scientific reports

Check for updates

OPEN

Detection and categorization of severe cardiac disorders based solely on heart period measurements

Shigeru Shinomoto[1,2,3,6✉], Yasuhiro Tsubo[3,4,6] & Yoshinori Marunaka[3,5]

Cardiac disorders are common conditions associated with a high mortality rate. Due to their potential for causing serious symptoms, it is desirable to constantly monitor cardiac status using an accessible device such as a smartwatch. While electrocardiograms (ECGs) can make the detailed diagnosis of cardiac disorders, the examination is typically performed only once a year for each individual during health checkups, and it requires expert medical practitioners to make comprehensive judgments. Here we describe a newly developed automated system for alerting individuals about cardiac disorders solely by measuring a series of heart periods. For this purpose, we examined two metrics of heart rate variability (HRV) and analyzed 1-day ECG recordings of more than 1,000 subjects in total. We found that a metric of local variation was more efficient than conventional HRV metrics for alerting cardiac disorders, and furthermore, that a newly introduced metric of local-global variation resulted in superior capacity for discriminating between premature contraction and atrial fibrillation. Even with a 1-minute recording of heart periods, our new detection system had a diagnostic performance even better than that of the conventional analysis method applied to a 1-day recording.

Heart disease has one of the highest mortality rates of any condition, which is similar to that of cancer[1–5]. Early detection of cardiac disorders might help prevent fatalities. One of the most basic means of detecting irregularities is pulse assessment, which was first used in ancient China and Egypt in centuries B.C.[6–8]. In the modern era, the electrocardiogram (ECG) provides a detailed assessment of cardiac electrical activity, reflecting depolarization and repolarization of the atria and ventricles[9]. Based on this information, a variety of diagnoses can be made, such as premature ventricular contraction (PVC), premature atrial contraction (PAC), and atrial fibrillation (AF).

Though efficient, ECG examinations are costly and inconvenient. While electrical signals are efficiently analyzed by automated algorithms, comprehensive diagnoses must be made by medical practitioners in the end[10,11]. Accordingly, they are typically carried out only once a year for each individual during physical checkups, and fewer people are recommended to perform Holter ECG examinations continuously for 1 day, only if some cardiac anomalies are found in the physical checkups. More serious is that those who were negative in those ECG examinations still have a chance of getting cardiac disorders.

Thus it is desirable to make it possible to alert all persons to cardiac dysfunction using a handy device such as a smartwatch. While full ECG information is not available, a handy device may make it possible to constantly monitor a series of pulse intervals, which may provide sufficient information for alerting cardiac disorders[12]. This has led us to an idea of identifying cardiac disorders by analyzing heart rate variability (HRV) through heart periods in as much detail, just as expert medical practitioners have been doing for centuries.

HRV is quantified by measuring fluctuations in heartbeat intervals, using metrics such as the standard deviation of intervals between successive cardiac R waves or RR intervals (SDRR) and the coefficient of variation (Cv), which is defined as the ratio of the standard deviation to the mean[13,14]. Nevertheless, the variation in heartbeat intervals is caused by not only cardiac disorders but also slow fluctuations in the heart rate that occur even in healthy people[15,16]. There have been many attempts for isolating cardiac disorders from such miscellaneous

[1]Brain Information Communication Research Laboratory Group, ATR Institute International, Kyoto 619-0288, Japan. [2]Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan. [3]Medical Research Institute, Kyoto Industrial Health Association, Kyoto 604-8472, Japan. [4]College of Information Science and Engineering, Ritsumeikan University, Shiga 525-8577, Japan. [5]Department of Molecular Cell Physiology, Kyoto Prefectural University of Medicine Graduate School of Medical Science, Kyoto 602-8566, Japan. [6]These authors contributed equally: Shigeru Shinomoto and Yasuhiro Tsubo. ✉email: shigerushinomoto@gmail.com

factors, using a variety of analysis methods such as linear, frequency domain, wavelet domain, and nonlinear methods[17–21]. Nevertheless, diagnosing cardiac symptoms still requires medical experts to consider various possible conditions when making comprehensive judgments.

Recently, artificial intelligence-aided methods based on deep learning algorithms have been adopted for detecting cardiac disorders from heartbeat signals[5,22–25]. While these methods achieve high performance, individual decisions depend on a huge number of parameters that were empirically determined with given training datasets. A problem of such artificial intelligence-aided methods is that we cannot explain the reason for individual diagnoses made for each case. Another problem is that such machine-learning diagnoses may depend sensitively on a huge number of model parameters, and accordingly, identical results may not be reproduced even from the same set of training data.

To improve these points, here we design an automatic diagnosing system by simply combining two metrics that measure specific aspects of heartbeat variability. On a plane of the two metrics, we specify decision boundaries for different types of cardiac disorders. From this, we can characterize the difference in various cardiac diseases explicitly in terms of two metrics.

As the first metric, we have adopted the metric of local variation (Lv), which was previously used to analyze neuronal firing in the brain[26–28]. This is because Lv captures the difference in adjacent pulse intervals by removing the influence of slow fluctuations in heart rate. We have confirmed that Lv was superior to the conventional Cv in detecting premature contraction. The working of Lv is similar to that of an HRV index $SD1^2$, in that they measure the variation of consecutive RR intervals. The superiority of Lv to Cv is similar to the superiority of $SD1^2$ to other linear HRV indices such as SDRR or SDSD[29,30].

Nevertheless, we also realized that Cv may still have captured some aspect independent of Lv particularly for atrial fibrillation. By identifying the differences between Lv and Cv, we sought to combine them in a manner that would maximize the ability to characterize cardiac status. Through this process, we invented a new metric, the local-global variation ratio (Lg), that effectively discriminates between premature contractions and fibrillation. Using these two metrics Lv and Lg, we constructed an algorithm to automatically diagnose cardiac symptoms. The new method exhibits outstanding performance in alerting AF, which is associated with serious diseases[31–34].

For this analysis, we used Holter ECG recordings obtained from more than 1,000 outpatients in total who had cardiological examinations at Clinic, Kyoto Industrial Health Association. Recordings contain a series of RR intervals monitored for 1 day, and are accompanied by diagnoses made by medical doctors and technicians, including not only AF but also PVC and PAC, the latter of which may take place even in healthy people[35,36]. By assuming that heart periods that can be measured by a handy device are noisy observation of RR intervals[37,38], we attempted to infer diagnoses of PVC, PAC, and AF from a set of RR intervals and examined how the inference can withstand the possible fluctuations added to the RR intervals, mimicking the noisy observations. To examine the generality of the analysis, we also applied the current method to the publicly available MIT-BIH databases, which have been conventionally used as standard datasets.

Results

For Holter ECG data recorded from 1,017 subjects in total (863 independent persons), we computed the variability metrics from each series of RR intervals and compared their ability to alert cardiac conditions, such as PVC, PAC, and AF.

Comparison of variability metrics.   Figure 1 depicts the distributions of PVC, PAC, and AF plotted against the RR interval metrics: the heart rate Hr, the logarithm of coefficient of variation $\log_{10} Cv$, and the logarithm of the local variation $\log_{10} Lv$ (METHODS), each of which is calculated every 10 minutes and averaged over 18 hours. Because Cv and Lv computed for heartbeats are much smaller than unity, we take their logarithm to focus on their difference.

Many subjects have nonzero PVC and PAC, in which premature heartbeat is initiated in ventricles and atria, respectively. These events may occur even in healthy people. In our dataset, individual PVC and PAC events were identified by an automatic detection algorithm of ECG, and those with the occurrence probability higher than 10^{-3} comprised 49% and 33% of the entire subjects, respectively. In contrast with these, AF accompanied by the irregular beating of the atrial chambers is considered a serious symptom leading to cardiac dysfunction. In our dataset, the symptom was determined manually by expert medical doctors and technicians, and 72 of 1,017 subjects (7%) exhibited nonzero AF.

For PVC, PAC, and AF, we categorized the diagnostic status of each subject as 0 or 1 according to whether the ratio of each condition was lower or higher than a threshold of 0.1, and predicted the dichotomic status based on Cv or Lv as obtained from pulsation signals. The numbers of the 1,017 subjects at high risk of PVC, PAC, and AF were 128 (12%), 34 (3%), and 53 (5%), respectively. Only one subject (0.1%) exhibited multiple symptoms (PAC and AF).
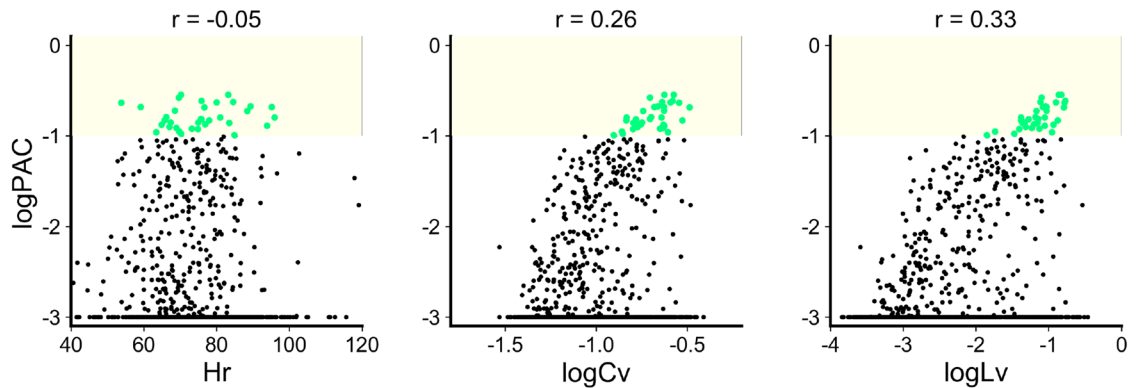
A conventional HRV metric Cv and our newly introduced metric Lv were strongly correlated with cardiac symptoms; subjects who exhibited the higher Cv or Lv were more likely to exhibit cardiac disorders. The average heart rate Hr was relatively weakly correlated with cardiac symptoms, and accordingly, we shall concentrate on Cv and Lv in the following analysis.

We first examined whether the high value of Cv or Lv could signal each cardiac disorder. We categorized data as "positive" or "negative" according to whether or not $\log_{10} Cv$ or $\log_{10} Lv$ was greater than a given threshold. Positive data were categorized as "true positive (TP)" or "false positive (FP)" according to whether or not subjects exhibited a given cardiac disorder (PVC, PAC, or AF). In contrast, negative data were categorized as "false negative (FN)" or "true negative (TN)" according to whether or not subjects exhibited a given cardiac disorder.
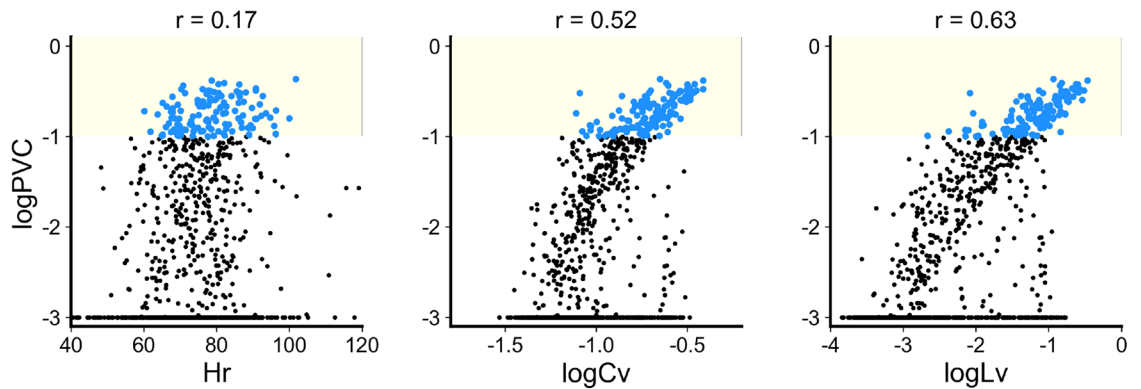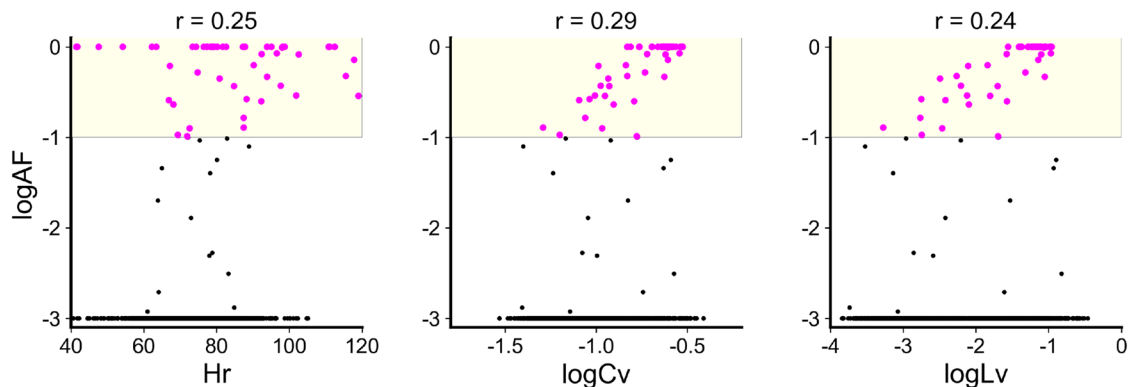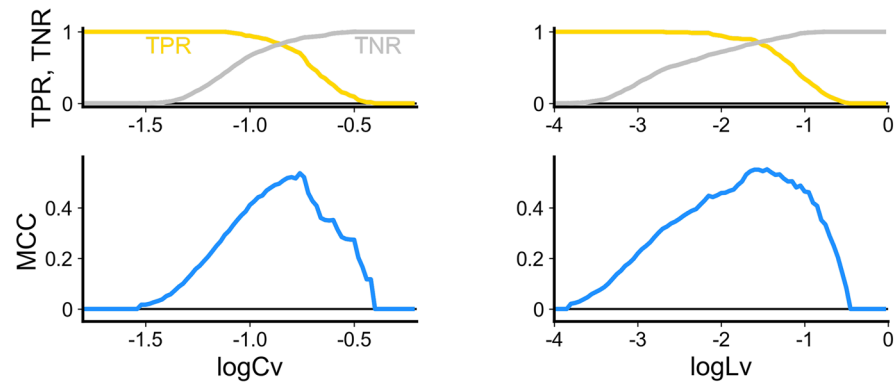
Figure 1. Cardiac disorders plotted against Hr, $\log_{10} Cv$, and $\log_{10} Lv$. Each dot represents the average of 10-minute statistics for a single subject over 18 hours. Vertical axes are the logarithm of (**a**) premature ventricular contraction (PVC); (**b**) premature atrial contraction (PAC); and (**c**) atrial fibrillation (AF). Values of the Pearson correlation r are indicated above. Yellow zones represent regions of > 0.1.

The upper rows in Fig. 2 depicts the true positive ratio $\text{TPR} = N_{\text{TP}}/(N_{\text{TP}} + N_{\text{FN}})$, and the true negative ratio $\text{TNR} = N_{\text{TN}}/(N_{\text{TN}} + N_{\text{FP}})$, where N_{TP}, N_{FN}, N_{TN}, and N_{FP} are the numbers of true-positive, false-negative, true-negative, and false-positive cases, respectively. Higher Cv or Lv results in higher TNR, implying that subjects in whom the metric is lower than the threshold are unlikely to exhibit cardiac disorders. However, a high threshold lowers TPR, implying that many cardiac disorders are missed.
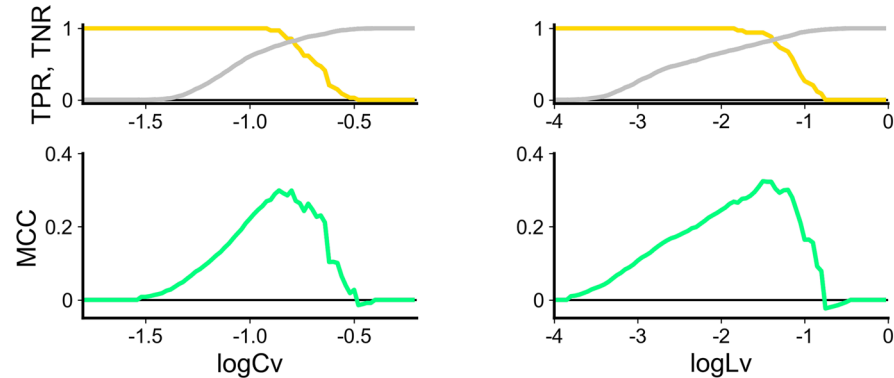
One means of achieving a reasonable balance between true and false positives and negatives is to maximize the Matthews correlation coefficient (MCC)[39] defined as

$$\text{MCC} = \frac{N_{\text{TP}}N_{\text{TN}} - N_{\text{FP}}N_{\text{FN}}}{\sqrt{(N_{\text{TP}} + N_{\text{FP}})(N_{\text{TP}} + N_{\text{FN}})(N_{\text{TN}} + N_{\text{FP}})(N_{\text{TN}} + N_{\text{FN}})}}.$$
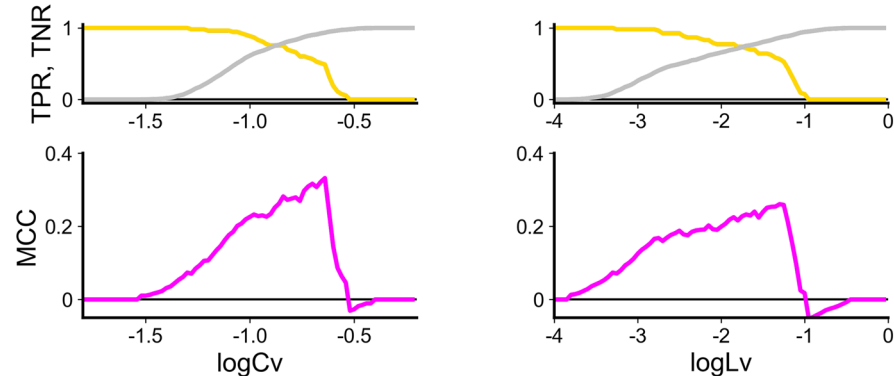
Figure 2. The true positive rate (TPR), true negative rate (TNR), and the Matthews correlation coefficient (MCC) measuring the ability of two individual metrics, thresholds of $\log_{10} Cv$ and $\log_{10} Lv$, to detect cardiac disorders: (**a**) PVC, (**b**) PAC, and (**c**) AF. For each cardiac disorder, each subject was categorized into positive or negative, according to whether the ratio of anomalous states identified by an automatic detection algorithm of ECG was higher or lower than 10%. TPR and TNR indicate the probability of each person being positive and negative according to whether the metric Cv or Lv is higher and lower than a given value, respectively. MCC measures a balance between true and false positives and negatives.

The lower rows in Fig. 2 depict the MCC values for PVC, PAC, and AF; peaks occur at an intermediate value for a threshold of $\log_{10} Cv$ or $\log_{10} Lv$. For PVC and PAC, the new metric Lv is more efficient than Cv in achieving higher MCC values. By contrast, Cv results in a higher MCC than Lv for AF.

Improving the detection of cardiac disorders.   In the aforementioned analysis, we have seen that Cv and Lv had different strengths for different disorders. Because these metrics capture different aspects of the RR-variability, their performance in alerting cardiac disorders might be increased if these metrics were suitably combined.
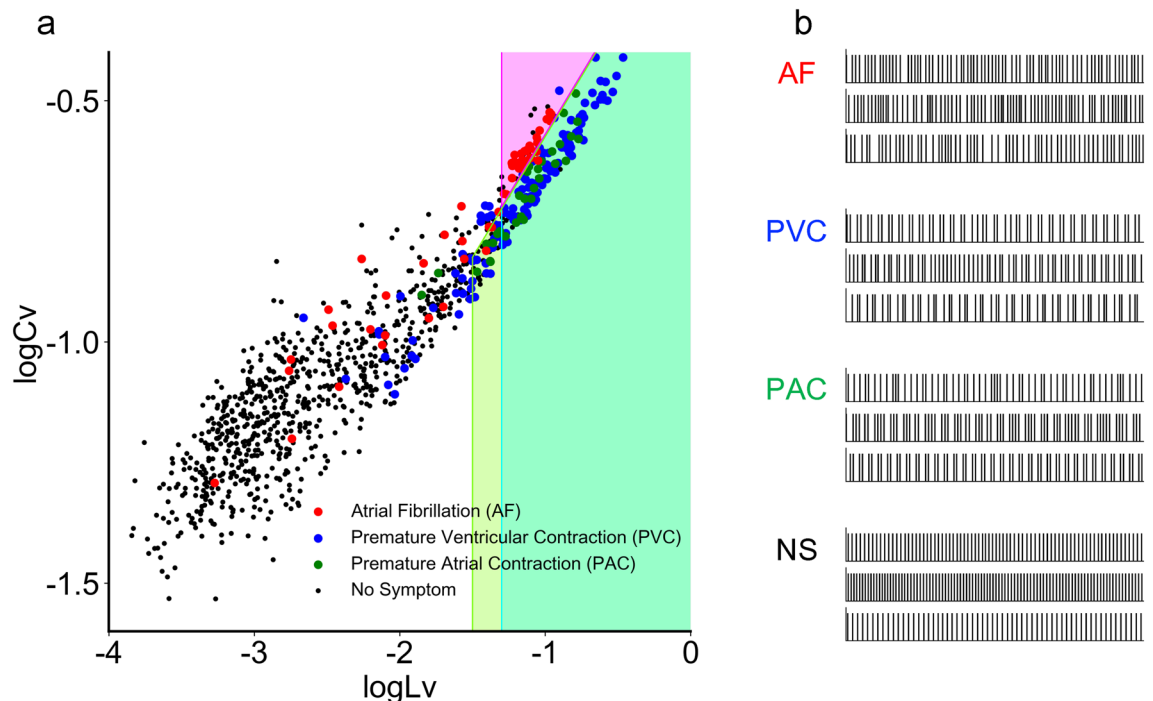
Figure 3. Different cardiac disorders plotted on a plane spanned by the new and old metrics $\log_{10} Lv$ and $\log_{10} Cv$. (**a**) Dots represent 1,017 subjects, who were diagnosed as atrial fibrillation (AF), premature ventricular contraction (PVC), premature atrial contraction (PAC), and negative, respectively colored in magenta, blue, and green, and black. The prediction zone for each cardiac disorders, given by the set of constraints $\log_{10} Lv > a$ and $\log_{10} Lg = \log_{10} Lv - 2\log_{10} Cv > b$ (or $< b$) is depicted by the appropriate colors. (**b**) Representative pulse sequences of 1 minute, respectively taken from four different cases. NS stands for no symptom.

Figure 3a displays the distribution of datapoints representing subjects with different cardiac diseases plotted on a plane spanned by $\log_{10} Lv$ and $\log_{10} Cv$. The distribution seemingly comprises different groups; there is one big cluster at the lower left, centered at low $\log_{10} Lv \approx -3$ and low $\log_{10} Cv \approx -1.2$, while the datasets with higher $\log_{10} Lv > -2$ tended to align linearly with the slope in parallel to $\log_{10} Lv - 2\log_{10} Cv$.

Measurements associated with cardiac disorders are located in the upper right region defined by high values of $\log_{10} Lv$ and $\log_{10} Cv$. On further analysis, the data points representing different cardiac disorders are distributed differently in this region; fibrillation (AF) cases are located on the upper side, whereas premature contraction (PVC and PAC) cases are located on the lower side. Representative pulse sequences of 1 minute taken from these diagnostic cases are depicted in Fig. 3b.

To determine how the local variation Lv and the coefficient of (global) variation Cv contribute to the separation of premature contractions and fibrillation, we newly introduce a metric of the local-global variation ratio (Lg), defined as

$$Lg = \frac{Lv}{Cv^2}.$$

In the denominator, Cv, which measures the standard deviation of intervals, is squared to conform to the power of deviation in the numerator Lv measures the squared deviation (METHODS). For a Poisson random pulse train, Lg exhibits the value unity because both Lv and Cv take the value unity. For a locally regular pulsation whose rate is slowly modulated, $Lg \ll 1$, because Lv is much smaller than Cv^2. For a perfectly regular pulsation in which RR intervals are identical, Lg is undefined because both the numerator Lv and the denominator Cv^2 are zero. We shall prove in the METHODS that Lg takes the value 3/2 for a sequence in which RR intervals of similar durations are arranged randomly. For a pulsation in which long and short intervals alternate, however, Lg takes the even higher value 3. In this way, the local-global variation ratio Lg can discriminate how intervals are arranged even if the intervals are close to each other (as is the case with heartbeats).

It is noteworthy that in Fig. 3, premature contractions (PVC and PAC) and fibrillation (AF) are well separated by the line $\log_{10} Lg = \log_{10} Lv - 2\log_{10} Cv \approx 0.15$. It is helpful to recall that $Lg = 3/2$ for randomly arranged RR intervals and that $\log_{10} 3/2 \approx 0.18$. As we have seen, Lg can be as large as 3 if long and short intervals alternate. This means that long and short pulse intervals tend to alternate in premature contractions, as they are distributed on the lower side of the line $\log_{10} Lg = \log_{10} Lv - 2\log_{10} Cv \approx 0.18$.

Considering these features, it is advantageous to infer the presence of each cardiac disorder based on a combination of variability metrics given by the following set of constraints,

$$\log_{10} Lv > a$$

Alerting method	PVC	PAC	AF
Cv	0.54	0.30	0.33
Lv	0.55	0.32	0.26
Combination	0.63	0.38	0.59

Table 1. MCC values measuring the performances of each alerting method.
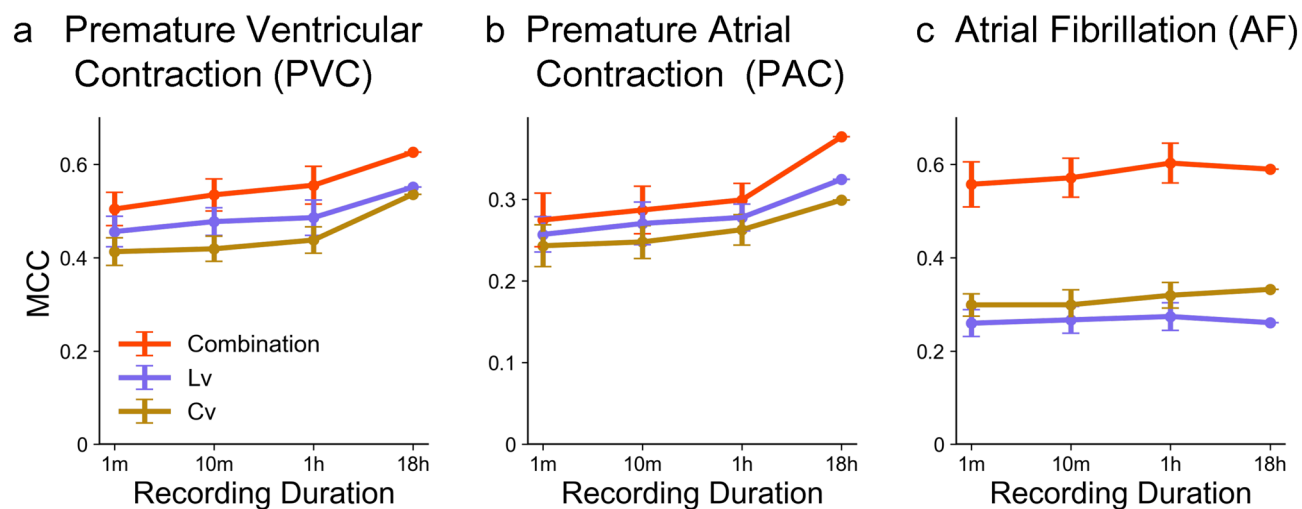

Figure 4. The MCCs for *Cv* and *Lv*, and the combination method based on datasets with various recording durations: 1 minute, 10 minutes, 1 hour, and 18 hours. (**a**) PVC, (**b**) PAC, and (**c**) AF. The MCCs for the shorter recording durations are represented with the mean and standard deviation.

and

$$\log_{10} Lg > b \ (\text{or} < b).$$

We selected the parameters *a* and *b* and the inequality direction so that the goodness index MCC is maximized for each cardiac disorder. The selected parameters are as follows: (PVC): $\log_{10} Lv > -1.3, \log_{10} Lg > 0.14$; (PAC): $\log_{10} Lv > -1.5, \log_{10} Lg > 0.15$; (AF): $\log_{10} Lv > -1.3, \log_{10} Lg < 0.15$. The selected zones are depicted in different colors in Fig. 3.

The method of combining *Lv* and *Lg* achieves MCC performances much higher than those obtained from simple thresholding of either *Cv* or *Lv* alone (Table 1). In particular, the ability to alert fibrillation (AF) was drastically improved by the combination method. Note that the obtained MCC performances would practically be unchanged even if we make leave-one-out cross-validation because the number of parameters (one or two) is much lower than the number of data points (higher than 1,000).

Alerting cardiac disorders based on a shorter recording period. So far we have compared alerting methods using 18-hour data of RR intervals extracted from 1-day Holter ECG. As the Holter recording is costly and burdensome, it would be ideal to be able to detect cardiac disorders even with a shorter interval. Here we are interested in how the detection performance degrades if the recording duration is shortened, such as 1 hour, 10 minutes, or even 1 minute. We clipped out such shorter recordings from full-length data and estimated the MCC values for detection performance using the single metrics *Cv* and *Lv* as well as the combination method. Figure 4 shows that our new combination method may provide good performance even based on shorter recordings, with an accuracy comparable to or even higher than that obtained by applying a conventional single metric such as *Cv* to an 18-hour data of RR intervals.

Figure 5 depicts distributions of HRV metrics computed from recording periods of 1 minute, 10 minutes, and 1 hour. While datasets of shorter durations are more scattered on a plane spanned by $\log_{10} Lv$ and $\log_{10} Cv$, the distributions of premature contraction (PVC and PAC), fibrillation (AF), and negative cases remain well separated.

Robustness against erroneous observation. If we monitor pulsations using a handy device such as a smartwatch, the measurement will be inaccurate since the pulse intervals are not identical to RR intervals. Considering possible fluctuations in the measurements, we created noisy data by jittering the original RR intervals with noises of the mean zero and the standard deviations of 10, 50, 100, 200, and 500 ms. Figure 6 demonstrates the manner in which the MCC values degrade with different noise durations. We can see that the combination method withstands the fluctuations if their noise level is less than 100 ms.
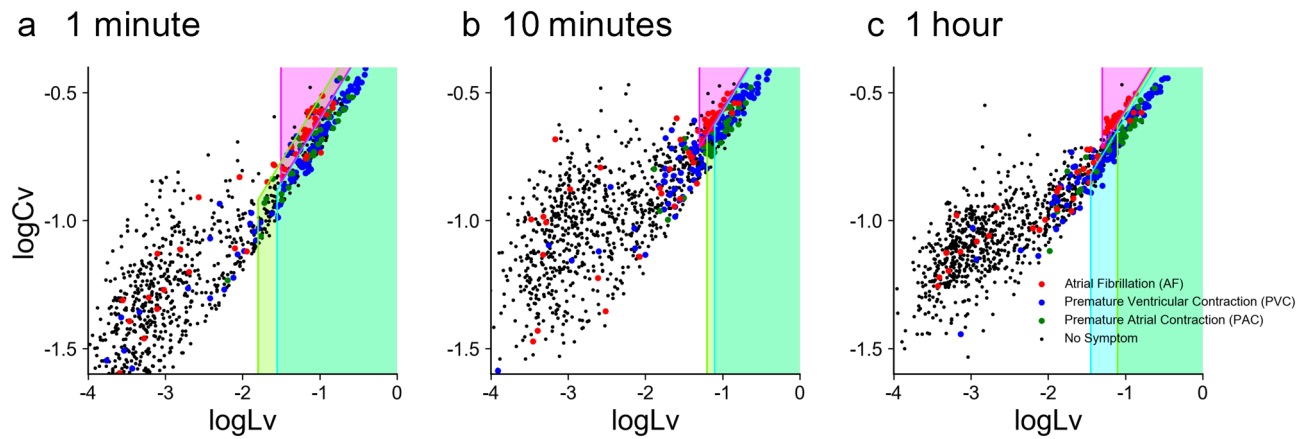
6

Figure 5. Two metrics $\log_{10} Lv$ and $\log_{10} Cv$ computed from shorter recording periods. (**a**), (**b**), and (**c**) Datasets of 1 minute, 10 minutes, and 1 hour, plotted on a plane spanned by $\log_{10} Lv$ and $\log_{10} Cv$. The parameters a and b of each prediction zone were adapted to respective datasets.
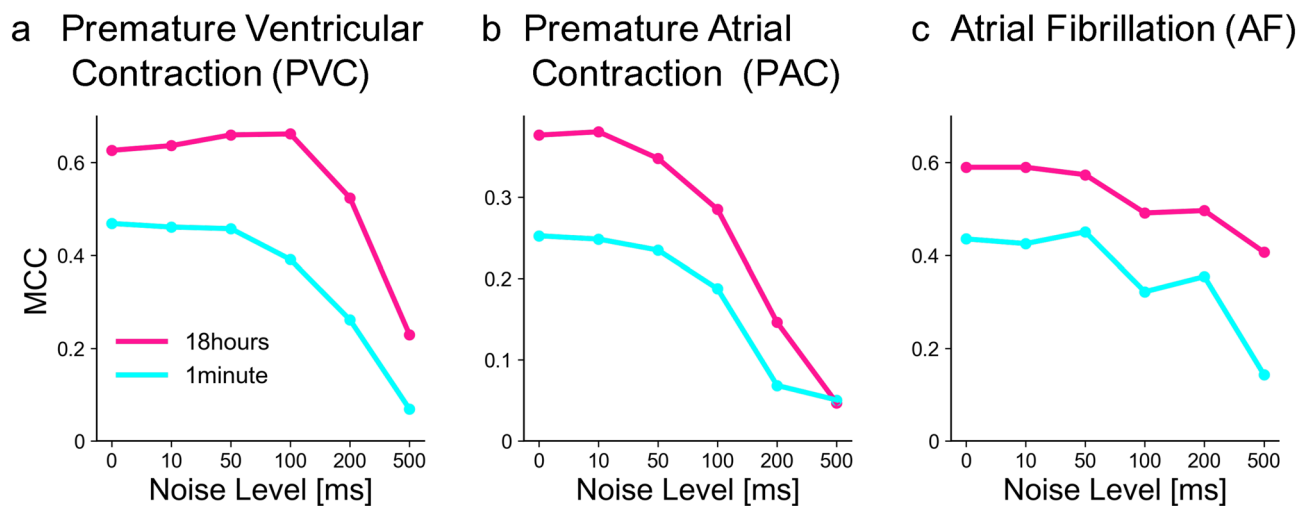

Figure 6. The MCCs for the combination method applied to recording duration of 1 minute and 18 hours. (**a**) PVC, (**b**) PAC, and (**c**) AF. Original pulse times were jittered with Gaussian noises of the standard deviation of 10, 50, 100, 200, and 500 ms.

Figure 7 depicts HRV metrics of noisy datasets, in which the original series of RR intervals of 18 hours were jittered with noises of the standard deviations of 10, 50, and 100 ms. While datasets of noisy RR intervals were shifted toward higher values of $\log_{10} Cv$ and $\log_{10} Lv$, the distributions of premature contraction (PVC and PAC), fibrillation (AF), and negative cases remain well separated for these cases.

Application to public databases. To examine if our method may also be valid for other datasets, we applied it to the data of the MIT-BIH databases, which have been deemed standard. Here we considered the "Atrial Fibrillation Database" as representing the symptoms similar to the AF cases selected by Kyoto Industrial Health Association, while the "Arrhythmia Database" and the "Normal Sinus Rhythm Database" represent those who were not AF. Figure 8 represents the distributions of $\log_{10} Lv$ versus $\log_{10} Cv$ of 1minute, 10minute, 1hour, and 18hour data of the MIT-BIH datasets. For each dataset, two parameters a and b of our prediction model were selected so that MCC is maximized. The MCC values obtained for the AF cases of MIT-BIH datasets were much higher than those obtained for our Kyoto datasets; for instance, the MCC value for 18-hour data was as high as 0.761. The difference may have arisen because diagnostic criteria and the population ratio of cardiac cases were different between these datasets. Nevertheless, it is noteworthy to see the two parameters a and b were similar to those of Kyoto database; for instance, $a = -0.115$ and $b = 0.27$ for 18 hour data. This implies that a similar criterion may be applicable even for different datasets.
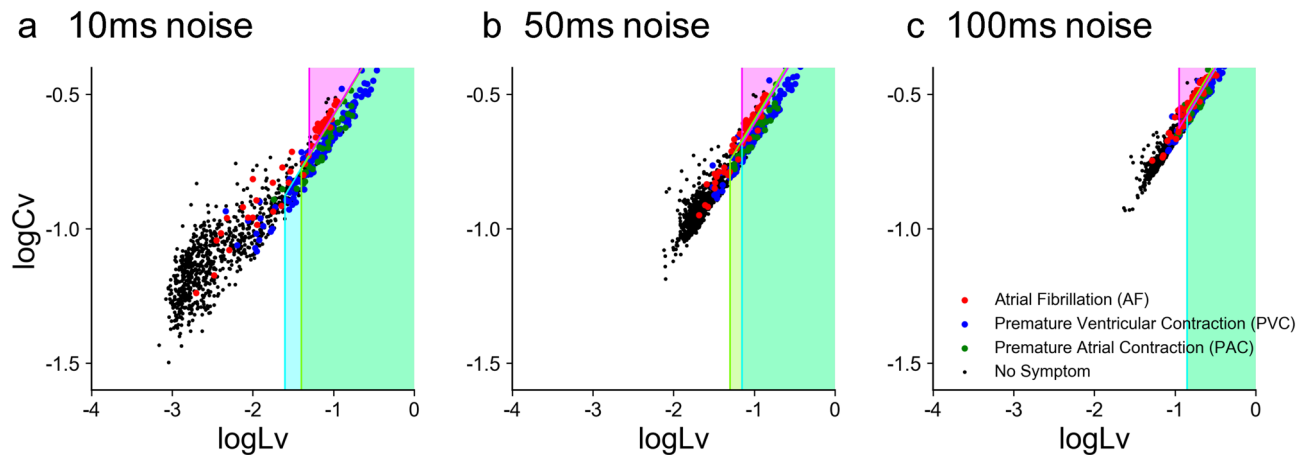
Figure 7. Noisy datasets of different cardiac disorders plotted on a plane spanned by $\log_{10} Lv$ and $\log_{10} Cv$. (**a**), (**b**), and (**c**): Datasets were obtained by jittering original pulse times with Gaussian noises of the standard deviation of 10 ms, 50 ms, and 100 ms. The parameters a and b of each prediction zone were adapted to respective noisy datasets.
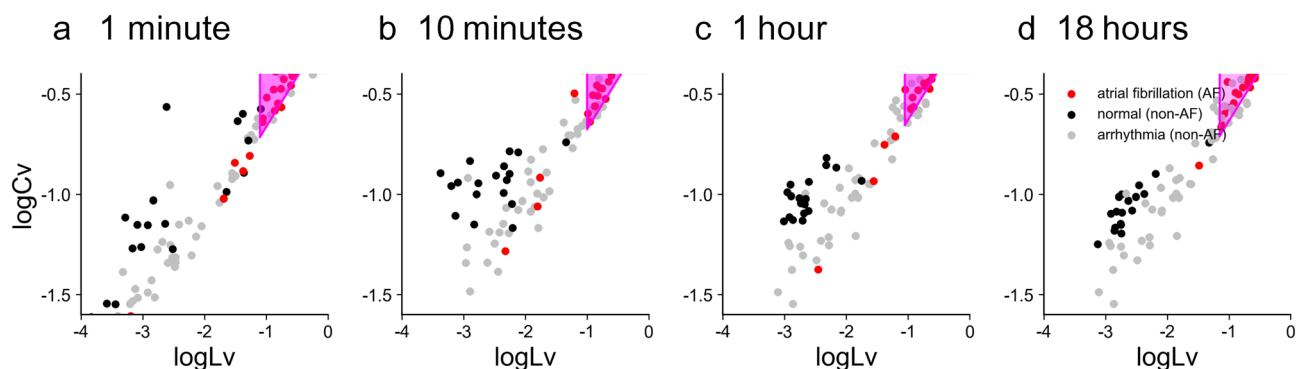

Figure 8. Two metrics $\log_{10} Lv$ and $\log_{10} Cv$ computed for three kinds of MIT-BIH databases. (**a**), (**b**), (**c**), and (**d**): Datasets of 1 minute, 10 minutes, 1 hour, and 18 hours plotted on a plane spanned by the two metrics. The parameters a and b of the prediction zone (magenta) for the AF cases were adapted to respective datasets.

Discussion

In this study, we constructed an automated algorithm for alerting cardiac symptoms from pulsation signals. We first compared Lv with Cv and found that Lv was superior for detecting premature contractions, while Cv was superior for identifying fibrillation. Considering the specificity of each metric, we introduced the local-global variation ratio $Lg = Lv/Cv^2$ and found that this metric effectively discriminates between premature contractions and fibrillation. With a method for combining Lv and Lg, we have obtained a performance far superior to that obtained by simple thresholding with single metrics.

Considering the cost and time needed to perform 1-day Holter ECG, it is desirable to be able to make a reasonable inference from a shorter recording, such as the few-minute ECG carried out during a regular medical checkup. We confirmed that applying our new combination method to 1-minute recordings provided detection performance that was comparable to or even better than that of the conventional HRV metric Cv applied to a 1-day recording.

Nevertheless, this favorable performance might not be exclusive to our method. Experienced medical practitioners should be able to achieve similar performance through pulse diagnosis, or by making full use of the many existing HRV metrics. Some of these metrics can detect features similar to Cv and Lv. For instance, if we map consecutive RR intervals, as was done in[40–42], the variability along a diagonal line in the map corresponds to the global variability as represented by Cv, while the variability along the orthogonal axis is similar to the local variability of consecutive intervals as represented by Lv, in which slow rate fluctuation is effectively mitigated.

We have seen that longer recording of pulses provides more reliable inference. This is partly because the estimation of variability metrics becomes more accurate with time. But the main reason may be that cardiac symptoms occur intermittently during 1-day recording and it is difficult to identify disorders if the analysis is performed while symptoms are absent. But if we can use a more convenient device such as a smartwatch that can constantly monitor heartbeats, cardiac disorders can be easily detected using our analysis algorithm.

By assuming that pulses reflect the RR intervals of cardiac beats, being accompanied by noises occurring downstream, we tested whether our analysis method was robust against noisy data. By jittering the original RR

intervals with fluctuating noises, we have confirmed that the detection of cardiac symptoms did not deteriorate if the fluctuation in the measurements was shorter than 100 ms. Thus our method might work for alerting to cardiac disorders from pulse signals obtained with a smartwatch. However, it should be noted that HRV analysis is limited in that it cannot classify arrhythmic events such as isolated premature ventricular contraction[43,44]. Once any malfunction is detected by HRV analysis, one should move on to ECG analysis for diagnosing the more detailed morphological features.

Here we selected the parameters for determining cardiac symptoms so that the performance was maximized in terms of the MCC. However, the MCC was based on a balance between true and false positives and negatives, in the sense that positives and negatives are considered to have equal weights. We may lower the threshold of Lv if we wish to be more cautious in predicting cardiac disorders, even though this results in an increase of false positives.

To test the generality of the analysis method, we applied it to the data of the MIT-BIH databases. We have confirmed that our method may detect AF cases with much higher MCC values, indicating the higher classification performance. It is noteworthy that we obtained a similar criterion for this dataset, although the diagnostic criteria may generally be different between different medical institutions. The advantage of our method may be its simplicity in representing and diagnosing cardiac status; our method has only two parameters a and b, and we may check whether they have been dependent on training datasets, which may be largely dependent on the diagnostic criterion of different institutions.

In this study, we achieved high detection performance by combining Cv and Lv, particularly by paying attention to the local-global variation ratio Lg. The essential point of improvement was to combine multiple metrics that detect different features of heartbeats. If we can obtain more sample data, it might be worthwhile to look for more detailed combinations of additional variability metrics.

Methods

ECG data.
The data of Holter ECG were obtained from outpatients who had cardiological examinations at Clinic, Kyoto Industrial Health Association, Kyoto 604-8472, Japan. Japanese institutions or corporations are requested to let workers undergo a medical checkup once a year. If potentially concerning results are observed during the ECG recording performed during the physical checkup, individuals are recommended to undergo 1-day Holter ECG monitoring. Accordingly, subjects who received Holter ECG may have been more likely to have cardiac disorders than the overall population. Nevertheless, the majority of these individuals were still "healthy" in that they were not diagnosed with heart failure, and some were asymptomatic.

In the analysis of 1-day Holter recording, each pulse was automatically diagnosed as PVC and/or PAC by software provided by Fukuda Denshi Co. Ltd. Based on the summary data, we calculated the fraction of pulses exhibiting PVC and/or PAC contained in an entire set of pulses. Medical doctors analyzed the full records of individual subjects, and if abnormalities were detected, they manually determined the periods of AF. We calculated the ratio of the AF period to the total measurement time, 18 hours. The ECG data obtained from a CM5 lead were analyzed. From the recorded RR intervals, we computed Cv and Lv. We discarded 14 datasets whose recording period was shorter than 18 hours. Accordingly, we had datasets of 1,017 subjects in total (863 independent persons, men: 625, women: 238), and analyzed the initial 18 hours. The ages of the subjects ranged from 20 to 90.

All methods were carried out in accordance with relevant guidelines and regulations. Informed consent was obtained from all subjects or their legal guardians. Categorization of cardiac symptoms of outpatients was performed by medical doctors who acquired the license as Physician and Surgeon in Japan. The present study was approved by Institutional Research Board, Kyoto Industrial Health Association (Permission No. S18-0006), and the Ethics Review Committee for Medical and Health Research involving Human Subjects, Ritsumeikan University (Permission No. BKC-LSMH-2021-039).

MIT-BIH ECG databases.
In addition to the Kyoto databases, we examined public databases MIT-BIH provided by the Harvard-MIT Division of Health Sciences and Technology[45–47]. Here we adopted three kinds of databases: MIT-BIH Atrial Fibrillation Database (2-channel Holter ECG recorded at a sampling frequency of 250 Hz for 10 hours); MIT-BIH Arrhythmia Database (2-channel Holter ECG recorded at a sampling frequency of 360 Hz for 24 hours. We adopted the 30-minute waveform data); and MIT-BIH Normal Sinus Rhythm Database (2-channel ambulatory ECG recorded at a sampling frequency of 128 Hz for 24 hours). We obtained RR intervals from the waveform data of the first channel. R-peaks were detected using the ecg_peaks function of neurokit2 software[48]. For datasets whose recording period was less than the specified period (1minute, 10minutes, 1hour, 18hours), we analyzed the entire recording period.

Metrics for measuring heartbeats.
Given a sequence of RR intervals, we computed variability metrics defined as follows.

- Average heart rate (Hr) The most basic metric for measuring heartbeats is the average heart rate Hr, defined as,

$$Hr = \text{number of heartbeats per minute.}$$

  The 18-hour average value was adopted as an HR statistic for each subject.
- Coefficient of variation (Cv)
  There are many conventional approaches for detecting HRV, such as linear, frequency domain, wavelet domain, and nonlinear methods. As a representative HRV metric, we used the coefficient of variation of RR intervals, defined by
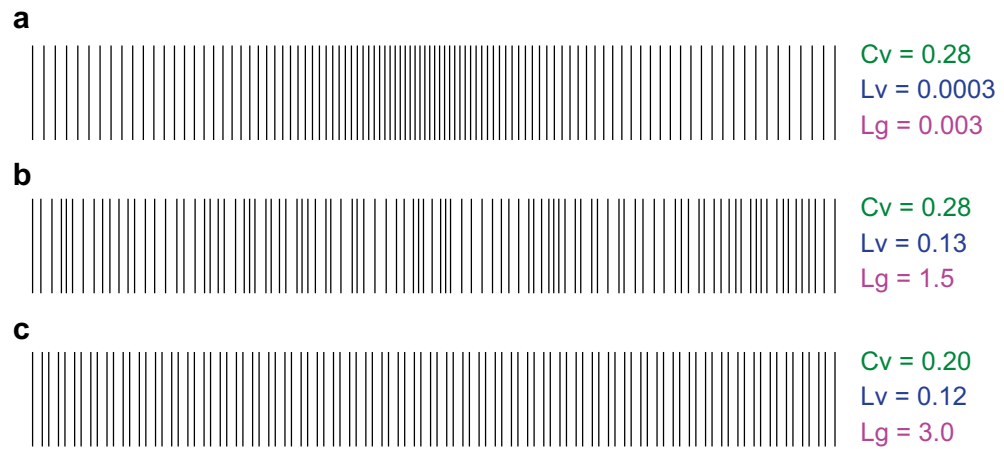
Figure 9. Synthetic pulse sequences and the computed values of the coefficient of variation Cv, the local variation Lv, and the local-global variation ratio Lg. (**a**) Pulsation is locally regular while the rate is slowly modulated from low to high and then back to low. (**b**) Pulsation is irregular, although the average rate is nearly constant. (**c**) Long and short RR intervals alternate with each other. Cv has the same value for (**a**) and (**b**), which are composed of identical sets of RR intervals, whereas Lv is small for (**a**), in which the RR intervals are well organized, and is large for (**b**), in which RR intervals are arranged randomly. Lg is very small for (**a**), in which the rate is slowly modulated. $Lg \approx 3/2$ for (**b**), in which RR intervals are randomly arranged. $Lg \approx 3$ for (**c**), in which long and short RR intervals alternate.

$$Cv = \frac{\Delta I}{\bar{I}},$$

where ΔI and $\bar{I}$ represent the standard deviation and the mean of RR intervals. These statistics are typically measured every 10 minutes (references). Unless specified in the main text, the 18-hour mean of the logarithm of the 10-minute mean Cv was adopted. In Fig. 4, the 1-minute Cv and the 10-minute Cv were used only one-shot data by averaged over the first 1 minute and 10 minutes duration after 10-minute transient time, respectively. The 1-hour Cv was estimated as the logarithm of the 10-minute mean Cv averaged over the first 1 hour after the 10-minute transient. The coefficient Cv is designed to exhibit a value of unity for a Poisson random pulse train and is zero for a perfectly regular pulsation signal.

- Local variation (Lv)

  The conventional method of measuring HRV is adversely affected by slow variations in the heart rate and is also sensitive to artifacts and errors. There have been efforts to remove these artifacts, and modern methods employ machine learning techniques such as the state-space method[49,50]. Here we employ a simple metric called the local variation Lv, which was introduced for measuring the firing irregularity of the neurons in the brain[26–28]. The local variation Lv is defined as,

$$Lv = \frac{3}{n-1} \sum_{i=1}^{n-1} \left(\frac{I_i - I_{i+1}}{I_i + I_{i+1}} \right)^2,$$

where I_i and I_{i+1} are the ith and $i + 1$st RR intervals, respectively, and n is the total number of the intervals in a given duration. Note that the heartbeat that makes up the end of the ith RR interval and the start of the $i + 1$st RR interval was within the duration. The coefficient 3 in the definition of Lv was chosen so that Lv gives the value of unity for a Poisson pulse train[26]. Lv is zero for a regular pulsation.

Whereas Cv represents the global variability of an entire sequence and is sensitive to rate fluctuations, Lv detects the instantaneous variability of intervals. To demonstrate the difference in the workings of these metrics, we created synthetic pulse sequences. In Fig. 9a, RR intervals are lined up in a regular manner from long to short and then short to long, while in Fig. 9b, the identical set of RR intervals is presented in a random sequence. Accordingly, (a) represents a locally regular pulsation while the rate is slowly modulated, whereas (b) represents irregular pulsation.

The coefficient of variation Cv has identical values for (a) and (b) because the standard deviation and the mean are the same for both sequences. However, Lv successfully ignores slow heart rate modulation, and accordingly, it identified the difference in local irregularity between the two sequences. Note that we made the variations in RR intervals much larger than those of real cardiac beats so that their differences would be apparent.

In addition to these metrics, we have newly introduced the local-global variation ratio $Lg = Lv/Cv^2$. Its characteristics are analyzed in the following subsection. Lg takes a value much smaller than unity for a locally regular pulsation whose rate is slowly modulated (Fig. 9a). Lg takes the value 3/2 for a sequence in which RR intervals of mutually similar values are arranged randomly (Fig. 9b). Lg takes an even higher value of 3 for a pulsation in which long and short intervals alternate (Fig. 9c).

Analytical calculations of *Cv*, *Lv*, and *Lg* for some limiting cases. *Renewal process with gamma-distributed interval.* Expectation values are analytically available for a wide class of renewal processes in which pulse intervals are derived from the gamma distribution,

$$p_z(I) = (I/\tau)^{z-1} \exp(-I/\tau) / \Gamma(z),$$

where $\Gamma(z)$ is the gamma function defined as $\Gamma(z) = \int_0^\infty dt \, t^{z-1} \exp(-t)$. The Poisson process corresponds to the case of $z = 1$, and we can generate more regular pulse trains with the larger z.

In this case, the expectation value of Lv is obtained analytically as $3/(2z+1)$, and the expectation value of Cv is obtained as $1/\sqrt{z}$. Accordingly, the local-global variation ratio Lg is

$$Lg = Lv/Cv^2 = 3z/(2z+1).$$

While $Lg = 1$ for the Poisson process with $z = 1$, $Lg = 3/2$ for a regular pulse train at a limit of $z \to \infty$.

Sequences consisting of long and short pulse intervals. The local-global variation ratio Lg is undefined for a perfectly regular pulse train, for which both Lv and Cv^2 are zero. Though we have obtained $Lg = 3/2$ for a regular pulse train in a limiting case of the renewal process with gamma interval distribution $z \to \infty$, it might be sensitive to the arrangement of pulse intervals. Here we consider long pulse trains in which equal numbers of long and short intervals are arranged in various orders. Let the long and short intervals be denoted as $\tau + \delta$ and $\tau - \delta$, respectively ($\tau > \delta > 0$). Because the mean interval and standard deviation are τ and δ, respectively, the coefficient of variation Cv is δ/τ. By denoting the switching probability between long and short intervals as p, the local variation Lv is given as $3p\,\delta^2/\tau^2$. Accordingly, the local-global variation ratio Lg is obtained as

$$Lg = Lv/Cv^2 = 3p.$$

Thus $Lg = 3/2$ if long and short intervals alternate randomly, in which case the switching probability between long and short is half, $p = 1/2$. But Lg can be as high as 3 if the long and short intervals always alternate ($p = 1$). Lg can be close to zero for a long pulse train in which switching between long and short intervals occurs very rarely ($p \approx 0$). In this way, the local-global variation ratio may discriminate the order of long and short intervals, even if the difference between the long and short intervals is very small ($\delta \ll \tau$).

Data availability

Raw patient-related data are protected and are not available due to data privacy laws. Processed data are available from Yoshinori Marunaka upon request under the permission of the ethical committees.

References

1. Go, A. S. *et al.* Prevalence of diagnosed atrial fibrillation in adults: National implications for rhythm management and stroke prevention: The anticoagulation and risk factors in atrial fibrillation (ATRIA) study. *JAMA* **285**, 2370. https://doi.org/10.1001/jama.285.18.2370 (2001).
2. Collaborators, G. . C. o. D., & Roth, G. A. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1736 https://doi.org/10.1016/s0140-6736(18)32203-7 (2018).
3. Benjamin, E. J. *et al.* American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics and Subcommittee, Heart disease and stroke statistics-2019 update: A report from the American heart association. *Circulation* **139**, e56. https://doi.org/10.1161/cir.0000000000000659 (2019).
4. Dagenais, G. R. *et al.* Variations in common diseases, hospital admissions, and deaths in middle-aged adults in 21 countries from five continents (PURE): a prospective cohort study. *Lancet* **395**, 785. https://doi.org/10.1016/s0140-6736(19)32007-0 (2020).
5. Ribeiro, A. H., Ribeiro, M. H., Paixão, G. M. M., Oliveira, D. M., Gomes, P. R., Canazart, J. A., Ferreira, M. P. S., Andersson, C. R., Macfarlane, P. W., M. W. Jr., Schön, T. B., & Ribeiro, A. L. P. Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nat. Commun.* **11**, 1760, https://doi.org/10.1038/s41467-020-15432-4 (2020).
6. Saba, M. M., Ventura, H. O., Saleh, M. & Mehra, M. R. Ancient Egyptian medicine and the concept of heart failure. *J. Cardiac Fail.* **12**, 416 (2006).
7. Hajar, R. The pulse in ancient medicine part 1. *Heart Views* **19**, 36 (2018).
8. Kalehoff, J. P., & Oparil, S. The story of the silent killer. A history of hypertension: Its discovery, diagnosis, treatment, and debates. *Curr. Hypertens. Rep.* **22**, 72 (2020).
9. Barold, S. S. Willem Einthoven and the birth of clinical electrocardiography a hundred years ago. *Card. Electrophysiol. Rev.* **7**, 99 (2003).
10. Blackburn, H., Keys, A., Simonson, E., Rautaharju, P. & Punsar, S. The electrocardiogram in population studies. *Circulation* **21**, 1160. https://doi.org/10.1161/01.cir.21.6.1160 (1960).
11. Somerville, S., Somerville, J., Croft, P. & Lewis, M. Atrial fibrillation: A comparison of methods to identify cases in general practice. *Br. J. Gen. Pract.* **50**, 727 (2000).
12. Langley, P. *et al.* Accuracy of algorithms for detection of atrial fibrillation from short duration beat interval recordings. *Med. Eng. Phys.* **34**, 1441. https://doi.org/10.1016/j.medengphy.2012.02.002 (2012).
13. Yeh, S. Y., Forsythe, A. & Hon, E. H. Quantification of fetal heart beat-to-beat interval differences. *Obstet. Gynecol.* **41**, 355 (1973).
14. Shaffer, F. & Ginsberg, J. P. An overview of heart rate variability metrics and norms. *Front. Public Health* **5**, 258. https://doi.org/10.3389/fpubh.2017.00258 (2017).
15. Malik, M. & Camm, A. J. Heart rate variability. *Clin. Cardiol.* **13**, 570 (1990).
16. Bonnemeier, H. *et al.* Circadian profile of cardiac autonomic nervous modulation in healthy subjects. *J. Cardiovasc. Electrophysiol.* **14**, 791 (2003).
17. Akselrod, S., Gordon, D., Ubel, F. A., Shannon, D.C., Berger, A., & Cohen, R.J., Power spectrum analysis of heart rate fluctuation: A quantitative probe of beat-to-beat cardiovascular control. *Science* **213**, 220 (1981).

18. Berntson, G. G. *et al.* Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology* **34**, 623 (1997).
19. Task Force of the European Society of Cardiology, the North American Society of Pacing and Electrophysiology, Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Eur. Heart J.* **17**, 354 https://doi.org/10.1093/oxfordjournals.eurheartj.a014868 (1996).
20. Acharya, U. R., Joseph, K. P., Kannathal, N., Lim, C. M. & Suri, J. S. Heart rate variability: A review. *Med. Biol. Eng. Comput.* **44**, 1031 (2006).
21. Malliani, A., Pagani, M., Lombardi, F. & Cerutti, S. Cardiovascular neural regulation explored in the frequency domain. *Circulation* **84**, 482 (1991).
22. Hannun, A. Y. *et al.* Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat. Med.* **25**, 65. https://doi.org/10.1038/s41591-018-0268-3 (2019).
23. Raghunath, S. *et al.* Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. *Nat. Med.* **26**, 886 https://doi.org/10.1038/s41591-020-0870-z (2020).
24. Aziz, S., Ahmed, S. & Alouini, M.-S. ECG-based machine-learning algorithms for heartbeat classification. *Sci. Rep.* **11**, 18738. https://doi.org/10.1038/s41598-021-97118-5 (2021).
25. Siontis, K. C., Noseworthy, P. A., Attia, Z. I. & Friedman, P. A. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nat. Rev. Cardiol.* **18**, 465. https://doi.org/10.1038/s41569-020-00503-2 (2021).
26. Shinomoto, S., Shima, K. & Tanji, J. Differences in spiking patterns among cortical neurons. *Neural Comput.* **15**, 2823 (2003).
27. Shinomoto, S. *et al.* Relating neuronal firing patterns to functional differentiation of cerebral cortex. *PLoS Comput. Biol.* **5**, e1000433 (2009).
28. Mochizuki, Y. *et al.* Similarity in neuronal firing regimes across mammalian species. *J. Neurosci.* **36**, 5736 (2016).
29. Brennan, M., Palaniswami, M. & Kamen, P. Do existing measures of Poincare plot geometry reflect nonlinear features of heart rate variability?. *IEEE Trans. Biomed. Eng.* **48**, 1342 (2001).
30. Porta, A. *et al.* Temporal asymmetries of short-term heart period variability are linked to autonomic regulation. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **295**, R550 (2008).
31. Go, A. S. *et al.*, Executive summary: Heart disease and stroke statistics-2014 update a report from the American heart association. *Circulation* **129**, 399. https://doi.org/10.1161/01.cir.0000442015.53336.12 (2014).
32. Wang, T. J. *et al.* Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality. *Circulation* **107**, 2920. https://doi.org/10.1161/01.cir.0000072767.89944.6e (2003).
33. Wolf, P. A., Abbott, R. D. & Kannel, W. B. Atrial fibrillation as an independent risk factor for stroke: The Framingham Study. *Stroke* **22**, 983. https://doi.org/10.1161/01.str.22.8.983 (1991).
34. Benjamin, E. J. *et al.* Impact of atrial fibrillation on the risk of death. *Circulation* **98**, 946. https://doi.org/10.1161/01.cir.98.10.946 (1998).
35. Conen, D. *et al.* Premature atrial contractions in the general population. *Circulation* **126**, 2302. https://doi.org/10.1161/circulationaha.112.112300 (2012).
36. Marcus, G. M. Evaluation and management of premature ventricular complexes. *Circulation* **141**, 1404. https://doi.org/10.1161/circulationaha.119.042434 (2020).
37. McManus, D. D. *et al.* PULSE-SMART: Pulse-based arrhythmia discrimination using a novel smartphone application. *J. Cardiovasc. Electrophysiol.* **27**, 51. https://doi.org/10.1111/jce.12842 (2016).
38. Bumgarner, J. M. *et al.* Smartwatch algorithm for automated detection of atrial fibrillation. *J. Am. Coll. Cardiol.* **71**, 2381. https://doi.org/10.1016/j.jacc.2018.03.003 (2018).
39. Matthews, B. W. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta* **405**, 442 (1975).
40. Tulppo, M. P., Makikallio, T. H., Takala, T., Seppanen, T. & Huikuri, H. V. Quantitative beat-to-beat analysis of heart rate dynamics during exercise. *Am. J. Physiol. Heart Circul. Physiol.* **271**, H244 (1996).
41. Kamen, P. W., Krum, H. & Tonkin, A. M. The correlation dimension of heart rate variability reflects cardiac autonomic activity. *Ann. Noninvasive Electrocardiol.* **2**, 206 (1997).
42. Kiviniemi, A. M. *et al.* Cardiac vagal outflow after aerobic training by analysis of high-frequency oscillation of the R-R interval. *Eur. J. Appl. Physiol.* **96**, 686 (2006).
43. Bauer, A. *et al.* Reference values of heart rate variability. *Heart Rhythm* **14**, 302 (2017).
44. La Rovere, M. T., Porta, A. & Schwartz, P. J. Autonomic control of the heart and its clinical impact. A personal perspective. *Front. Physiol.* **11**, 582 (2020).
45. Moody, G., & Mark, R., The MIT-BIH arrhythmia database on CD-ROM and software for use with it. *Proc. Comput. Cardiol.* https://doi.org/10.1109/cic.1990.144205 *(1990).*
46. Moody, G. B. & Mark, R. G. The impact of the MIT-BIH arrhythmia database. *IEEE Eng. Med. Biol. Mag.* **20**, 45. https://doi.org/10.1109/51.932724 (2001).
47. Goldberger, A. L. *et al.* PhysioBank, PhysioToolkit, and PhysioNet. *Circulation* **101**, E215. https://doi.org/10.1161/01.cir.101.23.e215 (2000).
48. Makowski, D. *et al.* NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behav. Res. Methods* https://doi.org/10.3758/s13428-020-01516-y *(2021).*
49. Barbieri, R., Matten, E. C., Alabi, A. A. & Brown, E. N. A point-process model of human heartbeat intervals: new definitions of heart rate and heart rate variability. *Am. J. Physiol. Heart Circul. Physiol.* **288**, H424 (2005).
50. Citi, L., Brown, E. N. & Barbieri, R. A real-time automated point-process method for the detection and correction of erroneous and ectopic heartbeats. *IEEE Trans. Biomed. Eng.* **59**, 2828 (2012).

Acknowledgements

Author contributions

S.S. and Y.T. contributed equally in performing analysis. Y.M. contributed data. S.S. designed the study and wrote the main manuscript text. All authors discussed the results and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.S.