



# An Open-Access Arrhythmia Database of Wearable Electrocardiogram

Qin Shen<sup>1</sup> · Hongxiang Gao<sup>2</sup> · Yuwen Li<sup>2</sup> · Qi Sun<sup>2</sup> · Minglong Chen<sup>1</sup> · Jianqing Li<sup>2,3</sup> · Chengyu Liu<sup>2</sup>

Received: 12 May 2020 / Accepted: 6 July 2020 / Published online: 22 July 2020  
© The Author(s) 2020

## Abstract

**Purpose** Heart arrhythmias result from any disturbance in the rate, regularity, and site of origin or conduction of the cardiac electric pulse. Sporadic and underappreciated characteristics make diagnosis less timely, leading to stroke, heart failure, or even sudden death. Wearable electrocardiogram (ECG) devices are gradually becoming the main trend of intelligent diagnosis alongside the improvement of information transmission and computation power of the hardware. Therefore, a database for arrhythmia detection was planned to construct.

**Methods** Collaborating with hospital, this study presents a bipolar limb two-lead wearable device by collecting a large amount of real-time data and giving rise to readily available ECG databases.

**Results** In total, the database contains 2000 30-s recordings of sinus, atrial and ventricular arrhythmias collected from more than 200 voluntary patients who had been diagnosed with heart diseases, ranging in age from 18 to 82. Meanwhile, manual annotations by cardiologists were proposed to benefit and instruct non-medical researchers to design the algorithm reasonably.

**Keywords** Electrocardiogram (ECG) · Wearable · Database · Arrhythmia · Annotations

## 1 Introduction

Cardiac arrhythmia is a large class of cardiovascular diseases (CVDs) in clinical, and it always endangers human health [1, 2]. There are many kinds of arrhythmias, which almost occupy more than half of the patients in the diagnosis of the surface electrocardiogram (ECG). The painless and unobvious natures of silent myocardial ischemia (SMI) are responsible for many sudden deaths, which creates the need for long-term monitoring of the specific patient [3]. Continuous monitoring of patients with certain CVDs, including

post-myocardial infarction (MI), tachycardia, arrhythmias, left ventricular dysfunction, etc., gives benefits to assess the condition of cardiovascular health effectively [4].

For the implementation of dynamic monitoring, the primary task is to collect ECG in real time. Nowadays, there are many approaches to measure/record ECG. Da Silva et al. [5] provided a taxonomy of state-of-the-art ECG measurement methods: *in-the-person*, *on-the-person* and *off-the-person*. The majority of devices used for ECG measurements belong to the *on-the-person category*. ECG signal can be acquired from multiple configurations of sensors on the body and is easily affected by body positions, reflecting health information of different parts of the heart. In 1961, Holter [6] introduced techniques for continuous recordings of ECG from ambulatory subjects over periods of many hours; the long-term ECG (Holter recording), typically with a duration of 24 h, has since become the standard technique for observing transient aspects of cardiac electrical activity [7]. Recorded signals are then analyzed off-line using dedicated diagnostic systems [3, 8–11]. As many arrhythmias are characterized by intermittent and short-lasting episodes that are not usually found early and therefore, have a very poor detection rate [12]. The diagnostic mechanism of Holter limits its applications in homecare which has raised extremely high demands in early and quick diagnosis of ECG arrhythmia. Therefore,

✉ Minglong Chen  
chenminglong2001@163.com

Jianqing Li  
ljq@seu.edu.cn

Chengyu Liu  
chengyu@seu.edu.cn

<sup>1</sup> The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

<sup>2</sup> The State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing, China

<sup>3</sup> School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, China

wearable real-time ECG for long-term monitoring has potential to be the most useful non-invasive device for assessing cardiac health.

De Chazal [13] demonstrated that similar effectiveness for ECG arrhythmia classification can be obtained at a lesser computational cost when using only one-lead, compared with methods using multiple leads [14]. Bipolar limb lead is part of the most utilized method which has the ability to display three of the most important waves: P wave, QRS complex and T wave. These waves correspond to the field induced by the electrical phenomena occurring on the heart surface, denominated atrial depolarization (P wave), ventricular depolarization (QRS complex) and repolarization (T wave) [15]. The patterns provoked by arrhythmias can deeply change these waves. Therefore, a wearable device with analog limb leads (lead I, II, III) is mainly used to monitor arrhythmias highlighted in this paper, such as, (1) sinus arrhythmia: sinus bradycardia, sinus tachycardia and sinus arrest; (2) atrial arrhythmias: premature atrial contractions (PACs), non-sustained/sustained atrial tachycardia (NSAT/SAT), and atrial fibrillation (AF); (3) ventricular arrhythmias: premature ventricular contractions (PVCs) and non-sustained/sustained ventricular tachycardia (NSVT/SVT).

Unfortunately, one of the fundamental problems associated with measuring dynamic ECG is the decrease in signal quality due to the unexpected environmental disturbances [16, 17]. Artifacts from physiological and nonphysiological sources are common. Wearable ECG monitoring devices are commonly based on single-lead measurements with dry metal plate [18], resulting in much smaller signal amplitude and noisy waveforms compared to wet adhesive electrodes. In the field of cardiology, there is an urgent need for such databases as they will play a great role in allowing manufacturers to design systems and hospitals to measure the performance of their systems against manufacturers' claims. In addition, several standard ECG databases are available to evaluate algorithms for different test purposes [19]. The most commonly used databases on published researches for arrhythmia are the MIT-BIH Arrhythmia Database, QT Database, CSE Database, and AHA Database [20]. Although classical they are, the characteristic of non-wearable makes them not perfect for dynamic automatic analysis algorithm designing [21–23].

The aim of this paper is to present two ECG databases suitable for development and testing of ECG classification methods. Signal quality database contains three different degrees of signal quality. Arrhythmias database contains kinds of subcategories, corresponding to various arrhythmias mode.

## 2 The Structure of Database

### 2.1 Data Acquisition

In China, patient's personal information and disease information are stored in the servers and hard disk devices of their respective hospitals. Subject to confidentiality agreement of collection equipment manufacturers and hospital ethics permission, retrieving existing recording is not desirable. Therefore, self-collection would be made the only feasible method.

Under the permission of the ethical certification of the First Affiliated Hospital of Nanjing Medical University, Southeast University jointly carried out the work of collecting ECG data with a wearable wireless ECG monitor, which has passed the FDA certification. The recordings acquired by the device are all 6-lead ECGs, digitized at 400 samples per second per channel with 12-bit resolution over a frequency response bandwidth of 0.05–40 Hz. More than 200 individuals with arrhythmias were tested, aged between 18 and 82. All subjects were trained to wear the wireless ECG monitor without other's help and retain at least 24 h even to 72 h to cover all possible onsets of arrhythmias. According to the agreement we reached, the data can be freely used, while the patient's identification should be kept anonymous. The specific operation process is shown in Fig. 1.

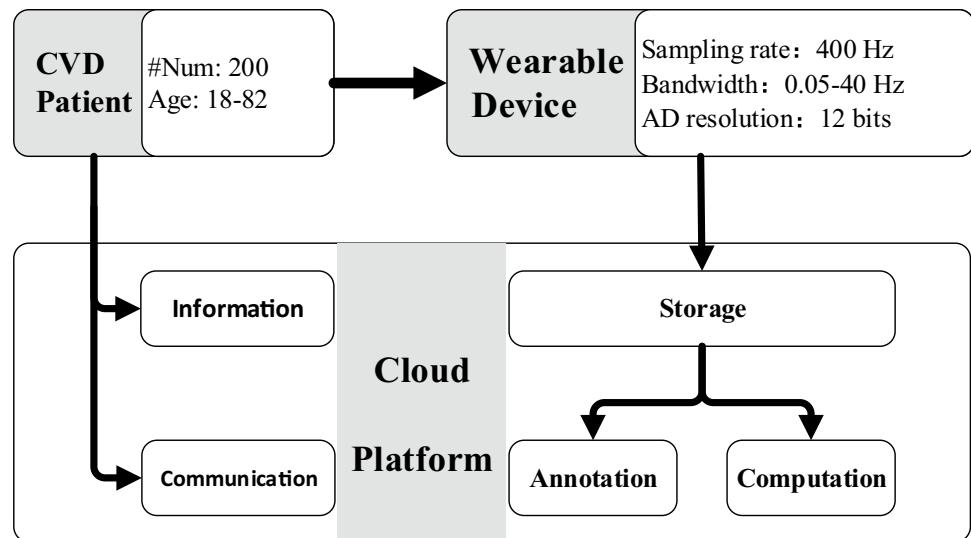
### 2.2 Annotation Workflow

For many bio-signal processing applications, the performance of algorithms and systems must be evaluated against reference or 'gold standard' annotations. It is common to have one or multiple experts' annotators to evaluate the data when this ground truth is not readily available. Thus, an annotation platform was developed by co-operation of automatic classification algorithms and three cardiologists. First, a huge amount of ECG recordings stored in cloud platform were uploaded to the annotation platform in the standard ECG drawing format. Then, an automatic step was applied to generate coarse annotation with commonly used algorithms. After that, two clinical cardiologists independently corrected the automatic labels. The third expert finally checked the results and identified the labels with different opinions and made a determination.

### 2.3 ECG Cloud Platform

The construction of an annotated open access database is a long-lasting task. Thus, it is wise to operate with an

**Fig. 1** The flow chart of data acquisition and processing



assistant platform. Herein, a cloud platform composed of five parts as followings was developed.

- (1) Information management platform: dealing with the personal information, such as account information, patient's identification medical history of doctors and users;
- (2) Communication management platform: transmitting real-time abnormal information generated by wearable monitor and feedbacks from doctor;
- (3) Storage management platform: unified management of raw data storage, modification, verification, search and forwarding, etc.;
- (4) Crowd-sourcing labeling platform: unified management of the ECG labeling works. Recordings stored in the storage clouds with no more than three different annotations will be released to this platform in a form of fee-for-service-based, to attract cardiologists and experts to annotation work. Double-blind measures are taken among experts to improve the reliability of labeling results. And yet, after an automatic annotation comparison, controversial issue will re-enter the label platform, others will be stored in another storage clouds;
- (5) Cloud computing platform: an automatic ECG analysis framework to make diagnosis, draw conclusions and generate diagnostic reports. A preprocessing step was brought in for denoising, mainly including power frequency interference, baseline drift, myoelectricity interference, motion artifacts, and electrode contact noise [24]. The annotation of QRS complex is the most important task, where an algorithm based on the Pan and Tompkins has a preferable performance [23, 25, 26]. Features are used to represent patterns with minima loss of important information. Then a neural

network classifier based on CNN and LSTM is chosen for the classification procedure [27, 28]. Eventually, a long-term ECG report will be exported in a standard format with the results of physician-assisted diagnosis.

## 2.4 Data Schema

Committed to signal quality assessment (SQA) and arrhythmia analysis on rhythm changes, two separate databases were constructed elaborately by several clinical cardiologists based on their diagnostic experiences and the criterion of ECG diagnosis. Signal quality was divided into three levels empirically, more detailed rules could be seen in Sect. 2.4.1. The arrhythmia database contains three categories, including sinus rhythm (category 'N'), atrial rhythm (category 'A'), and ventricular rhythm (category 'V'), and each one is composed by a variety of arrhythmias specifically interpreted in Sect. 2.4.2.

### 2.4.1 Signal Quality Database

In recent years, analysis and evaluation of various physiological signals, especially ECG signal quality, have been a hot topic [29–31]. The PhysioNet/Computing in Cardiology Challenge (CinC) 2011 [32] aimed to develop an efficient algorithm able to run in near real-time within a mobile phone [17], which could provide useful feedbacks to a layperson in the process of acquiring ECG recordings. Due to the poor signal quality caused by the dry electrodes [33, 34], SQA is considered as a main target. Besides, fewer leads means that there is a high probability of synchronous contaminations between different channels, preventing joint diagnosis with multi-leads. Herein, a specialized database contains 300 recordings lasting for 10 s was designed (see Table 1), divided into three categories of signal quality:

**Table 1** Specification for signal quality division

Category	Symbol	Definition
Good	A	Signal with apparent P-QRS-T morphologies Signal with slightly baseline drift or transient artefacts
Medium	B	A good recording contaminated severely in a narrow window A good recording with one or a few missing signals
Poor	C	A poor recording that may be interpretable with difficulty Signal usefulness in clinical applications (maybe caused by misplaced electrodes, poor skin-electrode contact)

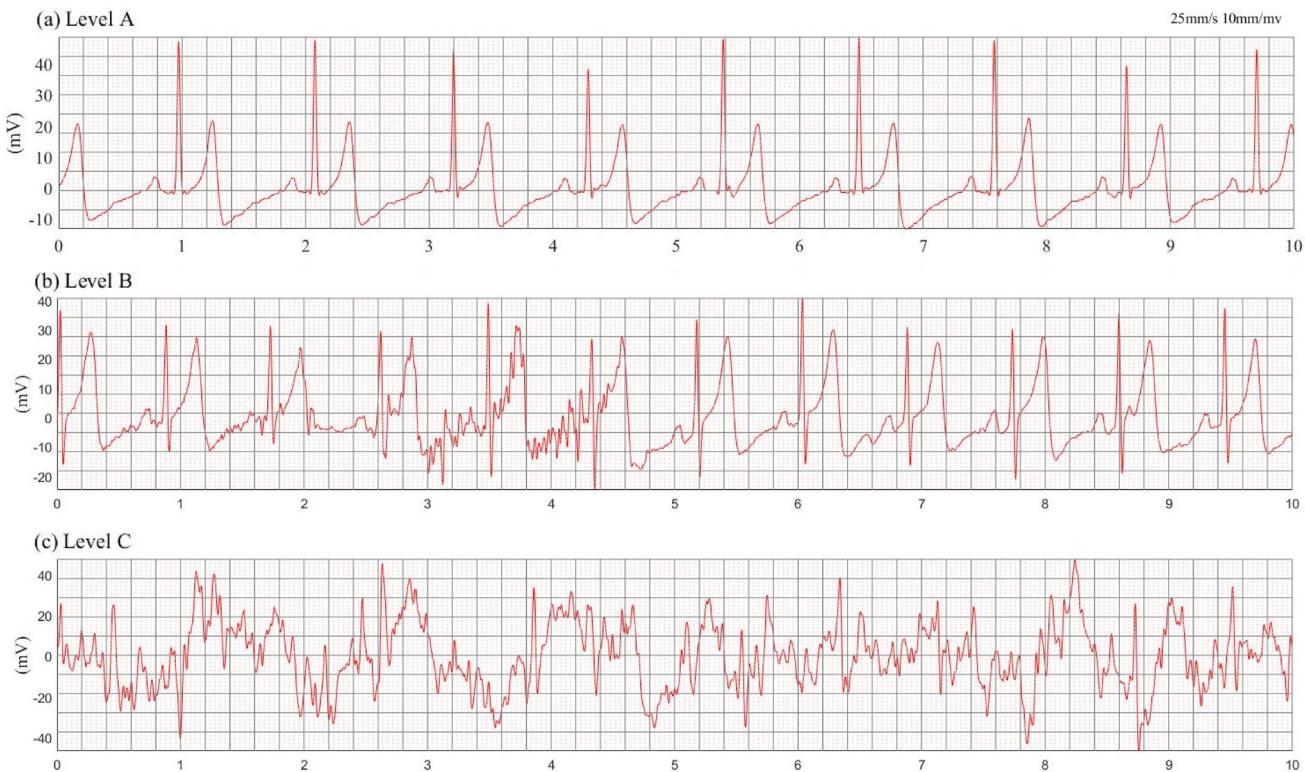
good signal quality (Type ‘A’), medium signal quality (Type ‘B’) and poor signal quality (Type ‘C’). Typical examples of ECG waveforms are shown in Fig. 2.

#### 2.4.2 Arrhythmias Database

Based on the above classification rule of signal quality, the arrhythmia database was set up by screening lead I ECG signal of type ‘A’. The ANSI/AAMI EC57:1998/(R) 2008 standard [35] specifies that records of patients using pacemakers should not be considered. In addition, segments of data containing ventricular flutter or fibrillation (VF) were also excluded from the analysis. Finally, according to the classification of rate, mechanism or duration, a database

containing three major categories and 18 subcategories was completed. Noted that, subject to the small sample of heart disease, several sub-types only contain a few recordings (see Table 2, followed by the corresponding sample image and the diagnostic mode). Sample pictures are presented in the form of standard ECG drawings lasting for 10 s extracted from the complete signal (Figs. 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14).

**2.4.2.1 Sinus Rhythm** For normal sinus rhythm, ECG appears periodically with a stable PR interval shorter than 0.2 s. Normal P waves in lead II usually show upright wave in morphology and consistent lasting time in duration. Bradycardia is the situation that resting heart rate (HR) less



**Fig. 2** Typical examples of the three signal quality categories. **a** Good signal quality; **b** medium signal quality; **c** poor signal quality. ECGs of good signal quality have clear and distinct P-QRS-T morphologies accompanied by slightly noise or artifacts occasionally. ECGs of

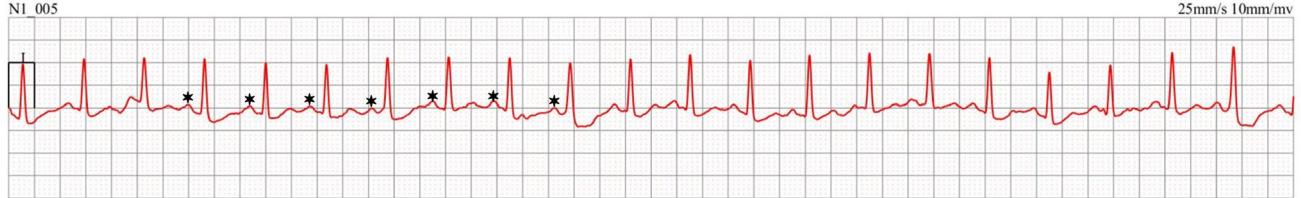
medium signal quality show obvious rhythmical characteristics, but with distinct signal noises and cannot be used for morphology diagnosis. ECGs of poor signal quality are totally unacceptable recordings, due to the large proportion of noise

**Table 2** The structure of the arrhythmias database

Class	Sub-type	Individual distribution			#Recordings	ECG characteristics
		Sex Male/female	Age (min–max)	#Num		
N	Sinus tachycardia (N1)	32/11	26–79	14	N1_(001-043)	Upright, consistent, and normal in morphology and duration of P waves in analog leads I and II Regular PP interval, and PR interval is between 120 and 200 ms HR less than 60 bpm
	Sinus bradycardia (N2)	73/130	37–76	20	N2_(001-203)	Upright, consistent, and normal in morphology and duration of P waves in analog leads I and II Regular PP interval, and PR interval is between 120 and 200 ms HR greater than or equal to 100 bpm
	Sinus pause (N3)	0/4	69–72	2	N3_(001-004)	Upright, consistent, and normal in morphology and duration of P waves in analog leads I and II Prolonged absence of sinus node activity (absent P waves)> 2 s No common multiple relationship between the longest PP interval and basic sinus PP interval
A	Single PAC (A11)	182/118	37–79	30	A11_(001-300)	P' waves, as opposed to sinus P waves, appear before the normal narrow QRS with normal axis
	Bigeminy (A12)	1/35	53–67	3	A12_(001-036)	P'R interval is equal or greater than PR interval which is between 120 and 200 ms
	Trigeminy (A13)	6/50	63–67	3	A13_(001-056)	Atrial premature beats are associated with an incomplete compensatory pause, meaning that the sinus RR interval between the preceding and following atrial premature beat is less than two complete cycles
	Atrial couplet and triplet (A2)	44/56	42–76	11	A2_(001-100)	If an atrial premature beat follows every sinus beat, atrial bigeminy is said to exist
	NSAT/SAT (A3)	62/38	37–76	9	A3_(001-100)	Two consecutive P' waves appear before the normal narrow QRS, followed by a long compensatory interval
	AF (A4)	256/144	32–76	18	A3_(001-400)	No less than three PACs appeared continuously and followed by a compensatory interval Non-sustained atrial tachycardia lasts no longer than 30 s, otherwise known as sustained atrial tachycardia Absence of distinct P waves and regular RR intervals Abrupt onset of an irregular narrow-complex tachycardia Small and irregular baseline wave with variable shape and amplitude, called F waves Frequency of F waves is about 350–600 bpm

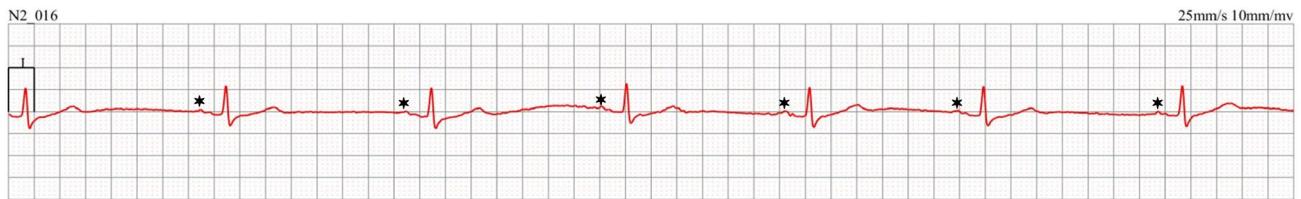
**Table 2** (continued)

Class	Sub-type	Individual distribution			#Recordings	ECG characteristics
		Sex	Age	#Num		
		Male/female	(min–max)			
V	Single PVC (V11)	99/201	25–82	35	V11_(001-300)	Occurs early on or shortly after the T wave of the preceding sinus beat, and is associated with a wide QRS complex; Accompanied by secondary ST-T change (ST segment down and T wave inversion); Followed by a long interval. The PP interval around the premature ventricular complex is twice the underlying sinus rate (PP interval), called a full compensatory pause; If a ventricular premature beat follows every sinus beat, ventricular bigeminy is said to exist.
	Bigeminy (V12)	81/19	35–53	7	V12_(001-100)	
	Trigeminy (V13)	42/58	32–60	9	V13_(001-100)	
	Interpolated (V14)	18/82	32–76	13	V14_(001-100)	
	Ventricular Couplet and Triplet (V2)	6/14	37–79	7	V2_(001-20)	Two consecutive wide and deformed QRS complexes, followed by a long compensatory interval.
	NSVT/SVT (V3)	1/0	68	1	V3_001	No less than three wide and deformed QRS complexes appeared continuously, followed by a long compensatory interval; Non-sustained ventricular tachycardia lasts no longer than 30 s, otherwise known as sustained ventricular tachycardia.



**Fig. 3** Sample graph of sinus tachycardia, where the black asterisks represent the position of sinus P wave. HR equals to 120 bpm, and PR interval is 0.16 s. The PR interval tends to shorten as heart rate

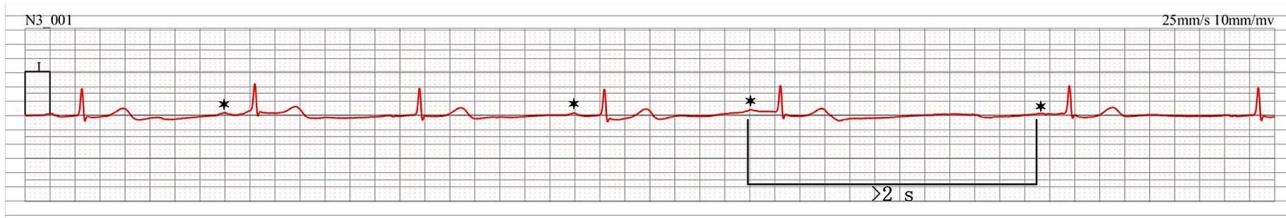
increases because of increased sympathetic tone speeds conduct through the atrioventricular (AV) node



**Fig. 4** Sample graph of sinus bradycardia, where the black asterisks represent the position of sinus P waves. HR equals to 40 bpm

than 60 bpm, on the contrary, the situation that resting HR greater than 100 bpm is called tachycardia. The condition is referred to as sinus pause if only one or two beats are missed and sinus arrest if more than two beats are missed. In

this situation, the sinus node ceases to generate the electrical impulses for a variable period of time.

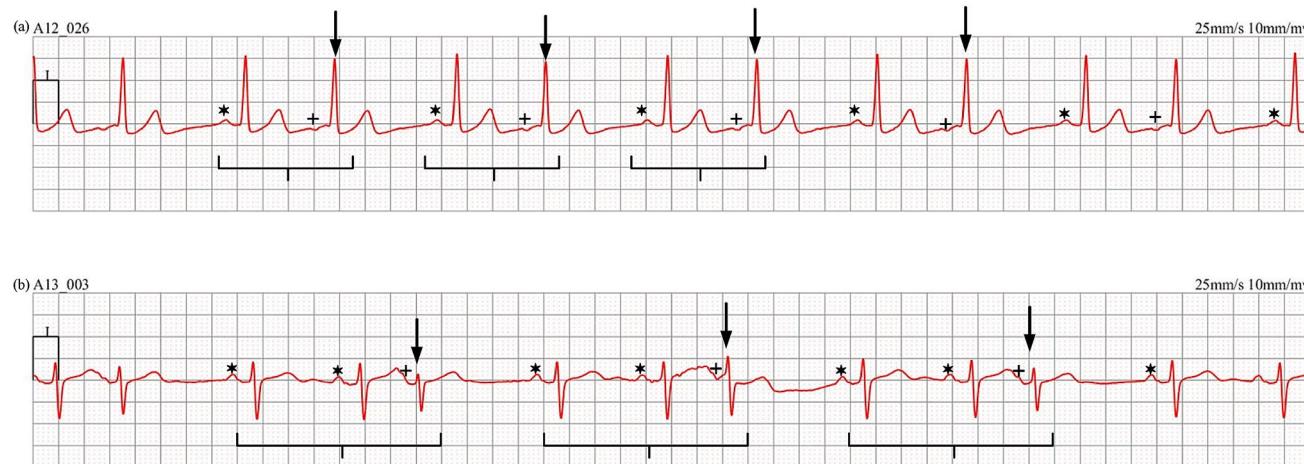


**Fig. 5** Sample graph of sinus pause, where the black asterisks represent the position of sinus P waves. There is a 2.2-s interval before the heart resume between the 5th and 6th peak, indicating a sinus pause



**Fig. 6** Sample graph of single PAC, where the black asterisks ‘\*’ represent the position of sinus P waves, while the black ‘+’ locates in the position of atrial P wave, and the black arrow points to atrial QRS complex. An abnormal (non-sinus) P’ wave (labeled with ‘+’) is followed by a QRS complex. The P’ wave typically has a different

morphology to the sinus P waves (labeled with black asterisks ‘\*’) due to the premature beat initiates outside the sinoatrial (SA) node. As shown in the Fig. 5, the SA node would reset when PAC reaches and depolarizes it, resulting in an incomplete compensatory before the next sinus beat arrives

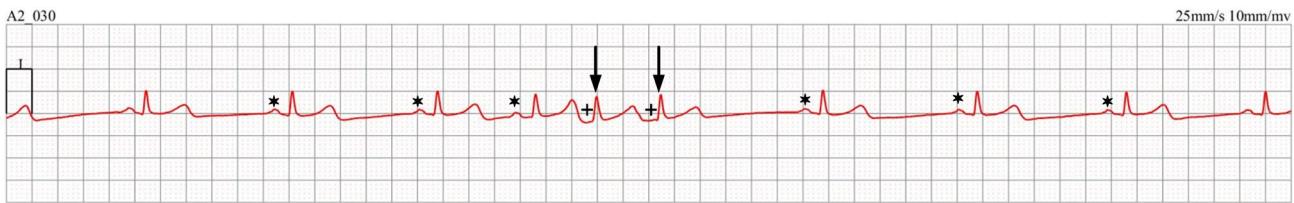


**Fig. 7** Sample graph of **a** atrial bigeminy and **b** trigeminy, where the black asterisks represent the position of sinus P waves, the black ‘+’ locates in the position of atrial P wave, and the black arrows point to atrial beats. PACs often occur in repeating patterns, and bigeminy means every other beat is a PAC, while every third PAC occurring means trigeminy. PACs arise from the premature atrial ectopic pacemaker. The low atrial ectopic pacemaker activates the atria retrogradely, producing an inverted P wave with a relatively short PR interval (see P waves in A12\_026). PACs arriving early in the cycle may be conducted aberrantly, usually with a RBBB morphology, as the right bundle branch has a longer refractory period than the left (see QRS wave in A13\_003). Furthermore, all abnormal P waves keep the same morphology, which means unifocal premature atrial contractions occurred

rogradely, producing an inverted P wave with a relatively short PR interval (see P waves in A12\_026). PACs arriving early in the cycle may be conducted aberrantly, usually with a RBBB morphology, as the right bundle branch has a longer refractory period than the left (see QRS wave in A13\_003). Furthermore, all abnormal P waves keep the same morphology, which means unifocal premature atrial contractions occurred

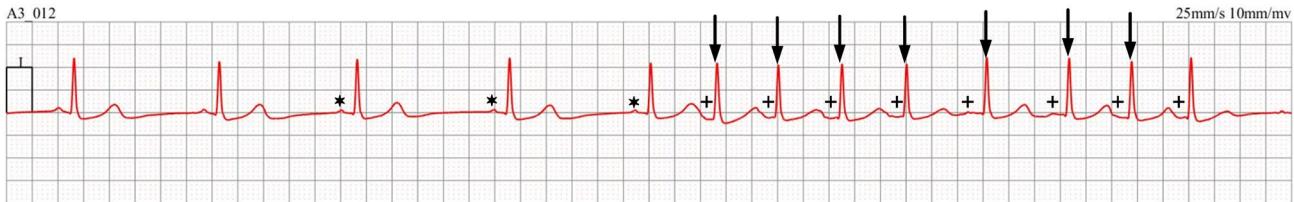
**2.4.2.2 Atrial Rhythm** Premature beat initiates outside the sinoatrial node. Atrial premature beat, also called PAC, is ectopic beat that originates in the atria. Typically, atrial impulse propagates normally through the AV node into the cardiac ventricles, resulting in a normal or narrow

QRS complex. Atrial premature beat is associated with an incomplete compensatory pause, meaning that the interval between the preceding and following sinus beats is less than twice the complete cycle.



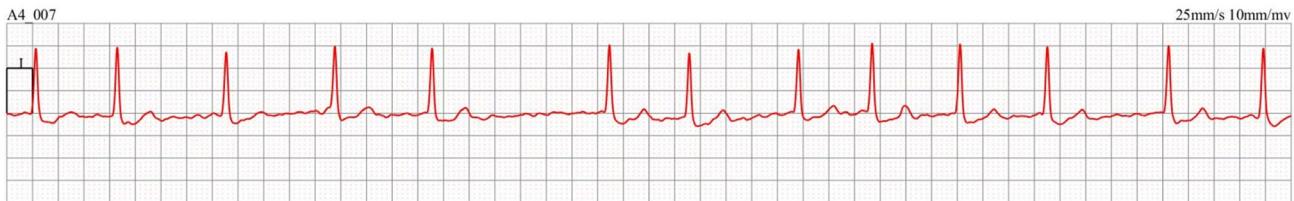
**Fig. 8** Sample graph of atrial couplet, where the black asterisks represent the position of sinus P waves, the black ‘+’ locate in the position of atrial P waves, and the black arrows point to atrial QRS

complexes. PACs may occur frequently or sporadically. Two PACs occurring consecutively are referred to as an atrial couplet

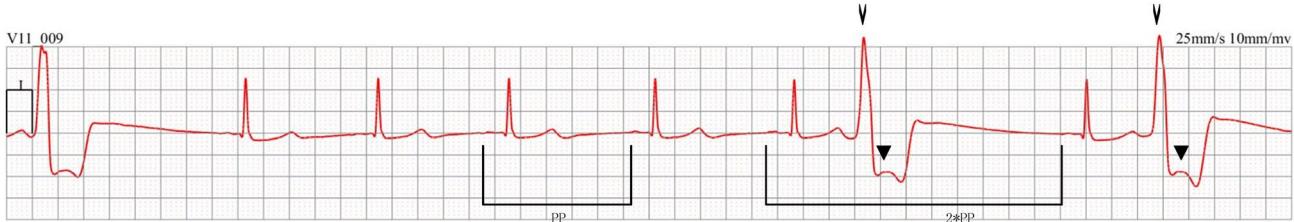


**Fig. 9** Sample graph of paroxysmal atrial tachycardia, where the black asterisks represent the position of sinus P waves, the black ‘+’ locate in the position of atrial P wave, and the black arrows point to atrial QRS complexes. ECG strip between the 1st and 5th QRS complexes is featured with a regular sinus rhythm at 62 bpm. There is a

narrow QRS complex tachycardia at 110 bpm between 6th and 13th QRS complexes; each QRS complex labeled with black arrow is preceded by an abnormal P wave (black ‘+’). Atrial P wave morphology is abnormal when compared with sinus P wave due to ectopic origin



**Fig. 10** Sample graph of atrial fibrillation with irregular baseline undulation. Irregularly irregular rhythm; absence of an isoelectric baseline and P waves. Coarse fibrillatory waves are visible with a high rate

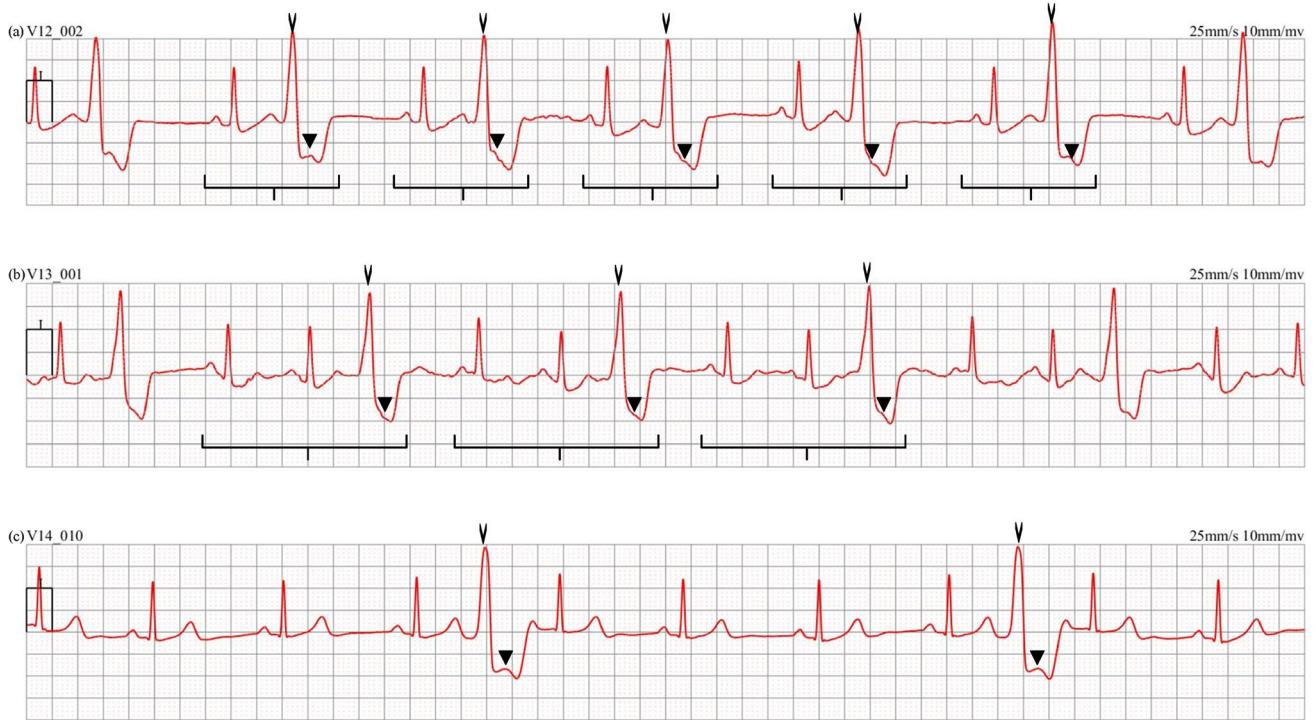


**Fig. 11** Sample graph of single PVC where the black ‘V’ represents the position of PVCs and the black inverted triangles present abnormal P waves. Sinus rhythm with PVCs of same morphologies (‘V’). With a full compensatory pause, the next normal beat arrives after an interval that is equal to double the preceding RR interval; Note the

appropriately discordant ST segment/ T waves. The ectopic impulse is conducted retrogradely through the AV node, producing atrial depolarization. Inverted P waves labeled with inverted triangles (‘retrograde P wave’) occurring after the QRS complexes

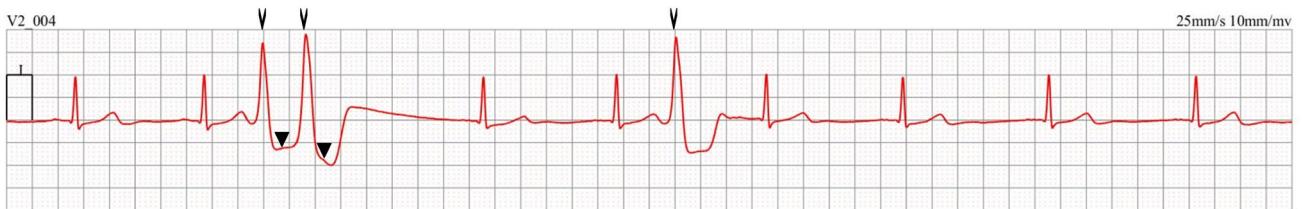
Single PAC clearly manifests a regular underlying rhythm, but there is a premature beat which can be identified by irregular P wave with different size and shape. Atrial

bigeminy is an abnormal pulse that each sinus beat is coupled to a premature atrial complex followed by a slight post-ectopic pause. PACs may occur frequently or sporadically.



**Fig. 12** Sample graphs of **a** bigeminy, **b** trigeminy and **c** interpolated PVC, where the black ‘V’ represent the position of PVCs and the black inverted triangles present abnormal P waves. Ventricular bigeminy means every other beat is a PVC, and every third PVC is

called ventricular trigeminy. In the third figure of Fig. 11, both two PVCs are interpolated, because they are sandwiched between two normal sinus beats without the compensatory pause that typically follows a PVC



**Fig. 13** Sample graph of ventricular couplets, where the black ‘V’ represents the position of PVCs and the black inverted triangles present abnormal P waves. PVC couplet is two PVCs in a row, the same morphology meaning they are unifocal. The 7th beat is an interpolated PVC.



**Fig. 14** Sample graphs of non-sustained ventricular tachycardia, where the black ‘V’ represent the position of PVCs. The four consecutive ventricular beats keep a HR of 150 bpm

Two PACs occurring consecutively are referred to an atrial couplet. Paroxysmal atrial tachycardia has a high regular rate of about 140–250 bpm. AF has an atrial rate of more than 400 bpm and is distinguishable due to its haphazardly irregular ventricular rate.

**2.4.2.3 Ventricular Rhythm** PVCs are premature ectopic beats arising from the right ventricle (RV) or left ventricle (LV) that can occur in a variety of patterns and can occasionally cause uncomfortable symptoms. PVCs are characterized by premature and bizarre shaped QRS complexes that usually last long (typically > 120 ms). These complexes are not preceded by a P wave, and the T wave is usually large and oriented in a direction opposite the major deflection of the QRS. Ventricular tachycardia (VT) is an ectopic ventricular rhythm with wide QRS complex (120 ms or greater), rate faster than 100 bpm, lasting for at least three beats that spontaneously resolves in less than 30 s.

## 2.5 Database Scale Expectation

The arrhythmias database would construct a complete standard annotated ECG database in the future. Every year the completed annotation work will be released to public, more information could be seen in this URL (<http://www.shelab.cn/Data>).

## 3 Discussion and Further Work

The size and diversity of databases play a more important role in machine learning than the learning algorithm and employed techniques. One of the obstacles in the research on fully automatic analysis in ECG is the insufficient quantity of available databases. Standard ECG database is created for validating algorithms and testing instruments on feature detection and disease diagnosing. ECG databases published in the PhysioNet platform basically collected with high quality in clinical environment, which is the first choice for major research. Researches of wearable devices proceed by painful lurches compared to many classical methods in literature, due to the unbalanced development of traditional databases and dynamic databases. The wearable monitoring of non-emergency arrhythmias raises a high demand on dynamic databases of signal quality and arrhythmias. This study organized database classification and annotation work carefully and put forward it freely. Signal quality database contains 300 recordings lasting for 10 s, sampled at 400 Hz, evenly divided into good signal quality, medium signal quality and poor signal quality. Arrhythmias database consists of 2000 single-channel arrhythmias ECG records, and each is 30-s long and sampled at a rate of 400 Hz. The database contains three categories: sinus arrhythmia, atrial arrhythmia

and ventricular arrhythmia. Such a database helps greatly in training algorithms of annotation and classification.

Unfortunately, only few data were obtained in several category of arrhythmias database, due to insufficient diversity of diseases among subjects and the rarity of particular disease. More patients with particular heart disease should be tested to cover the existing deficiencies. More cardiologists are needed because the recordings acquired in dynamic conditions are always very long (24 to 72 h), causing a formidable task with the beat-by-beat annotation. Demographics distributions are another important issue in disease prediction that uncommonly available. For all this, an ECG labeling crowd-sourcing platform will be released to accelerate the process. Furthermore, we look forward to more disclosure of annotated data and novel advances in existing databases.

**Acknowledgements** This work was supported by the Distinguished Young Scholars of Jiangsu Province (BK20190014), the National Natural Science Foundation of China (81871444), the Key Research and Development Programs of Jiangsu Province (BE2017735). We thank the supports from the State Key Laboratory of Bioelectronics in China and the Southeast-Lenovo Wearable Heart-Sleep-Emotion Intelligent monitoring Lab.

## Compliance with Ethical Standards

**Conflict of interest** There is no conflict of interest to this work.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Luz, E. J. D. S., Schwartz, W. R., Cámara-Chávez, G., & Menotti, D. (2016). ECG-based heartbeat classification for arrhythmia detection: A survey. *Computer Methods and Programs in Biomedicine*, 127(C), 144–164.
2. Tang, D. H., Gilligan, A. M., & Romero, K. (2014). Economic burden and disparities in healthcare resource use among adult patients with cardiac arrhythmia. *Applied Health Economics and Health Policy*, 12(1), 59–71.
3. Levin, R. I., et al. (1986). Potential for real-time processing of the continuously monitored electrocardiogram in the detection, quantitation, and intervention of silent myocardial ischemia. *Cardiology Clinics*, 4(4), 735–745.
4. Arora, B. M. N., & Vora, Y. (2019). Wearable ECG for real time complex P-QRS-T detection and classification of various arrhythmias. In *Proceedings of 2019 11th international conference on*

- communication systems & networks (COMSNETS)* (pp. 870–875). Los Alamitos: IEEE.
5. Da Silva, H. P., Carreiras, C., Lourenço, A., Fred, A., Das Neves, R. C., & Ferreira, R. (2015). Off-the-person electrocardiography: Performance assessment and clinical correlation. *Health and Technology*, 4(4), 309–318.
  6. Holter, N. J. (1961). New method for heart studies: Continuous electrocardiography of active subjects over long periods is now practical. *Science*, 134(3486), 1214–1220.
  7. Survivors of Out-Of-Hospital Cardiac Arrest With Apparently Normal Heart. (1997). Need for definition and standardized clinical evaluation. Consensus Statement of the Joint Steering Committees of the Unexplained Cardiac Arrest Registry of Europe and of the Idiopathic Ventricular Fibrillation Registry of the United States. *Circulation*, 95(1), 265–272.
  8. Subramanian, V. B. (1986). Clinical and research applications of ambulatory Holter ST-segment and heart rate monitoring. *The American Journal of Cardiology*, 58(4), B11–B20.
  9. Deanfield, J. E. (1987). Holter monitoring in assessment of angina pectoris. *American Journal of Cardiology*, 59(7), 18–22.
  10. Jovanov, T. M. E., & Raskovic, D. (2000). Issues in wearable computing for medical monitoring applications: A case study of a wearable ECG monitoring device. In *ISWC* (pp. 43–49). Los Alamitos: IEEE Computer Society.
  11. Krahm, A. D., Klein, G. J., Yee, R., & Skanes, A. C. (1999). Recording that elusive rhythm. *CMAJ: Canadian Medical Association Journal*, 161(11), 1424–1425.
  12. Steinberg, C., et al. (2019). A novel wearable device for continuous ambulatory ECG recording: Proof of concept and assessment of signal quality. *Biosensors*, 9(1), 17. <https://doi.org/10.3390/bios9010017>
  13. de Chazal, P. (2013). Detection of supraventricular and ventricular ectopic beats using a single lead ECG. In *Conference proceedings: Annual international conference of the IEEE engineering in medicine and biology society* (Vol. 2013, p. 45).
  14. De ChazalM, P., O'Dwyer, & Reilly, R. B. (2004). Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering*, 51(7), 1196–1206.
  15. Elgendi, M., Eskofier, B., Dokos, S., & Abbott, D. (2014). Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems. *PLoS ONE*, 9(1), e84011-e. <https://doi.org/10.1371/journal.pone.0084018>
  16. Liu, C., et al. (2019). Signal quality assessment and lightweight QRS detection for wearable ECG SmartVest system. *IEEE Internet of Things Journal*, 6(2), 1363–1374. <https://doi.org/10.1109/JIOT.2018.2844090>
  17. Clifford, G. D., Behar, J., Li, Q., & Rezek, I. (2012). "Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. *Physiological Measurement*, 33(9), 1419–1433.
  18. Majumder, S., Chen, L., Marinov, O., Chen, C., Mondal, T., & Deen, M. J. (2018). Noncontact wearable wireless ECG systems for long-term monitoring. *IEEE Reviews in Biomedical Engineering*, 11(9), 306–321.
  19. Flores, N., Avitia, R. L., Reyna, M. A., & García, C. (2018). Readily available ECG databases. *Journal of Electrocardiology*, 51(6), 1095–1097. doi:<https://doi.org/10.1016/j.jelectrocard.2018.09.012>.
  20. Goldberger, A. L., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23), 215–220.
  21. Martinez, J. P., Almeida, R., Olmos, S., Rocha, A. P., & Laguna, P. (2004). A wavelet-based ECG delineator: Evaluation on standard databases. *IEEE Transactions on Biomedical Engineering*, 51(4), 570–581.
  22. Liu, J. Z. X., & Dong, J. (2012). CCDD: An enhanced standard ECG database with its management and annotation tools. *International Journal of Artificial Intelligence*, 21(05), 1240020–1240021.
  23. Liu, F., et al. (2018). Performance analysis of ten common QRS detectors on different ECG application cases. *Journal of Healthcare Engineering*, 2018(2), 1–8.
  24. Clifford, G. D. (2006). ECG statistics, noise, artifacts, and missing data. In G. D. Clifford, F. Azuaje, & P. E. McSharry (Eds.), *Advanced methods and tools for ECG data analysis* (pp. 55–93). Boston: Artech House Publishing.
  25. Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering*, 32(3), 230–236.
  26. Hamilton, P. S., & Tompkins, W. J. (1986). Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. *IEEE Transactions on Biomedical Engineering*, 33(12), 1157–1165.
  27. Yildirim, Ö. (2018). A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. *Computers in Biology and Medicine*, 96, 189–202. <https://doi.org/10.1016/j.combiomed.2018.03.016>
  28. Oh, S. L., Ng, E. Y. K., Tan, R. S., & Acharya, U. R. (2018). Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats. *Computers in Biology and Medicine*, 102, 278–287. <https://doi.org/10.1016/j.combiomed.2018.06.002>
  29. Mark, Q. LiR. G., & Clifford, G. D. (2007). Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. *Physiological Measurement*, 29(1), 15–32.
  30. Behar, J., Oster, J., Li, Q., & Clifford, G. D. (2013). ECG signal quality during arrhythmia and its application to false alarm reduction. *IEEE Transactions on Biomedical Engineering*, 60(6), 1660–1666. <https://doi.org/10.1109/TBME.2013.2240452>
  31. Redmond, S. J., Lovell, N. H., Basilakis, J., & Celler, B. G. (2008). ECG quality measures in telecare monitoring. In *IEEE engineering in medicine and biology society. Annual conference* (Vol. 2008, pp. 2869–2872). <https://doi.org/10.1109/IEMBS.2008.4649801>.
  32. Moody, I. Silva, G. B., & Celi, L. (2011). Improving the quality of ECGs collected using mobile phones: The Physionet/computing in cardiology challenge 2011. In *Proceedings of 2011 computing in cardiology* (pp. 273–276). Los Alamitos: IEEE.
  33. Searle, A., & Kirkup, L. (2000). A direct comparison of wet, dry and insulating bioelectric recording electrodes. *Physiological Measurement*, 21(2), 271–283.
  34. Baek, J., An, J., Choi, J., Park, K., & Lee, S. (2008). Flexible polymeric dry electrodes for the long-term monitoring of ECG. *Sensors and Actuators A: Physical*, 143(2), 423–429.
  35. ANSI/AAMI EC57:1998/(R)2008. (2008). *Series (R) 2008-Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms*. Arlington: American National Standards Institute.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.