at time of hospitalization for stroke or heart failure.

Photoplethysmography (PPG) is a promising technology for long-term and continuous ambulatory monitoring of cardiovascular parameters such as blood pressure and heart rhythm. PPG measures changes in blood volume by optical means and is often integrated in wearable devices like smartwatches (Lemay *et al* 2020, Allen and Kyriacou 2021). Consequently, PPG-based devices have great potential for the early detection of CAs, leading to improved diagnosis, treatment and a reduction in complications.

Numerous studies have investigated the detection of atrial fibrillation (AF), the most common CA, affecting up to 34 million people worldwide (Chugh *et al* 2014, Hindricks *et al* 2021). Most of these studies relied on the analysis of irregularities in inter-beat intervals (IBIs). Besides IBIs, CAs also distort the morphology of individual PPG pulses. Such information can be extracted by pulse wave analysis (PWA) (Proença *et al* 2019) to improve the detection of CAs (Jeanningros *et al* 2022, Basza *et al* 2023). However, both IBIs and PWA rely on an accurate detection of heartbeats in the PPG signal. A suboptimal beat detection would introduce IBIs that contain two pulses (false negative detections) and pulses split in two IBIs (false positive detections). This would bias IBI-based measures of irregularity (Shannon entropy, RMSSD, pNN50, …) and compromise PWA computation.

Whereas beat detectors can be very accurate for healthy subjects (Charlton *et al* 2022), their performance has not been studied in the presence of different CAs. Only few studies focused on the evaluation of PPG beat detection performance during AF. Harju *et al* (2018) reported a mean absolute error (MAE) of 51 ms on IBI estimation from wrist-worn PPG in 21 subjects with AF. Their detection performance corresponds to an F_1 score of 96.5%. Väliaho *et al* (2019) reported performance equivalent to 94.5% F_1 score for pulse detection on 106 patients with AF. Recently, Charlton *et al* (2022) compared fifteen open-source beat detectors on multiple datasets associated with various conditions. Among them, the eight detectors that performed best overall achieved F_1 scores between 91.8% and 97.1% on 19 patients suffering from AF. Han *et al* (2022) developed a complex beat detector designed for HR estimation in presence of CAs. Their *SWEPD* algorithm detected IBIs with an F_1 score of 97.2% in 21 patients with AF and 97.8% when analyzing performance in the presence of frequent atrial and ventricular premature contractions.

To the best of our knowledge, there is no study that compared the performance of various beat detectors on various types of CAs. Considering CAs other than AF is important when screening large populations potentially displaying pathological rhythms, such as ventricular and atrial bigeminy or ventricular tachycardia. Hence, the choice of beat detectors can be a determining factor for the performance of CAs classificators based on IBIs and PWA.

In this study, we used the open-source *PPG-beats* framework developed by Charlton *et al* (2022) to benchmark the performance of 15 open-source beat detectors. The framework was applied on two newly acquired datasets containing 8 different types of CAs. The goals of this work are (1) to evaluate which beat detectors are effective and reliable in presence of verious types of CAs, and (2) to identify CAs for which heartbeat detection from wrist-PPG signals is limited.

2. Materials and methods

2.1. Datasets

This research was conducted in accordance with the principles embodied in the Declaration of Helsinki, as well as local statutory requirements. All participants gave written informed consent to participate in the study. Subjects were offered to take part in the study regardless of their sex. Hence, the proportion of males and females is supposed to reflect the frequency of medical interventions for each sex.

2.1.1. Clinical dataset

The first dataset includes 58 patients referred for diagnostic or therapeutic electrophysiological procedures at the Lausanne University Hospital (CHUV). This study has been accepted by the local ethics committee of Lausanne (CER-VD, Project-ID 2021–00586) and registered on http://ClinicalTrials.gov (NCT04884100).

PPG signals were acquired at 100 Hz from a proprietary wrist-bracelet (CSEM, Neuchâtel, Switzerland). Concurrently, 12-lead ECG signals were recorded using the Axiom Sensis XP® System (Siemens®, Munich, Germany) at 2 kHz sampling frequency and bandpass filter settings of 0.5–200 Hz. ECG signals were used for gold standard annotations of both R-peaks (beats) and CAs.

2.1.2. Ambulatory dataset

The second dataset includes 44 subjects referred for an ambulatory Holter ECG recording for either 24 h (40 subjects) or 7 days (4 subjects). The clinical study has been conducted at Inselspital in Bern and is still ongoing. It has been accepted by the local ethical committee KEK-BERN (Project-ID 2021–02117). PPG signals were recorded with the same proprietary wrist-bracelet from CSEM as for the clinical dataset, together with a 3-lead Holter ECG monitor Lifecard CF (Spacelabs Healthcare®, Issaquah, Washington, USA). R-peaks and CAs were annotated by the software Sentinel from Spacelabs Healthcare®. To exclude PPG signals corrupted by motion artifacts, only periods for which motion was continuously low were selected. To this end, a moving average filter of 2000 s window was applied every 60 s on the absolute value of the differences in normed 3D accelerometer signals. Periods where the moving average was below 0.15 mG s^{-1} were considered as low motion. Only periods

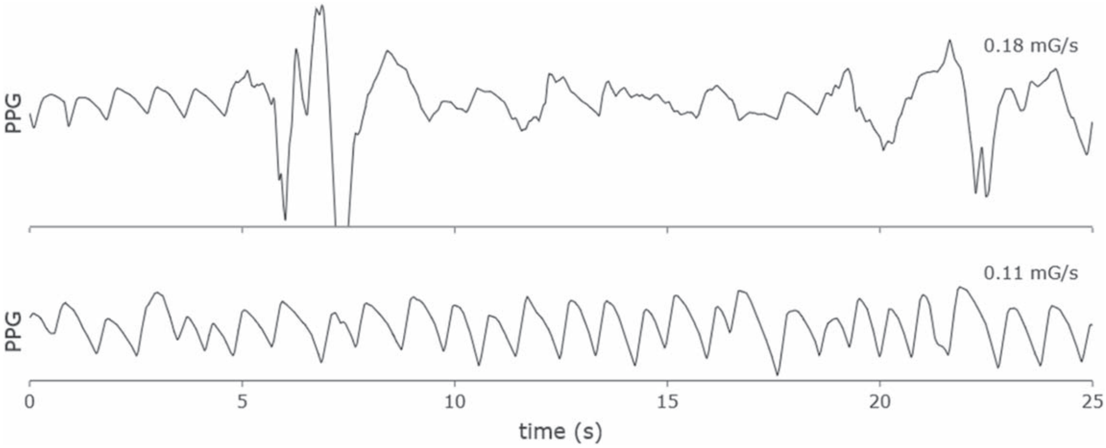

Figure 1. Example of motion influence on PPG signals from the ambulatory dataset. The top row shows a PPG signal with a motion level of 0.18 mG s^{-1} (with respect to the moving average described in section 2.1.2). The bottom row shows a PPG signal from the same patient with a motion level of 0.11 mG s^{-1}. The threshold to reject periods corrupted by motion was set at 0.15 mG s^{-1}.

Table 1. PPG beat detectors evaluated in the present study.

Beat detector	Original author
ABD: automatic beat detection	Aboy *et al* (2005)
AMPD: automatic multiscale peak detection	Scholkmann *et al* (2012)
ATM: adaptative threshold method	Shin *et al* (2009)
COPPG: percentile peak detector	Orphanidou *et al* (2015)
ERMA: event-related moving averages	Elgendi *et al* (2013)
HEARTPY	van Gent *et al* (2019)
IMS: incremental merge segmentation	Karlen *et al* (2012)
MSPTD: multi-scale peak and trough detection	Bishop and Ercole (2018)
PDA: peak detection algorithm	Argüello Prada and Serna Maldonado (2018)
PULSES: PPG pulses detector	Lázaro *et al* (2014)
QPPG: adapted onset detector	Zong *et al* (2003)
SWT: stationary wavelet transform	Vadrevu and Manikandan (2019)
WFD: wavelet foot delineation	Conn and Borkholder (2013)

lasting more than 10 min were kept for analysis. PPG signals with a moving average below and above 0.15 mG s^{-1} are shown in figure 1.

2.1.3. Cardiac arrhythmia labelling

ECG signals of the clinical dataset were annotated by a medical expert who manually identified CAs. In contrast, ECG signals from the ambulatory dataset have been automatically annotated by the software Sentinel from Spacelabs Healthcare® and corrected by a cardiologist. Independently of the dataset, both atrial and ventricular bigeminy, as well as trigeminy and quadrigeminy, or any combination of these rhythms, were indistinctly labeled as bigeminy. The label AVRT includes both atrioventricular reetrant tachycardia and atrioventricular nodal reentrant tachycardia. Finally, single atrial and ventricular premature contractions were not considered as CAs and were therefore ignored in this study.

2.2. PPG beat detector evaluation

The *PPG-beats* framework[5] provided by Charlton and colleagues (Charlton *et al* 2022) was applied. The methods used to evaluate PPG beat detectors are identical to those of the original paper (Charlton *et al* 2022). The essential steps are summarized in the following.

The PPG signals underwent bandpass filtering between 0.67 and 8.0 Hz to eliminate non-cardiac frequencies. Then, beats were detected using thirteen open-source detectors listed in table 1. The *PPG-beats* framework (Charlton *et al* 2022) provides two additional detectors (*SPAR* and *PWD*) which had to be removed from analysis because of runtime errors for several signals. To apply PPG beat detection, the PPG signals were

[5] http://github.com/peterhcharlton/ppg-beats

Table 2. List of cardiac arrhythmias with corresponding demographic and quantitative statistics. Demographic statistics are specified for males (M) and females (F). Durations include only motion-free periods.

	Cardiac arrhythmia	Subjects (F/M)	Duration (h)
	Clinical dataset	58 (18/40)	81.4
AF	Atrial fibrillation	12 (5/7)	5.4
AFL	Atrial flutter	9 (1/8)	7.8
AT	Atrial tachycardia	3 (1/2)	1.2
AVB	Atrioventricular block	2 (1/1)	0.5
AVRT	Atrioventricular reentrant tachycardia	8 (3/5)	0.3
Bi	Bigeminy (atrial or ventricular)	10 (3/7)	4.6
SR	Sinus rhythm (normal)	58 (18/40)	58.8
VT	Ventricular tachycardia	10 (2/8)	2.9
	Ambulatory dataset	*44 (20/24)*	*684.5*
AF	Atrial Fibrillation	8 (4/4)	69.9
Bi	Bigeminy (atrial and ventricular)	11 (3/8)	17.4
SR	Sinus rhythm (normal)	37 (16/21)	597.2

segmented into 20 s windows with a 5 s overlap. Duplicate beats within overlapping segments were removed. This method guaranteed that no beat detectors were penalized for missing beats at the end or the start of the window (e.g. during initialization of the detector).

Depending on the detector, timings of detected beats could either correspond to the pulse foot, the systolic peak, or the maximum of the first derivative. In order to perform an analysis that is comparable for all detectors, the middle-amplitude point of systolic upslope, defined as the timing associated with the mean amplitude of the pulse foot and the systolic peak, was used for analysis. To do so, for each beat, the preceding minimum (pulse foot) and subsequent maximum (systolic peak) were extracted if not yet provided by the detector. To synchronize PPG beats with reference ECG beats, ECG beats were considered correctly identified if at least one PPG beat was closer than 150 ms. The lag associated with the maximum number of correctly identified ECG beats was used to align the two beat time series. The synchronization step was directly applied on the full records of the clinical dataset and on low-motion periods (>10 min) of the ambulatory dataset. The performance of the beat detectors was evaluated based on the number of reference ECG beats (n_{ref}), estimated PPG beats (n_{PPG}), and correctly identified beats (n_{correct}) to calculate sensitivity (Se), positive predictive value (PPV) and F_1 score (F_1) as follows: $\text{Se} = \frac{n_{\text{correct}}}{n_{\text{ref}}} \times 100$

$$PPV = \frac{n_{correct}}{n_{PPG}} \times 100$$

$$F_1 = \frac{2 \times \text{PPV} \times \text{Se}}{\text{PPV} + \text{Se}} \times 100.$$

The performance metrics were calculated on a per-rhythm basis, both for the entire cohort and individually for each subject. To achieve this, reference ECG beats, estimated PPG beats, and correctly identified beats were aggregated by rhythm if they belonged to a homogeneous rhythmic event lasting at least 25 s.

3. Results

3.1. Datasets

Table 2 details the seven different types of CA that were recorded in the clinical dataset and the two types of CA present in the ambulatory dataset together with the corresponding cumulative duration of arrhythmic events and the number of patients experiencing the specific CA. Among 58 subjects involved in the clinical dataset, 40 were men and 18 were women with a mean age of 56 ± 16 years. Skin color was categorized according to Fitzpatrick scale, as I (5 patients), II (26), III (9), IV (1), V (1), VI (1) and 1 patient had missing data. The ambulatory dataset consisted of 24 men and 20 women, with a mean age of 56 ± 16 years. Their skin colors were I (24), II (18), III (11), IV (1), VI (1) and 3 patients with missing data. The imbalance between the number of male and female is in accordance with the prevalence of CAs that affect males more frequently than females (Khurshid *et al* 2018). However the imbalance is very large in the clinical dataset, but no other reason than randomness can be identified to explain this difference.

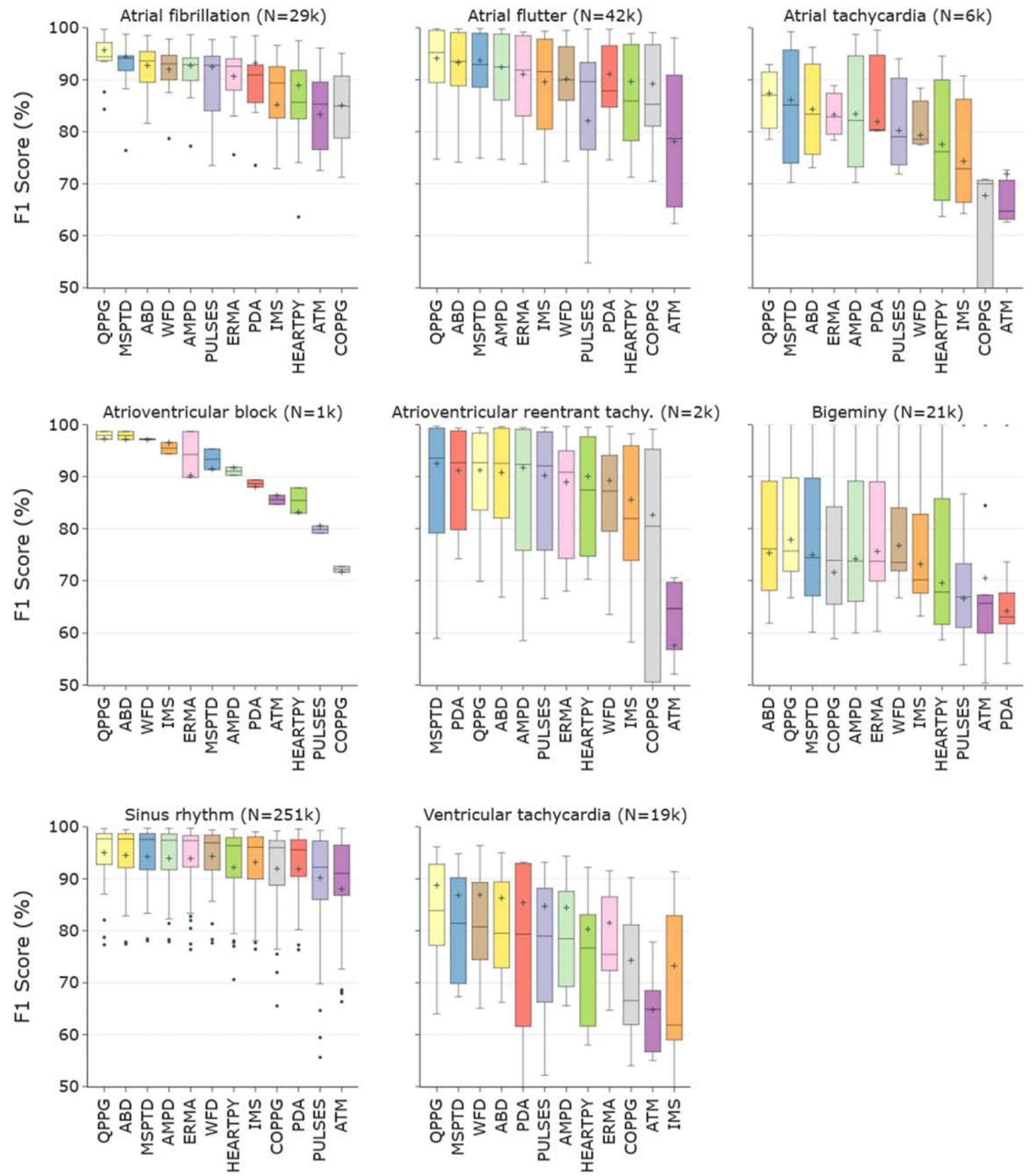

Figure 2. Beat detector performance (F_1 score) comparison by cardiac arrhythmia on the clinical dataset. The number of reference beats (N) per CA is written next to each subtitle in thousands (k). Black dots represent outlier subjects, boxes show the median, 1st and 4th quartiles, 10th and 90th percentiles of F_1 scores obtained per subjects while the black cross indicates F_1 score calculated across all subjects. Detectors are ordered by decreasing median of F_1 score.

3.2. Beat detector performance

Given the unequal proportions between the number of subjects of the two sexes for the majority of CA (see table 2), the restricted total number of subjects, and the important variability of inter-subject performance, the results are not separately detailed for both sexes.

3.2.1. Clinical dataset

F_1 scores obtained from the clinical dataset are shown in figure 2 and detailed in table 3, along with additional metrics. Medians of F_1 scores on normal sinus rhythm range from 89.6% to 97.7%, with five beat detectors that show similar scores (>97.3%): *QPPG*, *ABD*, *MSPTD*, *AMPD* and *ERMA*. The loss of accuracy when detecting beats during AF or atrial flutter is visible. *QPPG* and *MSPTD* are the best detectors with respectively 94.4% and 94.1% medians of F_1 scores during AF. Beat detection is more unequal across subjects during atrial flutter for which *QPPG* stands out from other detectors with a median F_1 score of 95.2%. Atrial tachycardia and ventricular tachycardia obtain the most spread-out performances of beat detectors. F_1 scores of *QPPG* (87.0%) and *MSPTD* (85.1%) slightly stand out from others on atrial tachycardia. Performances on ventricular tachycardia are highly

Table 3. Beat detector performance on clinical dataset. The medians across subjects of F_1 score, sensitivity (Sens.), and positive predictive value (PPV) in percent (%) are detailed for each cardiac arrhythmia: atrial fibrillation (AF), atrial flutter (AFL), atrial tachycardia (AT), atrioventricular blocks of 2nd and 3rd degree (AVB), atrioventricular reentrant (nodal and non-nodal) tachycardia (AVRT), atrial and ventricular bigeminy (Bi), normal sinus rhythm (SR) and ventricular tachycardia (VT).

		ABD	AMPD	ATM	COPPG	ERMA	HEARTPY	IMS	MSPTD	PDA	PULSES	QPPG	SWT	WFD
AF	F_1 score	93.6	92.9	85.3	84.9	92.6	85.7	89.4	94.1	90.9	92.8	94.4	72.9	93.1
	Sens.	89.5	89.8	80.6	77.6	89.8	79.7	82.8	91.3	87.4	87.2	92.0	63.6	90.0
	PPV	98.3	95.8	93.2	96.3	96.0	94.4	97.8	96.2	95.4	98.2	98.6	92.7	97.5
AFL	F_1 score	93.5	92.5	78.7	85.3	91.8	85.9	91.5	92.9	87.8	89.6	95.2	70.4	90.0
	Sens.	95.0	92.6	70.3	80.4	94.6	82.5	85.6	94.6	84.5	83.6	96.5	62.8	89.7
	PPV	98.4	96.1	90.7	95.6	94.8	94.1	97.1	96.6	97.1	96.6	99.2	88.7	96.2
AT	F_1 score	83.4	82.2	64.7	70.0	82.9	76.2	72.9	85.1	80.3	79.0	87.0	52.4	78.5
	Sens.	76.7	75.4	48.9	55.3	75.9	67.5	59.9	80.3	72.3	70.1	81.1	37.1	70.4
	PPV	99.8	98.6	93.9	98.6	96.7	92.8	98.4	98.6	92.2	99.8	98.9	92.1	97.9
AVB	F_1 score	97.9	91.0	85.6	72.2	94.3	85.4	95.5	93.3	88.7	79.8	97.9	88.4	97.2
	Sens.	96.7	90.4	82.5	62.7	96.9	84.4	91.8	94.4	87.7	70.7	96.7	80.8	95.2
	PPV	99.2	92.3	90.1	85.1	92.0	86.6	99.5	92.6	90.7	91.8	99.2	98.4	99.2
AVRT	F_1 score	92.6	92.3	64.6	80.5	90.9	87.4	81.9	93.5	92.7	92.1	92.7	49.0	87.2
	Sens.	88.2	86.4	49.8	67.4	84.1	80.6	70.3	88.5	86.8	86.2	87.3	33.3	79.7
	PPV	100.0	99.7	94.2	99.9	99.2	99.4	99.9	100.0	100.0	99.9	100.0	97.6	100.0
Bi	F_1 score	76.1	73.8	65.6	73.9	73.7	67.8	70.1	74.4	63.1	66.9	75.7	68.4	73.5
	Sens.	72.6	71.4	52.4	63.6	71.9	58.8	56.2	72.3	53.6	50.6	68.6	52.4	67.3
	PPV	88.5	88.9	87.2	93.1	86.9	86.7	97.1	88.5	74.1	98.2	90.8	98.1	91.7
SR	F_1 score	97.6	97.4	91.0	96.0	97.3	96.4	96.1	97.6	95.6	92.2	97.7	89.6	96.9
	Sens.	96.7	96.7	88.4	93.7	96.9	95.0	93.9	97.0	93.9	90.1	97.1	84.5	96.5
	PPV	98.4	98.0	94.4	98.0	97.7	97.8	99.0	98.1	97.4	97.3	98.5	96.5	98.2
VT	F_1 score	79.5	78.5	64.9	66.6	75.5	76.7	61.8	81.4	79.3	79.0	83.9	51.9	80.7
	Sens.	67.2	66.9	49.6	50.9	62.4	63.3	45.4	71.2	69.7	66.0	73.4	36.4	69.4
	PPV	95.6	94.7	91.4	94.6	94.4	94.4	95.9	95.0	93.7	97.6	97.3	90.3	95.2

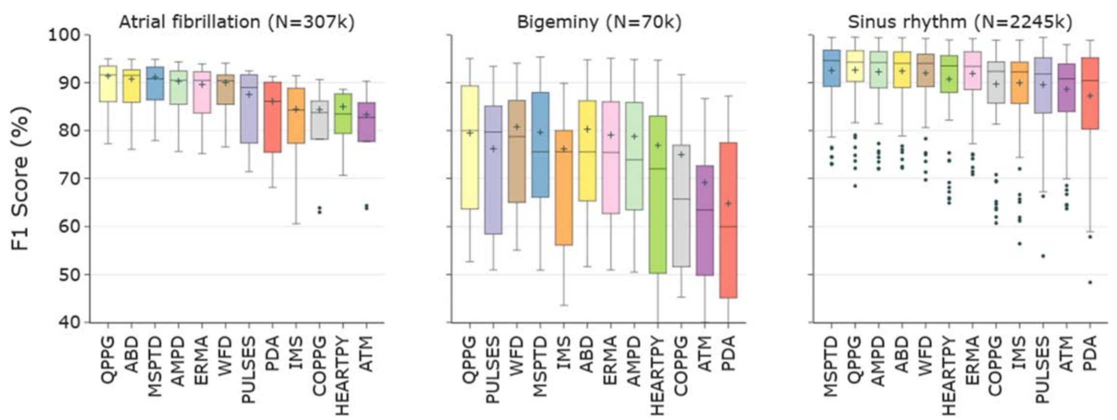

Figure 3. Beat detector performance (F_1 score) by cardiac arrhythmia on the ambulatory dataset. The number of reference beats (N) is indicated in thousands (k) next to each subtitle. Black dots represent outlier subjects, boxes show the median, 1st and 4th quartiles, 10th and 90th percentiles of F_1 scores obtained per subjects while the black cross indicates F_1 score calculated across all subjects. Detectors are ordered by decreasing median of F_1 score.

variable across subjects with some very inaccurate detections. *QPPG* is again top ranked with 83.9% median F_1 score. Bigeminy beats often remain undetected as well depending on the subject. Indeed, bigeminy shows the worst performance, the best detectors being *ABD* and *QPPG* with median F_1 scores of 76.1% and 75.7% respectively. Finally, top ranked beat detectors achieve high performance for both atrioventricular blocks and atrioventricular reentrant tachycardias. *QPPG*, *ABD* and *WFD* get medians F_1 scores between 97.2% and 97.9% for AV blocks. *MSPTD* is the best detector for AVRT with a median F_1 scores of 93.5% closely followed by *PDA*, *QPPG*, *ABD*, *AMPD* and *PULSES* (>92.1%).

3.2.2. Ambulatory dataset

To assess detector performance, only periods characterized by low motion were retained, leading to the exclusion of 695.7 h of signals, which accounted for 51.9% of the total duration. The subsequent assessment of performance was carried out on the remaining 684.5 h of motion-free PPG, as outlined in table 2.

The evaluation of beat detector performance on the ambulatory dataset is shown in figure 3, with comprehensive metrics provided in table 4. On AF segments, *QPPG* is top ranked with a median F_1 score of 91.6%, closely followed by *ABD* and *MSPTD* (>90.8%). Half of the beat detectors perform similarly well on normal sinus rhythm, with *MSPTD* top-ranked at 94.6% and *QPPG*, *AMPD*, *ABD*, and *WFD* achieving medians of F_1 scores superior to 94.0%. The beats of bigeminy are once again poorly detected. *QPPG*, *PULSES* and *WFD* slightly stand out from other detectors with medians of F_1 scores between 80.0% and 78.8%.

4. Discussion

The aim of this study was to assess the performance of several open-source detectors for various types of CAs. Our findings help determine the type of detectors most suitable for the monitoring of CA in every-day life, but also highlight potential limitations in the detection of heartbeats for given CAs.

4.1. Beat detector performance

ABD, *MSPTD* and *QPPG* detectors were consistently ranked among the best detectors for various CAs in both clinical and ambulatory conditions without any failure on specific CAs. These results are in line with the study of Charlton and colleagues (Charlton *et al* 2022), which concluded that *MSPTD* and *QPPG* detectors were performing best within various conditions (hospital, daily-life, emotions, atrial fibrillation, neonates and skin colors). Our analyzes highlighted the superior performance of the *QPPG* beat detector performance in hospital conditions (clinical dataset). This is likely due to the excellent sensitivity of *QPPG*, which is optimal for detecting beats occurring early in the cardiac cycle. It provides a clear advantage for CAs such as atrial and ventricular tachycardias, atrial flutter and AF without a significant loss in PPV, as it is the case with bigeminy for other detectors. This hypothesis was supported by the performance results obtained from the ambulatory dataset. Indeed, *QPPG* was top ranked in an ambulatory setting for CAs showing premature contractions (AF and bigeminy) and was very good for detecting normal sinus beats. *MSPTD* was the best beat detector for sinus rhythm. It showed very good performance during AF as well but was less efficient for detecting bigeminy beats.

Table 4. Beat detector performance on ambulatory dataset. The medians across subjects of F_1 score, sensitivity (Sens.), and positive predictive value (PPV) in percent (%) are detailed for each cardiac arrhythmia: atrial fibrillation (AF), atrial and ventricular bigeminy (Bi) and normal sinus rhythm (SR).

		ABD	AMPD	ATM	COPPG	ERMA	HEARTPY	IMS	MSPTD	PDA	PULSES	QPPG	SWT	WFD
AF	F_1 score	91.5	90.5	82.7	83.7	90.5	83.4	84.3	90.8	86.1	89.0	91.6	68.9	90.4
	Sens.	88.0	87.9	76.5	77.5	86.8	79.3	77.5	89.7	81.6	84.8	90.8	55.5	88.4
	PPV	91.5	90.5	89.2	90.3	90.3	88.0	93.7	90.4	89.3	90.7	90.7	91.1	90.6
Bi	F_1 score	75.5	73.9	63.4	65.7	75.4	72.0	75.5	75.5	60.0	79.7	80.0	65.0	78.8
	Sens.	73.7	65.5	54.1	50.0	73.8	64.3	64.1	74.1	49.5	66.7	71.2	49.4	70.5
	PPV	85.5	82.1	74.1	86.5	84.9	75.9	95.9	85.3	77.7	94.8	88.2	95.1	84.0
SR	F_1 score	94.0	94.2	90.8	92.4	93.4	93.5	92.2	94.6	90.4	91.8	94.3	88.6	94.0
	Sens.	94.7	94.3	88.7	90.3	94.4	91.7	88.6	95.1	88.5	89.9	94.8	81.2	94.3
	PPV	95.4	95.7	94.4	96.1	95.2	96.1	97.1	95.4	93.5	94.7	95.2	97.4	94.6

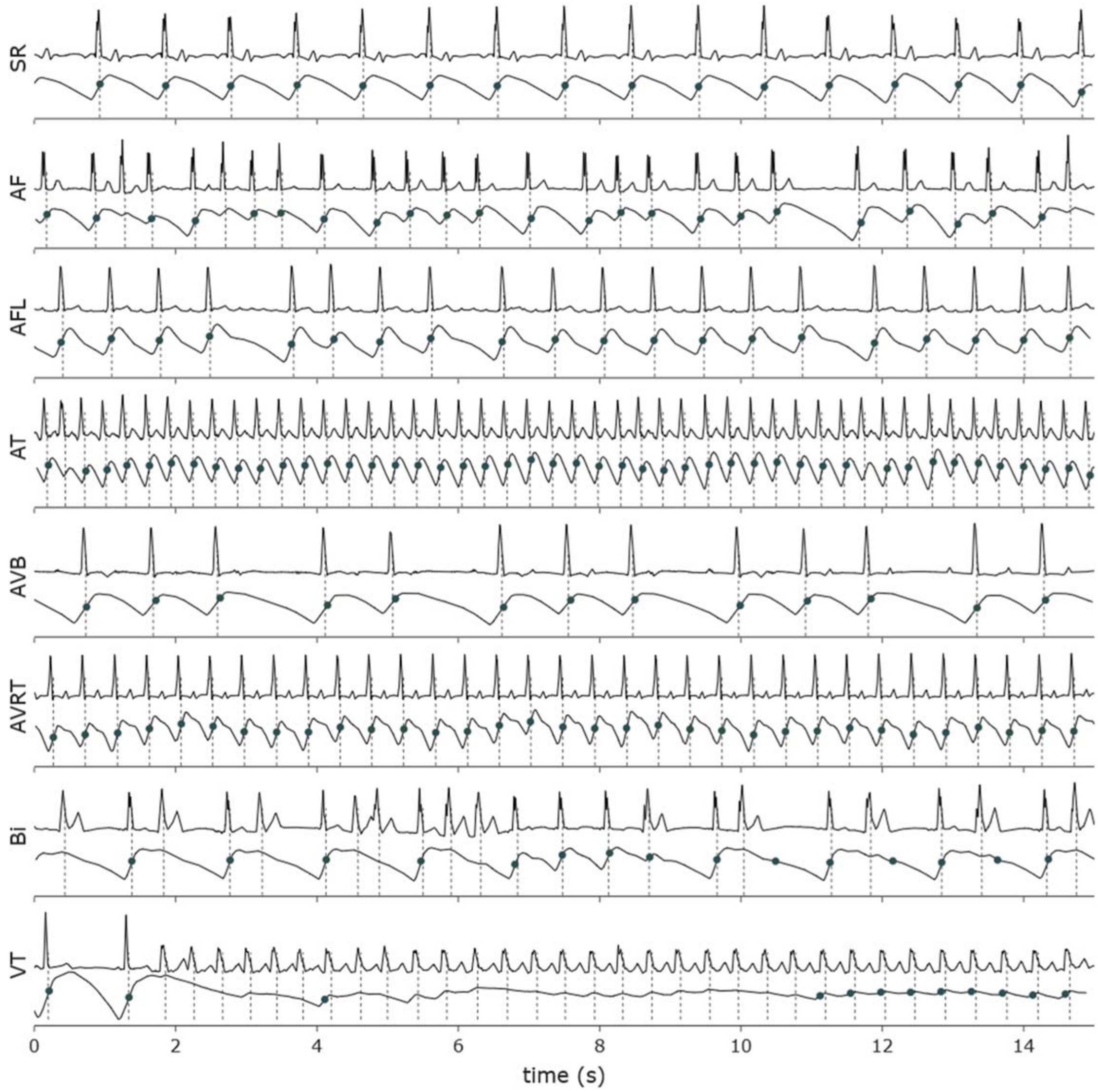

Figure 4. Example signals of the 8 distinct cardiac arrhythmias from the clinical dataset. Each row shows the ECG signal (top curve) and the simultaneous PPG signal (bottom curve). Dotted vertical lines indicate the timing of detected ECG beats, and dots on PPG show the timing of detected beats by the *QPPG* detector. Each row shows the example of one cardiac arrhythmia: sinus rhythm (SR), atrial fibrillation (AF), atrial flutter (AFL), atrial tachycardia (AT), atrioventricular block (AVB), atrioventricular (nodal or not) re-entrant tachycardia (AVRT), bigeminy (Bi), and ventricular tachycardia (VT).

Both *QPPG* and *MSPTD* require low computational efforts and might be suited for embedding in a wearable device. This last point is crucial for the screening of large population with small devices and low battery consumption.

4.2. Limitations of beat detection in cardiac arrhythmias

All beat detectors show lower sensitivity in presence of ventricular tachycardia (VT), one of the fastest CAs. The onset of VT can be very abrupt, which results in PPG waves of decreased amplitude as illustrated in the last row of figure 4. This certainly induces strong differences between outputs of detectors that use different adaptive scaling mechanisms. Slow adaptation to abrupt changes in amplitude, such as those due to onsets of ventricular tachycardia, results in numerous missed detections.

The detection of bigeminy beats in both datasets was particularly poor compared to that of other types of CAs. This is due to premature contractions that occur very early in the cardiac cycle, leading to heartbeats that do not necessarily generate a pressure wave. The resulting changes in the PPG signal—reflecting blood volume changes in the peripheral arteries—are minimal, comparable to that of a dicrotic notch. Examples of bigeminy in figures 4 and 5 show that it is very difficult to detect such premature beats. It is therefore rather an intrinsic physiological limitation for the detection of heartbeats from blood volume variations in the peripheral vascular system. This opinion is in line with the work of Han *et al* (2020), which identified patterns formed by successive IBIs in a Poincaré plot to detect premature contractions. If this method was conclusive for the detection of

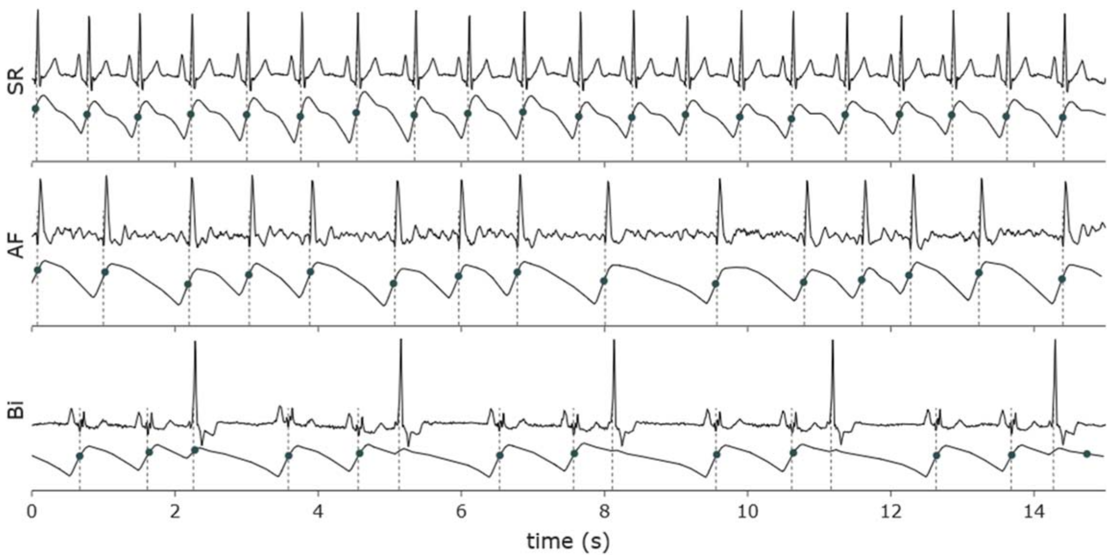

Figure 5. Example signals of the 3 distinct cardiac arrhythmias from the ambulatory dataset. Each row shows the ECG signal (top curve) and the simultaneous PPG signal (bottom curve). Dotted vertical lines indicate the timing of detected ECG beats, and dots on PPG show the timing of detected beats by the *QPPG* detector. Each row shows the example of one cardiac arrhythmia: sinus rhythm (SR), atrial fibrillation (AF) and bigeminy (Bi).

isolated premature contractions, trigeminy and quadrigeminy, it was not the case for the detection of bigeminy with silent premature contractions. However, one possibility would be an in-depth analysis of the PPG waveform, to characterize it as typical bigeminy and deduce that it contains a hidden premature contraction.

4.3. Study limitations

Our work is limited by the inclusion of only five different types of CAs. The number of arrhythmic events of atrioventricular blocks (of any degree) and atrioventricular re-entrant (nodal or not) tachycardias was too small to draw significant conclusions in these two groups of CA. In addition, for the ambulatory dataset, the present analysis was limited to motion-free periods resulting in the rejection of 51.9% of data. In a future study, the influence of motion on the heartbeat detection performance should be investigated in more detail. Finally, ECG-based labelling of CAs have been annotated by one single expert (for the clinical dataset) or software annotations have been corrected by a single cardiologist (for the ambulatory dataset). Annotations that are more reliable could be obtained by systematically involving two cardiologists and keeping only periods of the data where both annotators agree.

5. Conclusion

In this work, we evaluated the performance of thirteen open-source PPG beat detectors in the presence of CAs. *QPPG* showed highest performance in terms of F_1 score. In addition, our evaluation revealed the reduced performances of beat detectors in presence of bigeminy and ventricular tachycardia.

This study provides solid support for selecting a beat detector for continuous monitoring of cardiac arrhythmias in every-day life.

Data availability statement

The data cannot be made publicly available upon publication due to legal restrictions preventing unrestricted public distribution. The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Adrian Luca https://orcid.org/0000-0002-0040-8191

References

Aboy M, McNames J, Thong T, Tsunami D, Ellenby M S and Goldstein B 2005 An automatic beat detection algorithm for pressure signals *IEEE Trans. Biomed. Eng.* **52** 1662–70

Allen J and Kyriacou P A 2021 *Photoplethysmography: Technology, Signal Analysis and Applications* (Academic)

Argüello Prada E J and Serna Maldonado R D 2018 A novel and low-complexity peak detection algorithm for heart rate estimation from low-amplitude photoplethysmographic (PPG) signals *J. Med. Eng. Technol.* **42** 569–77

Basza M *et al* 2023 Photoplethysmography wave morphology in patients with atrial fibrillation *Physiol. Meas.* **44** 045001

Bishop S M and Ercole A 2018 Multi-scale peak and trough detection optimised for periodic and quasi-periodic neuroscience data *Intracranial Pressure & Neuromonitoring XVI* ed T Heldt (Springer International Publishing) (*Acta Neurochirurgica Supplement*) 189–95

Charlton P H *et al* 2022 Detecting beats in the photoplethysmogram: benchmarking open-source algorithms *Physiol. Meas.* **43** 085007

Chugh S S *et al* 2014 Worldwide epidemiology of atrial fibrillation *Circulation* **129** 837–47

Conn N J and Borkholder D A 2013 Wavelet based photoplethysmogram foot delineation for heart rate variability applications *2013 IEEE Signal Processing in Medicine and Biology Symp. (SPMB)* pp 1–5 (https://doi.org/10.1109/SPMB.2013.6736782)

Elgendi M, Norton I, Brearley M, Abbott D and Schuurmans D 2013 Systolic peak detection in acceleration photoplethysmograms measured from emergency responders in tropical conditions *PLoS One* **8** e76585

Gorenek (chair) B *et al* 2017 Device-detected subclinical atrial tachyarrhythmias: definition, implications and management—an european heart rhythm association (EHRA) consensus document, endorsed by heart rhythm society (HRS), asia pacific heart rhythm society (APHRS) and Sociedad Latinoamericana de estimulación cardíaca y electrofisiología (SOLEACE) *EP Europace* **19** 1556–78

Han D *et al* 2020 Premature atrial and ventricular contraction detection using photoplethysmographic data from a smartwatch *Sensors* **20** 5683

Han D *et al* 2022 A real-time PPG peak detection method for accurate determination of heart rate during sinus rhythm and cardiac arrhythmia *Biosensors* **12** 82

Harju J, Tarniceriu A, Parak J, Vehkaoja A, Yli-Hankala A and Korhonen I 2018 Monitoring of heart rate and inter-beat intervals with wrist plethysmography in patients with atrial fibrillation *Physiol. Meas.* **39** 065007

Hindricks G *et al* 2021 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the european association for cardio-thoracic surgery (EACTS): the task force for the diagnosis and management of atrial fibrillation of the european society of cardiology (ESC) developed with the special contribution of the european heart rhythm association (EHRA) of the ESC *Eur. Heart J.* **42** 373–498

Jeanningros L *et al* 2022 Pulse wave analysis of photoplethysmography signals to enhance classification of cardiac arrhythmias *2022 Computing in Cardiology (CinC)* 49

Karlen W, Ansermino J M and Dumont G 2012 Adaptive pulse segmentation and artifact detection in photoplethysmography for mobile applications *2012 Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society* pp 3131–4

Khurshid S *et al* 2018 Frequency of cardiac rhythm abnormalities in a half million adults *Circ. Arrhythm Electrophysiol.* **11** 006273

Lázaro J, Gil E, Vergara J M and Laguna P 2014 Pulse rate variability analysis for discrimination of sleep-apnea-related decreases in the amplitude fluctuations of pulse photoplethysmographic signal in children *IEEE J. Biomed. Health Inform.* **18** 240–6

Lemay M *et al* 2020 Applications of optical cardiovascular monitoring *Wearable Sensors* ed E Sazonov (Elsevier) 2nd ednpp 487–517

Orphanidou C, Bonnici T, Charlton P, Clifton D, Vallance D and Tarassenko L 2015 Signal-quality indices for the electrocardiogram and photoplethysmogram: derivation and applications to wireless monitoring *IEEE J. Biomed. Health Inform.* **19** 832–8

Proença M *et al* 2019 Pulse wave analysis techniques *The Handbook of Cuffless Blood Pressure Monitoring: A Practical Guide for Clinicians, Researchers, and Engineers* ed J Solà and R Delgado-Gonzalo (Springer International Publishing) pp 107–37

Rho R W and Page R L 2005 Asymptomatic atrial fibrillation *Prog. Cardiovasc. Dis.* **48** 79–87

Scholkmann F, Boss J and Wolf M 2012 An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals *Algorithms* **5** 588–603

Shin H S, Lee C and Lee M 2009 Adaptive threshold method for the peak detection of photoplethysmographic waveform *Comput. Biol. Med.* **39** 1145–52

Tsao C W *et al* 2023 Heart disease and stroke statistics—2023 update: a report from the american heart association *Circulation* **147** e93–621

Vadrevu S and Manikandan M S 2019 A Robust pulse onset and peak detection method for automated PPG signal analysis system *IEEE Trans. Instrum. Meas.* **68** 807–17

Väliaho E-S *et al* 2019 Wrist band photoplethysmography in detection of individual pulses in atrial fibrillation and algorithm-based detection of atrial fibrillation *EP Europace* **21** 1031–8

van Gent P, Farah H, van Nes N and van Arem B 2019 HeartPy: a novel heart rate algorithm for the analysis of noisy signals *Transportation Res.*F **66** 368–78

Zeppenfeld K *et al* 2022 2022 ESC guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: developed by the task force for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death of the european society of cardiology (ESC) endorsed by the association for european paediatric and congenital cardiology (AEPC) *Eur. Heart J.* **43** 3997–4126

Zong W, Heldt T, Moody G B and Mark R G 2003 An open-source algorithm to detect onset of arterial blood pressure pulses *Comput. Cardiol.* **2003** 259–62