Computers in Biology and Medicine 109 (2019) 280–289

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Novel gridded descriptors of poincaré plot for analyzing heartbeat interval time-series

Chang Yan[a], Peng Li[a,b,1,*], Changchun Liu[a,**], Xinpei Wang[a], Chunyan Yin[a], Lianke Yao[a]

[a] School of Control Science and Engineering, Shandong University, Jinan, 250061, China
[b] Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, 02115, USA

ARTICLE INFO

Keywords:
Poincaré plot
Gridded distribution rate
Gridded distribution entropy
Short-term time-series
Coronary artery disease

ABSTRACT

A Poincaré plot is a return map that geometrically elucidates the progression of a time-series. It has frequently been used in heart rate variability analyses. However, algorithms for dedicatedly dissecting the shape of this geometrical plot are yet to be established. In this study, we proposed a gridded Poincaré plot by coarse-graining the original graph and using the newly proposed one, defined two novel measures, namely gridded distribution rate (GDR) and gridded distribution entropy (GDE). The GDR essentially represents the percentage of grids with points, while the GDE estimates the Shannon entropy of the grid weight; that is, the number of points in each grid. The performances of the two measures were examined using both theoretical data with known dynamics and experimental short-term RR interval time-series, and they were compared with several existing metrics. Simulation tests demonstrated that both the GDR and GDE could distinguish among different dynamics, while all the compared methods failed. The experimental results further indicated the ability of the GDR and GDE to differentiate healthy young people from healthy aged adults as well as distinguish healthy subjects from patients with coronary artery disease. Our results suggest that the proposed GDR and GDE may better characterize the Poincaré plot in terms of differentiating between varying dynamical regimes, and between human physiological or pathological conditions. Further studies are warranted to establish their feasibility in evaluating cardiovascular functions in clinical practice.

1. Introduction

Heart rate variability (HRV) represents the tiny beat-to-beat fluctuations in the heart rate. It has been widely accepted as a noninvasive physiological biomarker for evaluating cardiac autonomic control. Various mathematical algorithms have been established to analyze HRV quantitatively [1,2]. Importantly, during the past decades, numerous nonlinear approaches have been proven to be capable of offering new insights into the changes in the HRV under different physiological or pathological conditions [2–6].

The Poincaré plot is a frequently used graphical tool for characterizing the nonlinear features of the HRV [7–12]. It plots each RR interval against the subsequent interval to visualize the acceleration and deceleration heartbeat patterns. The Poincaré plot has been interpreted from several different angles to understand its geometric patterns. For example, ellipse fitting has been used to approximate the graph shape, and thereby, the major and minor axes of the best fitting ellipse have been defined as quantitative indices [6,12,13]. The aging effect and sex difference in these indices have been observed [14]. Moreover, their changes in diseased subjects, such as patients with dilated cardiomyopathy [15], cardiac arrhythmias [16], and heart failure [11], have been studied. The geometric asymmetry of the Poincaré plot is a further angle that has been applied intensively for quantitatively describing the graph [17–19]. It is accepted that the beat-to-beat increase or decrease in the heart rate is physiologically asymmetric [20]. Several indices have been established to assess this asymmetry feature, including the Porta's index (PI) [21], Guzik's index (GI) [10], slope index (SI) [22], and area index (AI) [23]. Based on these measures, previous studies have documented changes in the asymmetry of the Poincaré plot under several pathological conditions; for example, myocardial infarction [24,25], post-myocardial infarction [26], obstructive sleep apnea [27], and type 1 diabetes [28].

The geometric shape itself or the manner in which points are distributed in the Poincaré plot may also be considerably important. Woo

* Corresponding author. School of Control Science and Engineering, Shandong University, Jinan, 250061, China.
** Corresponding author. School of Control Science and Engineering, Shandong University, Jinan, 250061, China.
E-mail addresses: pli@sdu.edu.cn (P. Li), changchunliu@sdu.edu.cn (C. Liu).
[1] Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, 02115, USA.

https://doi.org/10.1016/j.compbiomed.2019.04.015
Received 24 December 2018; Received in revised form 18 April 2019; Accepted 18 April 2019
0010-4825/

et al. [11,29] qualitatively demonstrated that the Poincaré plot for healthy subjects exhibited a “comet-like” shape. Park et al. [30] found that the Poincaré plots of atrial fibrillation patients exhibited irregular shapes. Despite the existence of these qualitative studies, quantitative algorithms are yet to be established. In this study, two novel quantitative indices were proposed, namely the gridded distribution rate (GDR) and gridded distribution entropy (GDE), based on a gridded characterization of the Poincaré plot. Their performance in studying aged individuals and patients with coronary artery disease (CAD) was demonstrated, and compared with ellipse fit-based metrics and asymmetry indices.

2. Method

2.1. Proposed GDR and GDE algorithms

For an RR interval time-series $\{RR_1, RR_2, ..., RR_i, RR_{i+1}, ..., RR_N, 1 \le i \le N\}$ consisting of N RR intervals, the GDR and GDE can be calculated according to the following steps:

1) Anomalous interval removal

The RR interval sequence is first subjected to an anomalous interval removal process. In this study, an RR interval is considered abnormal if it is shorter than 0.3 s or longer than 2 s, or if it exhibits a change of > 20% compared to the previous RR interval. The square filter combined with a quotient filter is used to remove these identified anomalous intervals [31]. This process is similar to the procedure of removing the anomalous-related points in the Poincaré plot; for example, for an identified anomalous interval RR_i, the two related points [RR_{i-1}, RR_i] [RR_i, RR_{i+1}] in the Poincaré plot will both be removed so that it will not produce an artificial point [RR_{i-1}, RR_{i+1}].

2) Normalization

After removing the anomalous intervals, the RR interval sequence is normalized to be within the 0–1 range, based on a min-max normalization:

$$RR_i^N = (RR_i - \min(RR_i))/(\max(RR_i) - \min(RR_i)), \quad (1)$$

where RR_i indicates the ith RR interval; RR^N means the normalized RR interval sequence; and RR_i^N denotes the ith element in sequence RR^N.

3) Construction of gridded Poincaré plot

The original Poincaré plot is first obtained by plotting each RR interval against its subsequent interval. Once the sequence has been normalized, the graph should span a 1 × 1 square on the two-dimensional coordinate plane. We coarse-grain this graph along both the x- and y-directions, using $n \times n$ evenly distributed horizontal and vertical grid lines (see Fig. 1) to construct the gridded Poincaré plot. The variable n serves as a coarse-graining parameter that determines how fine the grid is.

4) Calculation

The GDR is defined as the average number of grids filled with at least one point; that is,

$$GDR = \frac{a}{n^2}, \quad (2)$$

where a is the number of grids filled with at least one point.

The GDE is defined by the formula for the Shannon entropy:

$$GDE = -\sum_{j=1}^{n^2} p_j \log p_j, \quad (3)$$

where j is the jth grid in the gridded Poincaré plot, and the probability p_j indicates the frequency of points in each grid, defined by

$$p_j = \frac{b}{N-1}, \quad (4)$$

where b is the number of points in the jth grid.

Fig. 1 provides a schematic explanation of the above four steps.

2.2. Selection of parameters

The coarse-graining level n is the only parameter used in the GDR and GDE calculations. It is obvious that, with $n = 1$ or 2, each grid will have been filled with points for most time-series types. In the current work, $n = 3$ was used through a sufficiently larger number, depending on the signal length (length of RR interval time-series), and sequences of the GDR and GDE were calculated (that is, for each assignment of n, a GDR value and GDE value were obtained).

2.3. Existing indices for poincaré plot

2.3.1. Ellipse fit-based metrics

In particular, the major and minor axes of the best fitting ellipse (SD1 and SD2), and their ratio (SD1/SD2), are used as quantitative measures [32]. These can be calculated by

$$SD1 = \sqrt{VAR\frac{RR_i - RR_{i+1}}{\sqrt{2}}}, \quad (5)$$

$$SD2 = \sqrt{VAR\frac{RR_i + RR_{i+1}}{\sqrt{2}}}, \quad (6)$$

$$SD1/SD2 = \frac{SD1}{SD2}, \quad (7)$$

where VAR is the variance.

2.3.2. Asymmetry indices

A. PI

The PI is the percentage of points below the line of identity (LI) [21]; that is,

$$PI = \frac{b}{m} \times 100, \quad (8)$$

where b is the number of points below the LI and m is the total number of points, except for those on the LI.

B. GI

The GI is proportional to the distance of points above the LI [10]; that is,

$$GI = \frac{\sum_{i=1}^{l} D_i}{\sum_{i=1}^{m} D_i} \times 100, \quad (9)$$

where l represents the number of points above the LI; m is the number of points in the Poincaré plot, except for those on the LI; and D_i is the distance from point P_i to the LI, which can be calculated as

$$D_i = \frac{|RR_{i+1} - RR_i|}{\sqrt{2}}. \quad (10)$$

C. SI

The SI is the ratio of phase angles (with respect to the LI) of the points above the LI [22]. That is,

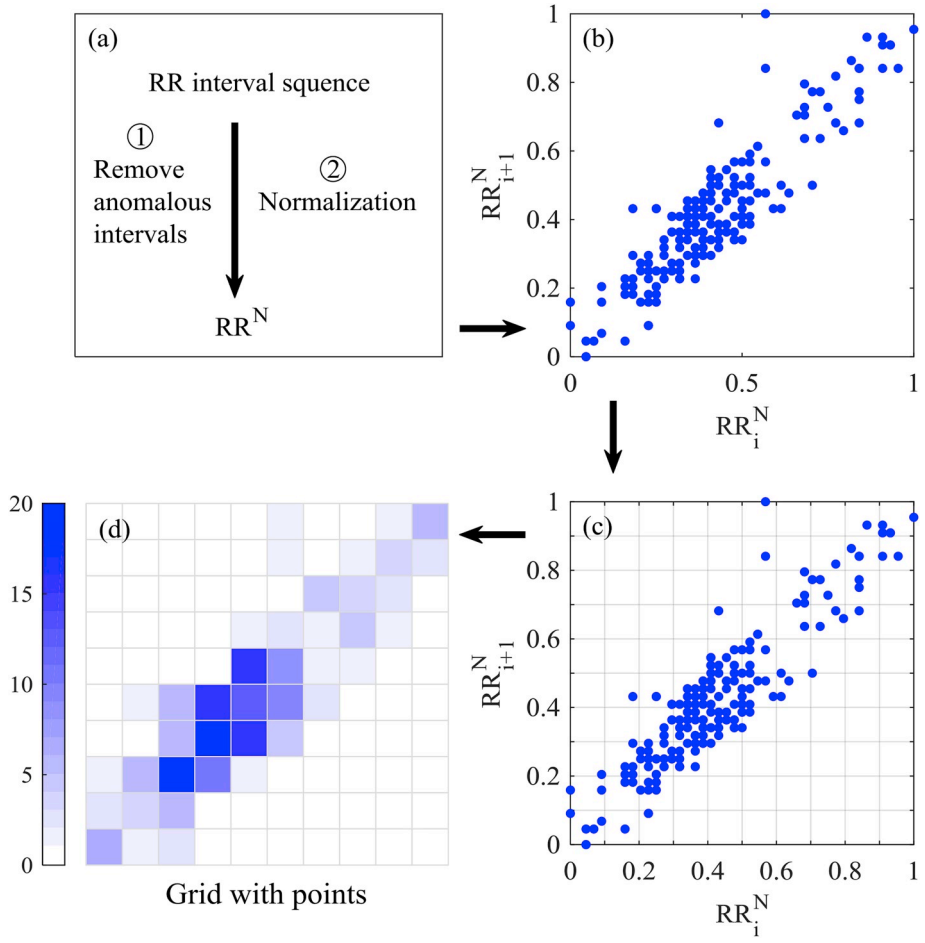

Fig. 1. Schematic representation of proposed GDR and GDE algorithms: (a) preprocessing of RR intervals; (b) RR^N (normalized RR record) in Poincaré plot; (c) even division of Poincaré plot into n*n blocks (n = 10 indicated); and (d) grids with points colorcoded for weight of each grid (number of points).

$$SI = \frac{\sum_{i=1}^{l} |R\theta_i|}{\sum_{i=1}^{m} |R\theta_i|} \times 100, \tag{11}$$

where l is the number of points above the LI and m is the number of points not on the LI. Moreover, $R\theta_i = \theta_{LI} - \theta_i$, where θ_{LI} is the phase angle of the LI and θ_i is the phase angle of every point, which is defined as $\theta_i = atan(RR_{i+1}/RR_i)$.

D. AI

The AI is the ratio of the cumulative area of the sectors corresponding to the points located above the LI [23]; that is,

$$AI = \frac{\sum_{i=1}^{l} |S_i|}{\sum_{i=1}^{m} |S_i|} \times 100, \tag{12}$$

where l and m represent the number of points above the LI and total number of points in the Poincaré plot not on the LI, respectively. Furthermore, S_i is the area of the sector, which can be calculated by

$$S_i = \frac{1}{2} \times R\theta_i \times r^2, \tag{13}$$

where r is the radius of the sector.

2.4. Simulation tests

2.4.1. Synthetic data

Examining the GDR and GDE using synthetic signals with known and predictable dynamical properties may aid in obtaining an improved understanding of their performance. In this study, we applied simulated chaotic time-series, Gaussian noises, MIX(p) processes, and periodic signals. The logistic attractor $x(n+1) = \omega \times x(n) \times (1 - x(n))$ with $\omega = 4.0$ was considered for the chaotic series. The MIX(p) process is, by nature, a sinusoidal signal with length N, in which $N \times p$ randomly selected points are replaced with independent, identically distributed random noise [33]. In this case, we used $p = 0$, 0.1, and 0.2 to produce a periodic signal and two time-series with different randomness levels. The Gaussian noise was generated by the random number function (randn) in MATLAB. All synthetic data were generated by customized MATLAB programs (version R2016a, Mathworks Inc., MA, USA). Fig. 2 presents examples of the five synthetic time-series and their corresponding Poincaré plots.

To eliminate random factors, 30 realizations of each model were performed. Specifically, 30 randomly selected initial values were adopted for the logistic attractor to generate 30 realizations of chaotic time-series, while 30 independent realizations were directly performed for the MIX(p) processes, Gaussian noise, and sinusoidal signal. Moreover, for each realization of the logistic map, we allowed a transient period (200 data points) prior to sampling a dataset to ensure that the dynamics had settled down on its trajectory. Considering that we were focusing on short-term RR interval time-series (essentially 5 min RR interval data), the length of each synthetic signal was predefined as N = 400 points.

2.4.2. GDR and GDE calculations

The GDR and GDE were calculated for each realization of the synthetic signals. As illustrated in Fig. 2, we expected a higher GDR value

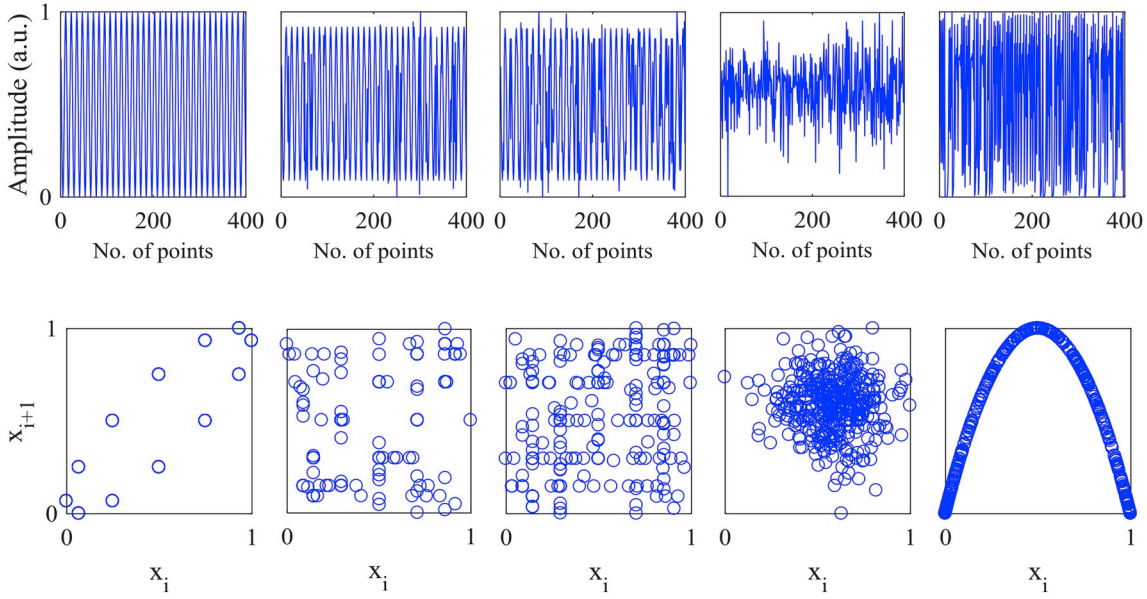

Fig. 2. Synthetic data. The top panels illustrate the five synthetic time-series: (from left to right) periodic series, MIX(0.1) process, MIX(0.2) process, Gaussian noise, and logistic chaotic time-series. The corresponding Poincaré plots of the five time-series are presented in the bottom panels, where x_i means the ith value of a time-series.

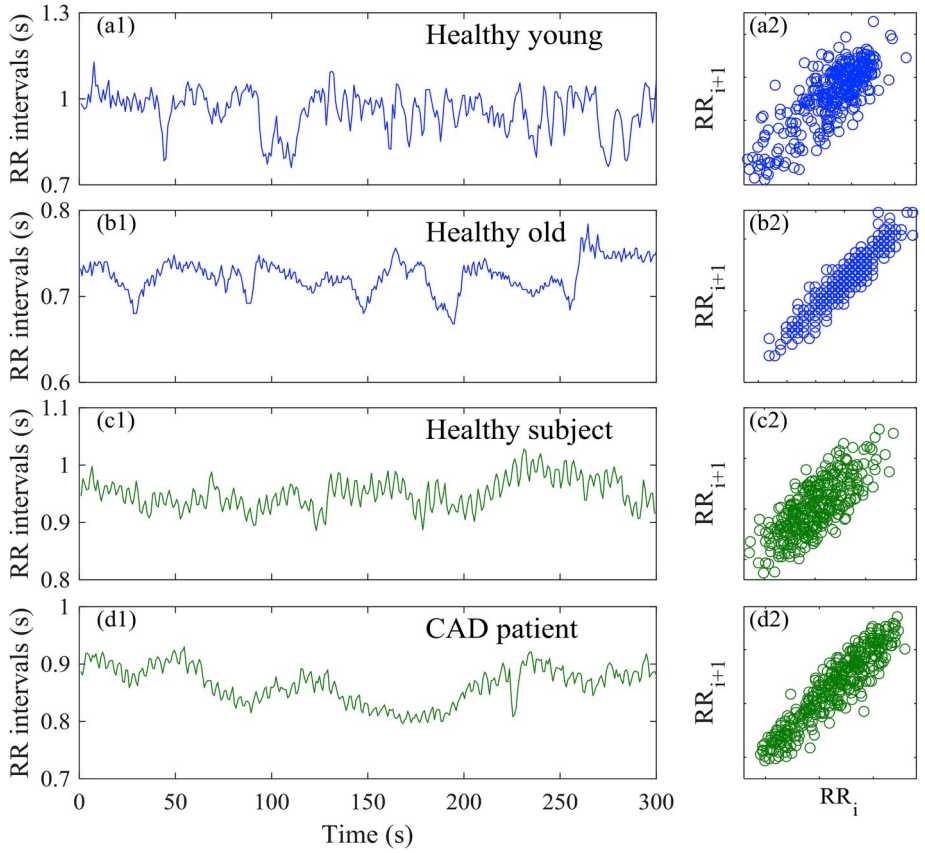

Fig. 3. Examples of RR interval time-series and corresponding Poincaré plots: (a) signal from representative healthy young subject; (b) signal from representative healthy old subject; (c) signal from representative healthy subject; and (d) the signal from representative CAD patient.

in the Gaussian noise, followed by the two MIX(p) processes, and finally, the sinusoidal signal. The logistic chaotic time-series could initially have a GDR as low as the level for the sinusoidal signal. However, as the coarse-graining level increased, the GDR could reach beyond the levels of the MIX(p) processes, as the points in the graph appear to be continuously distributed, rather than aggregated into clusters. There could ultimately be more grids with points filled in the Poincaré plot of the logistic chaotic time-series than those in the two tested MIX(p) processes. The same inference could also apply to the GDE.

Table 1
Basic characteristics of subjects.

Variables	Healthy controls	CAD patients	*p*
No. (female/male)	35 (16/19)	30 (10/20)	0.31
Age (years)	58.3 ± 6.1	59.4 ± 9.3	0.77
Height (cm)	168.0 ± 7.2	166.2 ± 8.7	0.83
Weight (kg)	65.6 ± 10.5	67.9 ± 9.9	0.39
BMI (kg/m^2)	23.8 ± 2.8	24.5 ± 2.5	0.30
HR (s)	53.8 ± 6.1	57.2 ± 8.7	0.07

Data are expressed as number (female/male) or mean ± standard deviation. Abbreviations: No. = number, BMI = body mass index, HR = heart rate.

2.5. Experimental tests

2.5.1. Comparison of healthy aged and healthy young groups

The Fantasia database was used, which is publicly available online from the PhysioBank Archives [34]. It consists of 20 elderly subjects (74.6 ± 4.5 [mean ± standard deviation, unless otherwise indicated] years old) and 20 young subjects (26.0 ± 4.3 years old; $p < 0.0001$ versus the elderly subjects) with rigorously screened health statuses. All subjects underwent a 120 min continuous ECG collection while lying straight on their backs. The ECG was digitized at 250 Hz. Previous studies have documented differences between the two groups in various RR interval metrics, including time- and frequency-domain measures, entropy measures, and symbolic dynamics [35].

The database also includes annotated heartbeat information to facilitate the analysis of RR interval time-series. These heartbeats were identified by means of an automated algorithm, followed by expert inspections [34]. Fig. 3 (a-b) shows two examples of the RR interval time-series from a young and an old subject, respectively, and their corresponding Poincaré plots. To understand the performance of the GDR and GDE in handling short-term RR interval time-series, two analysis protocols were applied: 1) using a short-term HRV segment randomly selected from the complete HRV recording (corresponding to a 5 min ECG segment); and 2) using 10 subsequent short-term HRV segments without overlap (corresponding to 10 subsequent 5 min ECG episodes, starting from the beginning of the ECG).

2.5.2. Comparison of CAD patients and health control subjects

To examine the performance of the proposed GDR and GDE algorithms further, we revisited the 5 min ECG data collected from our previous human study [36–40]. A total of 30 CAD patients and 35 age- and sex-matched healthy volunteers were recruited and provided informed consent in that study. Their basic characteristics are presented in Table 1. All healthy subjects underwent routine ECG and echocardiography examinations, along with medical history questionnaires, to confirm their health status. The CAD patients were recruited from those who were scheduled for interventional surgery and data were collected one day prior to the surgery. Patients were enrolled whose coronary angiography suggested that > 50% stenosis presented in at least one main coronary branch. Standard limb lead II ECG data were collected continuously for 5 min in a quiet, temperature-controlled clinical measurement room (25 ± 3 °C) at the Shandong Provincial Qianfoshan Hospital by a cardiovascular function detection device (CV FD-II, Huiyironggong Technology Co., Ltd, Jinan, China). The data were sampled at 1000 Hz. The study obtained full approval from the Institutional Review Board of Shandong Provincial Qianfoshan Hospital, and it was conducted according to the principles in the Declaration of Helsinki and its following amendments. The R-peaks were detected and ectopic R-peaks identified based on a template-matching approach [38], followed by visual inspections by experts. Intervals of consecutive normal R-waves formed the RR interval time-series. Two examples of RR interval time-series from a healthy subject and a CAD patient, respectively, and their corresponding Poincaré plots are shown in Fig. 3 (c-d).

2.6. Statistical analyses

We implemented both the proposed and existing algorithms in MATLAB. The results were first subjected to the Shapiro–Wilk test to examine the normality [41]. The student's *t*-test was used to examine the group differences (young versus old; healthy versus CAD) in the GDR and GDE if the normality hypothesis was not rejected; the Mann–Whitney *U* test was used otherwise. The effect size was estimated by Cohen's *d* static. The statistical significance was accepted at $p < 0.05$. An effect size of $d > 0.5$ was considered as large [42,43]. The area under the receiver operating characteristic (ROC) curve was used for further examination of the classification performance. All statistical analyses were performed in MATLAB.

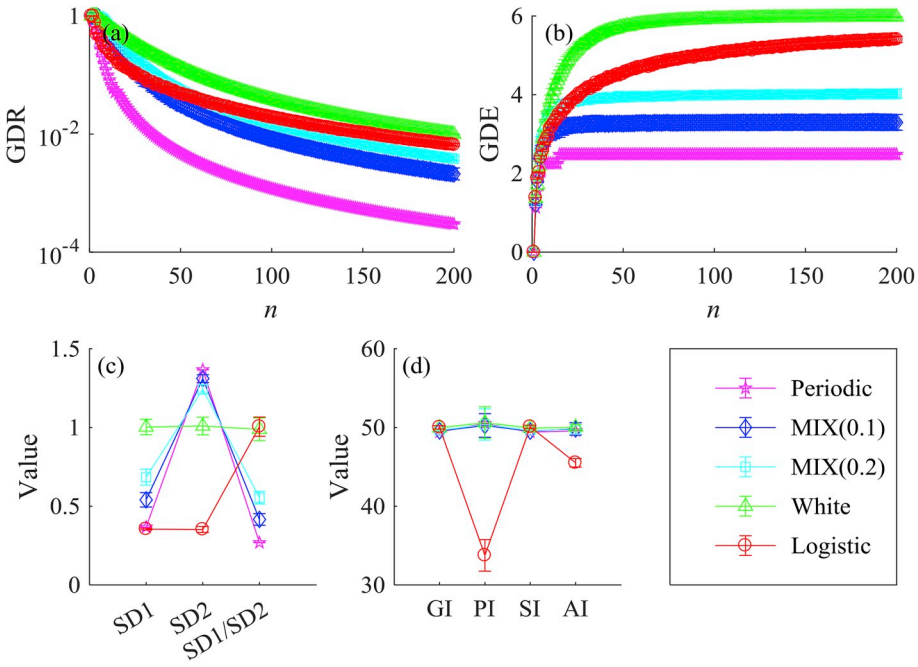

Fig. 4. Results of Poincaré plot metrics for Gaussian noises, chaotic time-series, MIX(0.2), MIX(0.1), and periodic time-series: (a) GDR results (base-10 logarithmic transformed) with *n* ranging from 1 to 200; (b) GDE results with the same range of *n*; (c) elliptical indices (SD1, SD2, SD1/SD2); and (d) asymmetry indices (GI, PI, SI, and AI). The error bar indicates the standard error from 30 realizations.

Table 2
Results of Poincaré plot metrics for Gaussian noises, chaotic time-series, MIX(0.2), MIX(0.1), and periodic time-series.

Index	Gaussian noises		Logistic chaotic time-series		MIX(0.2)		MIX(0.1)		Periodic time-series	
	Mean ± SD	Median ± IQR	Mean ± SD	Median ± IQR	Mean ± SD	Median ± IQR	Mean ± SD	Median ± IQR	Mean ± SD	Median ± IQR
GDR	−1.75 ± 0.00	−1.75 ± 0.01	−1.98 ± 0.02	−1.98 ± 0.03	−2.17 ± 0.04	−2.16 ± 0.05	−2.41 ± 0.05	−2.41 ± 0.07	−3.26 ± 0.00	−3.26 ± 0.00
GDE	5.96 ± 0.01	5.96 ± 0.02	5.25 ± 0.07	5.24 ± 0.10	3.98 ± 0.13	3.99 ± 0.17	3.31 ± 0.11	3.28 ± 0.14	2.48 ± 0.00	2.48 ± 0.00
SD1	1.00 ± 0.05	1.00 ± 0.07	0.35 ± 0.01	0.35 ± 0.01	0.67 ± 0.04	0.67 ± 0.06	0.55 ± 0.03	0.55 ± 0.04	0.37 ± 0.00	0.37 ± 0.00
SD2	1.00 ± 0.05	1.00 ± 0.07	0.35 ± 0.01	0.36 ± 0.02	1.25 ± 0.02	1.25 ± 0.02	1.30 ± 0.02	1.30 ± 0.02	1.37 ± 0.00	1.37 ± 0.00
SD1/SD2	1.00 ± 0.05	0.99 ± 0.06	1.00 ± 0.05	0.99 ± 0.08	0.53 ± 0.04	0.54 ± 0.05	0.42 ± 0.03	0.42 ± 0.04	0.27 ± 0.00	0.27 ± 0.00
GI	49.99 ± 0.15	49.99 ± 0.19	50.00 ± 0.17	49.98 ± 0.24	49.72 ± 0.16	49.64 ± 0.12	49.56 ± 0.06	49.56 ± 0.03	49.44 ± 0.00	49.44 ± 0.00
PI	50.13 ± 1.59	49.88 ± 1.50	33.38 ± 1.5	33.38 ± 1.25	50.29 ± 1.33	50.25 ± 1.75	50.33 ± 0.94	50.25 ± 1.50	50.50 ± 0.00	50.50 ± 0.00
SI	50.01 ± 0.18	50.01 ± 0.18	50.06 ± 0.17	50.05 ± 0.24	49.7 ± 0.18	49.64 ± 0.14	49.54 ± 0.08	49.52 ± 0.10	49.40 ± 0.00	49.40 ± 0.00
AI	49.94 ± 0.23	49.96 ± 0.34	45.45 ± 0.39	45.45 ± 0.61	49.75 ± 0.38	49.70 ± 0.43	49.61 ± 0.37	49.73 ± 0.62	49.57 ± 0.00	49.57 ± 0.00

SD: standard deviation; IQR: interquartile range; GDR and GDE: average of n = 100 to 200.

3. Results

3.1. Simulation results

Fig. 4 summarizes the GDR and GDE results for the synthetic signals, as well as the results of the existing metrics of the Poincaré plot. Moreover, the results are presented in Table 2 in terms of the mean ± standard deviation (SD) and median ± interquartile range (IQR). As the coarse-graining parameter n increased from 1 to 200, the GDR values all exhibited an initial sharp decrease and then became stable at approximately n = 80. To provide improved visualization of this changing pattern, we illustrate the base-10 logarithmic transform of the GDR in Fig. 4, instead of using its original linear scale. As expected, during the initial unstable stage, the GDR was the highest in terms of Gaussian noises and decreased gradually in the order of MIX (0.2), MIX(0.1), logistic chaotic time-series, and sinusoidal signals (Fig. 4a). At approximately n = 25, the GDR for the logistic chaotic time-series reached beyond the level for MIX(0.1), then beyond the level for MIX(0.2), and finally settled down. Furthermore, the GDE exhibited an initial sharp change before converging towards each stable level (Fig. 4b). This suggests the same ranking as for the GDR; that is, the GDE level gradually decreased in the order of Gaussian noises, logistic chaotic time-series, MIX(0.2), MIX(0.1), and sinusoidal signals. Surprisingly, none of the traditional metrics could fully discriminate among these, as illustrated in Fig. 4c and d, suggesting that these traditional metrics may not be sensitive to the underlying time-series dynamics, or cannot capture the relevant geometrical features. To understand the different abilities of the proposed metrics and existing traditional metrics in capturing the dynamic structural time-series characteristics further, we performed surrogate data tests. Specifically, the logistic chaotic time-series were randomly shuffled, and all metrics were recalculated for the shuffled data. The results are summarized in Fig. 5, which clearly indicates that both the GDR and GDE were quite sensitive to the changes in dynamical patterns caused by the random shuffling, while all traditional metrics failed, except for the PI and AI.

3.2. Comparison between healthy aged and healthy young groups

3.2.1. Using 5 min HRV segment randomly selected from complete time-series

Using the experimental data, similar transient sharp changes followed by steady stages were identified in the GDR and GDE, as illustrated in Fig. 6a1 and 6a2. Both the GDR and GDE in the stable stages suggested a decrease in the healthy aged group compared to the healthy young group. Of the seven traditional Poincaré plot indices (Fig. 6b1 to 6b9), only SD1 and SD2 exhibited significant differences. The corresponding p values, area under the curve (AUC) levels, and Cohen's d statics are summarized in Fig. 6c1 to 6c3, which clearly suggest superior performance of the GDR and GDE over the traditional metrics (Fig. 6d1 to 6d3).

3.2.2. Using 10 subsequent 5 min HRV segments without overlap

Relatively consistent performance of the GDR and GDE was achieved across the 10 HRV segments, as illustrated in Fig. 7. The p values, AUC values, and effect sizes behaved similarly with the change in n across all 10 segments, suggesting that the significantly improved performance of the GDR and GDE illustrated above was not by chance. This could be further clarified when demonstrating their performance together with the traditional metrics, as explained in Fig. 7(a3) to (c3). In all 10 HRV segments, the GDR and GDE outperformed the existing metrics. Note that the performance of SD1 and SD2 also appeared to be relatively stable, and significant group differences could be identified for the majority of the 10 segments. However, both the AUC values and effect sizes, as reflected by the Cohen's d static, were relatively low, which is consistent with the findings when using only one 5 min HRV segment (Fig. 6).

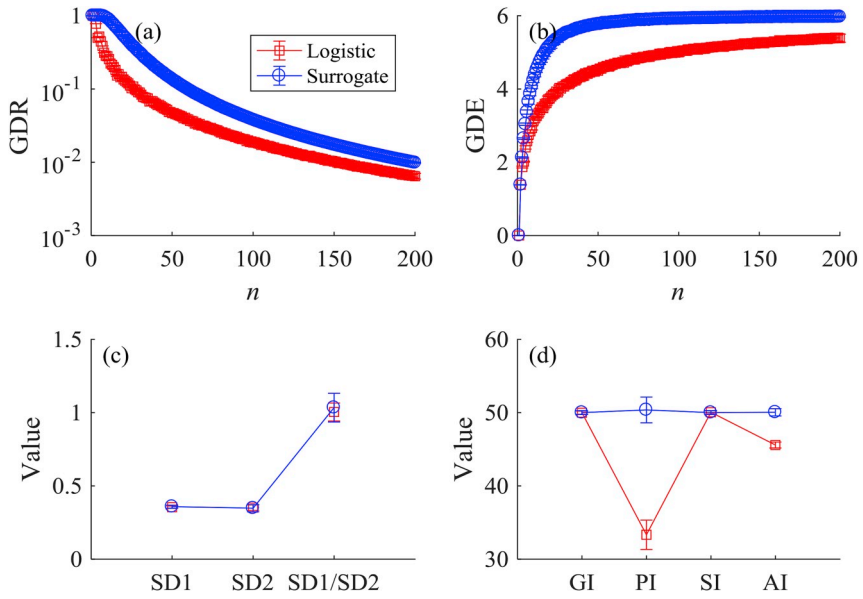

Fig. 5. Results of surrogate data tests: (a) GDR results (base-10 logarithmic transformed) with n ranging from 1 to 200; (b) GDE results with the same range of n; (c) elliptical indices (SD1, SD2, and SD1/SD2); and (d) asymmetry indices (GI, PI, SI, and AI). The error bar indicates the standard error of 30 realizations for each series.

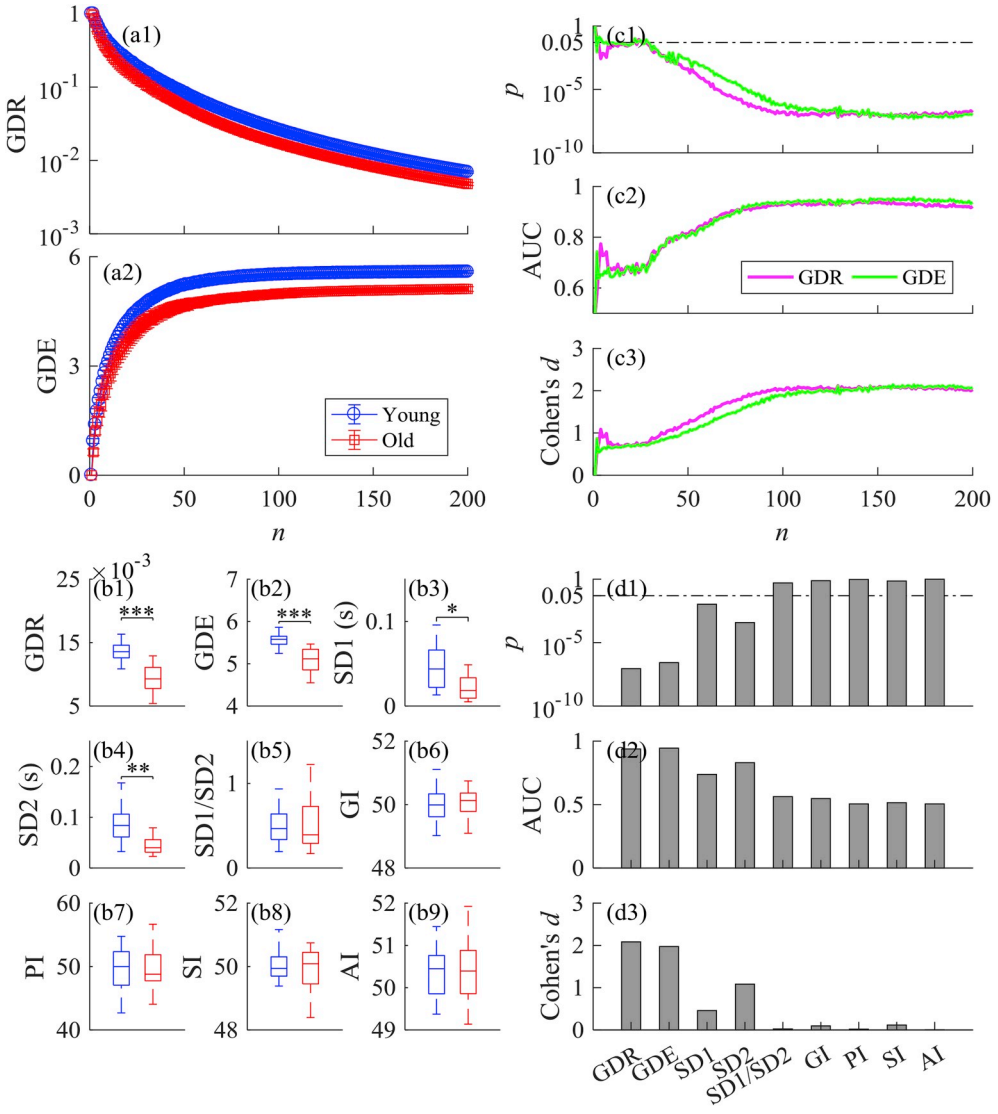

Fig. 6. Results of Poincaré plot measures for RR interval time-series of healthy young and old adults: (a1) GDR results (base-10 logarithmic transformed) with n ranging from 1 to 200; (a2) GDE results with the same range of n, where error bars indicate standard errors; (b1 to b9) comparisons of GDR (average of $n = 100$ to 200), GDE (average of $n = 100$ to 200), SD1, SD2, SD1/SD2, GI, PI, SI, and AI, respectively, between young healthy subjects and aging healthy group; (c1 to c3) performance of GDR and GDE for comparing young healthy subjects with aging healthy group in terms of p values, AUC results, and Cohen's d, respectively; and (d1 to d3) performance of existing indices in terms of p values, AUC results, and Cohen's d, respectively. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.0001$.

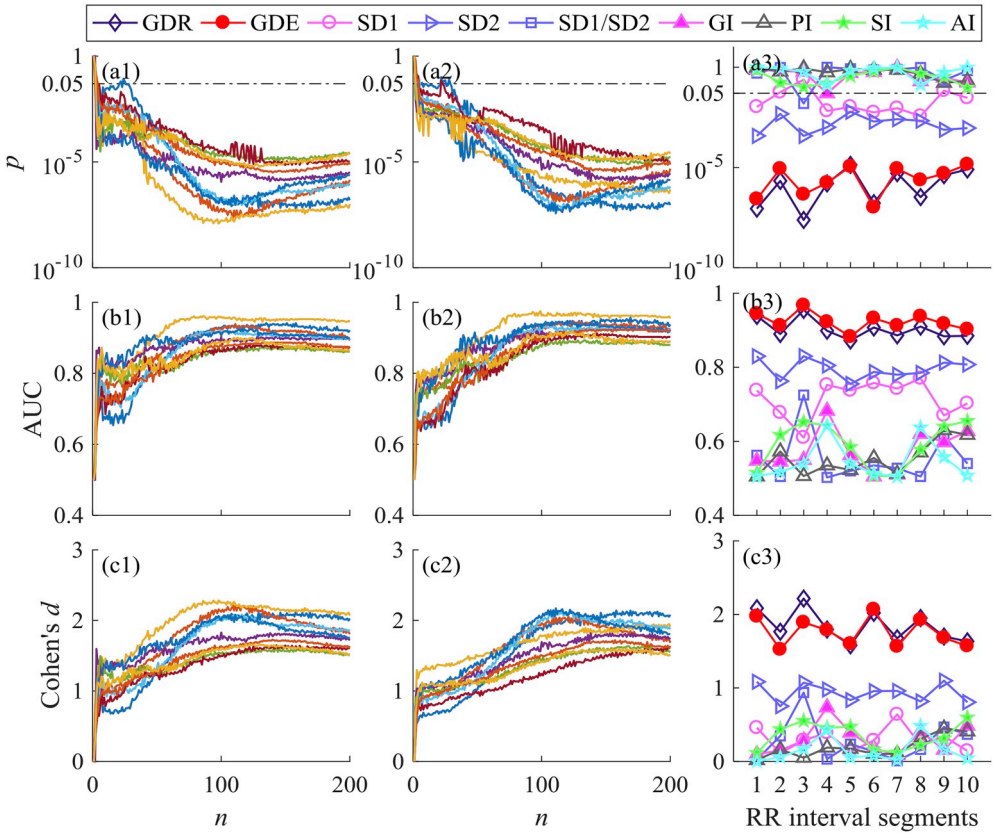

Fig. 7. Performance of GDR and GDE for comparing healthy young and healthy old groups in terms of p values, AUC results, and Cohen's d statics. The GDR and GDE results presented in the right panels are based on mean values of $n = 100$ to 200.

3.3. Comparison between CAD patients and healthy control subjects

Using our experimental data, similar transient sharp changes followed by converged stages were also identified in the GDR and GDE, as illustrated in Fig. 8. Both the GDR and GDE in the stable stages suggested a decrease in the CAD group compared to the healthy control group. Of the seven traditional Poincaré plot indices, only SD1/SD2 exhibited significant differences.

The corresponding p values, AUC levels and Cohen's d statics are summarized in Fig. 8, which also clearly suggests superior performance of our proposed indices, particularly the GDE, compared to the traditional metrics.

4. Discussion

Two novel measures, namely the GDR and GDE, were established in this study to characterize the shape of the Poincaré plot quantitatively. These were applied to both synthetic data with known properties and experimental ECG RR interval time-series, and their performance compared with seven existing Poincaré plot measures. The simulation tests demonstrated that the proposed GDR and GDE could effectively differentiate among the synthetic signals while none of the existing measures could, suggesting that the GDR and GDE may be able to capture the underlying dynamic or certain intrinsic geometric properties of a signal. These observations were also supported by the subsequent experimental tests. Specifically, the GDR and GDE could effectively discriminate the Poincaré plots of short-term RR interval time-series between healthy young and healthy aged subjects. In comparison, only two of the seven existing measures could achieve this, albeit with significantly inferior performance. Moreover, the GDR and GDE exhibited fairly strong performance in differentiating subjects with CAD from healthy controls, while only one of the seven existing methods could achieve this.

Our results indicated that the GDR and GDE of young subjects were greater than those of older subjects, suggesting a more dispersed distribution of the Poincaré plot for young people. Furthermore, the converged GDR and GDE values for healthy people appeared to be quite close to those for the synthetic chaotic signal, while the changing directions of the GDR and GDE in CAD patients were towards the level of less irregular/complex regimes (that is, periodic or the two studied mix processes) (see Figs. 4 and 8, and Table 2). Together, these results are in line with the studies of Lipsitz et al. [44] and Iyengar et al. [45], which demonstrated a reduction in the fractal organization of HRV owing to aging, as well as the study of Goldberger et al. [46], in which decreased complexity of the HRV was observed in a healthy aging group. Furthermore, the results suggested that the points in the Poincaré plot of CAD patients were clustered more intensively than those of the healthy controls (Fig. 8), which could indicate that CAD patients have reduced variance in the beat-to-beat heart rate [47,48], while also suggesting a certain shape distortion of the Poincaré plot, which could easily be overlooked by existing methods.

The proposed algorithms, particularly the GDE, can take full advantage of the Poincaré plot. In particular, they weight the grid areas according to the number of points, and further take all of these grids into consideration. By this means, the GDR and GDE are less likely to be affected or determined by a single outlier. The ellipse fit-based metrics SD1 and SD2 may offer a similar advantage. However, compared to the finely tuned grids and rates applied in the GDR and GDE, the ellipse fit may only offer a rather rough description of the shape, which is far from sensitive. However, towards another extreme, the four studied asymmetry metrics may be too sensitive to the positions of points, thereby being quite easily affected by the reallocation of even a single point. Our proposed GDR and GDE manage to composite the performance and are therefore neither too abstract nor too vulnerable.

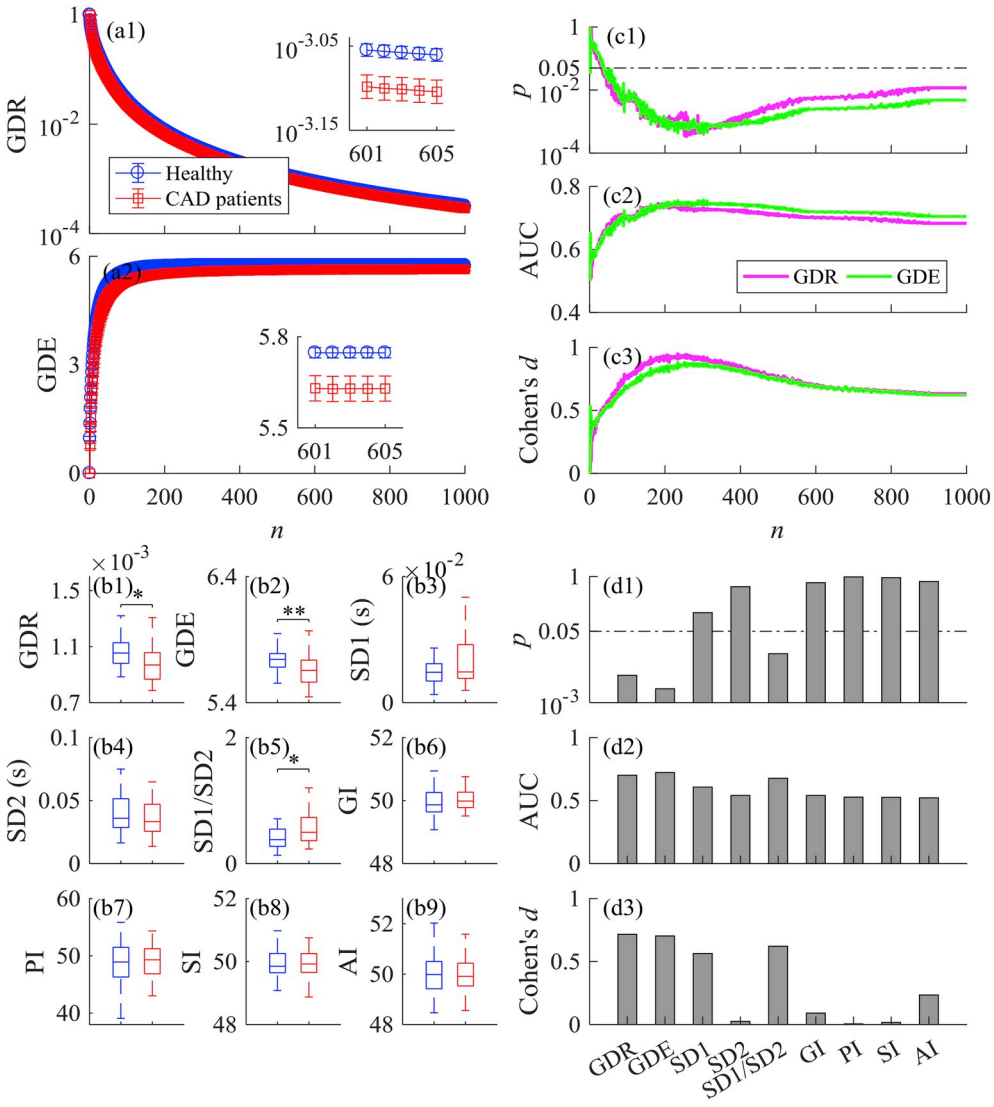

Fig. 8. Results of Poincaré plot metrics for clinical data: (a1) GDR results (base-10 logarithmic transformed) with n ranging from 1 to 1000; (a2) GDE results with the same range of n, where error bars indicate standard error; (b1 to b9) comparisons of GDR, GDE (average of n = 300 to 1000), SD1, SD2, SD1/SD2, GI, PI, SI, and AI, respectively, between healthy control and CAD groups; (c1 to c3) performance of GDR and GDE for comparing healthy control with CAD groups in terms of p values, AUC results, and effect size Cohen's d, respectively; (d1 to d3) performance of existing indices in terms p values, AUC results, and effect size Cohen's d, respectively. *: $p < 0.05$, **: $p < 0.01$.

Moreover, it is worth noting that another well-established graphical analytic tool, namely the recurrence plot or recurrence quantification analysis, offers similar metrics, such as entropy [49,50]. The recurrence plot is based on delay embedded reconstruction and aimed at indicating at which times the trajectory visits roughly the same areas in the phase space, by demonstrating how the distances between different vectors (of length m – the dimension parameter) are distributed (either within a threshold parameter r or not). Conceptually, the Poincaré plot is a type of one-dimensional recurrence plot, with the difference that it represents data in the original magnitude instead of illustrated the dichotomized distances between data points (as a real one-dimensional recurrence plot would do), which may offer additional information regarding the dynamics compared to dichotomized values [39,40]. Another obvious benefit of the GDR and GDE is that the construction of the gridded Poincaré plot requires only one input parameter, while the recurrence plot requires at least three (the embedding dimension m, time delay τ, and threshold r), leading to the potential disadvantage of parameter dependence, particularly for short-term data. To understand all of the pros and cons fully and objectively, a systemic comparison among these methods is warranted.

The coarse-graining parameter n is the only input parameter for the GDR and GDE. With increases in n, both measures exhibited an obvious and unique evolving profile, as explained by the initial transient sharp change (either an increase or a decrease), followed by a converged steady stage. The extensive steady stage suggests that the two features, namely GDR and GDE, are not trivial and should explain certain intrinsic properties. The steady stage in the GDR and GDE profiles should be the exact region on which we focus. Moreover, the wide range of n values with the steady results renders feasibility in the selection of this input parameter in practice. However, it should be noted that the length of the RR interval time-series may play a role, as it may interact with the parameter n to make the evolving profiles of the GDR and GDE more complicated. As mainly short-term, specifically 5 min, data were the focus of this study, we will attempt to investigate the influence of the data length and the appropriate selection of the parameter n in future works.

We note that, according to the flow of the algorithms illustrated in Fig. 1, the HRV time-series are normalized with a min-max normalization and hence, both the x-axis and y-axis range from 0 to 1. Outliers may thus potentially affect the calculations of the GDR and GDE. For example, extremities would certainly cause the actual RR intervals to be clustered more closely to one another in this normalized plane. In the current study, we applied the outlier removal prior to the normalization process, which possibly minimized this effect. A further possible limitation of this study is that, as we have applied open-source data, an effectively performing metric may simply be exhibited by statistical chance. However, the results when using our own experimental data suggest consistent performance of the proposed algorithms, which

essentially aided in alleviating this limitation.

In conclusion, we established two novel measures, namely the GDR and GDE, for quantitatively describing the shape of a Poincaré plot for short-term heartbeat interval time-series. We applied both simulation and experimental tests to validate these two algorithms. Our results clearly suggest that our proposed GDR and GDE outperform many existing Poincaré plot indices, including ellipse fit-based metrics (SD1, SD2, and SD1/SD2) and asymmetry indices (PI, GI, SI, and AI). Therefore, these two novel indices may be potentially applicable in the clinical and noninvasive evaluation of cardiovascular function.

Conflicts of interest

The authors declare no competing interests.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (Nos. 61471223, 61601263, and 61501280), and the Shandong Provincial Science and Technology Project (No. 2016GSF201037).

References

[1] Marek Malik, et al., Heart rate variability standards of measurement, physiological interpretation, and clinical use, Eur. Heart J. 17 (1996) 354–381.
[2] U. Rajendra Acharya, et al., Heart rate variability: a review, Med. Biol. Eng. Comput. 44 (12) (2006) 1031–1051.
[3] S. Akselrod, et al., Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control, Science 213 (4504) (1981) 220–222.
[4] P.K. Stein, et al., Heart rate variability: a measure of cardiac autonomic tone, Am. Heart J. 127 (5) (1994) 1376–1381.
[5] N. Karim, J.A. Hasan, S.S. Ali, Heart rate variability–a review, J. Basic Appl. Sci. 7 (1) (2011) 71–77.
[6] R.A. Hoshi, et al., Poincaré plot indexes of heart rate variability: relationships with other nonlinear variables, Auton. Neurosci. 177 (2) (2013) 271–274.
[7] M. Brennan, M. Palaniswami, P. Kamen, Poincaré plot interpretation using a physiological model of HRV based on a network of oscillators, Am. J. Physiol. Heart Circ. Physiol. 283 (5) (2002) H1873–H1886.
[8] P. Shi, S. Hu, H. Yu, Recovery of heart rate variability after treadmill exercise analyzed by lagged Poincaré plot and spectral characteristics, Med. Biol. Eng. Comput. 56 (2) (2018) 221–231.
[9] P. Kamen, A.M. Tonkin, Application of the Poincaré plot to heart rate variability: a new measure of functional status in heart failure, Aust. N. Z. J. Med. 25 (1) (1995) 18–26.
[10] P. Guzik, et al., Heart rate asymmetry by Poincaré plots of RR intervals, Biomed. Tech. 51 (4) (2006) 272–275.
[11] M.A. Woo, et al., Patterns of beat-to-beat heart rate variability in advanced heart failure, Am. Heart J. 123 (3) (1992) 704–710.
[12] M.P. Tulppo, et al., Quantitative beat-to-beat analysis of heart rate dynamics during exercise, Am. J. Physiol. 271 (1 Pt 2) (1996) H244–H252.
[13] M. Brennan, M. Palaniswami, P. Kamen, Do existing measures of Poincare plot geometry reflect nonlinear features of heart rate variability? IEEE Trans. Biomed. Eng. 48 (11) (2001) 1342–1347.
[14] A. Voss, et al., Short-term heart rate variability—influence of gender and age in healthy subjects, PLoS One 10 (3) (2015) e0118308.
[15] A. Voss, et al., Lagged segmented Poincaré plot analysis for risk stratification in patients with dilated cardiomyopathy, Med. Biol. Eng. Comput. 50 (7) (2012) 727–736.
[16] Z. Lijuan, et al., Automatic recognition of cardiac arrhythmias based on the geometric patterns of Poincaré plots, Physiol. Meas. 36 (2) (2015) 283.
[17] M. Costa, A.L. Goldberger, C.-K. Peng, Broken asymmetry of the human heartbeat: loss of time irreversibility in aging and disease, Phys. Rev. Lett. 95 (19) (2005) 198102.
[18] J. Piskorski, P. Guzik, Geometry of the Poincaré plot of RR intervals and its asymmetry in healthy adults, Physiol. Meas. 28 (3) (2007) 287.
[19] J. Piskorski, P. Guzik, Compensatory properties of heart rate asymmetry, J. Electrocardiol. 45 (3) (2012) 220–224.
[20] D.R. Chialvo, M.M. Millonas, Asymmetric unbiased fluctuations are sufficient for the operation of a correlation ratchet, Phys. Lett. 209 (1) (1995) 26–30.
[21] A. Porta, et al., Temporal asymmetries of short-term heart period variability are linked to autonomic regulation, Am. J. Physiol. Regul. Integr. Comp. Physiol. 295 (2) (2008) R550–R557.
[22] C.K. Karmakar, A.H. Khandoker, M. Palaniswami, Phase asymmetry of heart rate variability signal, Physiol. Meas. 36 (2) (2015) 303–314.
[23] C. Yan, et al., Area asymmetry of heart rate variability signal, Biomed. Eng. Online 16 (1) (2017) 112.
[24] P.K. Stein, et al., Traditional and nonlinear heart rate variability are each independently associated with mortality after myocardial infarction, J. Cardiovasc. Electrophysiol. 16 (1) (2005) 13–20.
[25] H.V. Huikuri, et al., Abnormalities in beat-to-beat dynamics of heart rate before the spontaneous onset of life-threatening ventricular tachyarrhythmias in patients with prior myocardial infarction, Circulation 93 (10) (1996) 1836–1844.
[26] T.H. Mäkikallio, et al., Dynamic analysis of heart rate may predict subsequent ventricular tachycardia after myocardial Infarctionfn1fn1This study was supported by grants from the Finnish foundation for cardiovascular research, Helsinki, Finland, national aeronautics and space administration, Washington D.C., and the G. Harold and leila Y. Mathers charitable foundation, Mt. Kisco, New York, Am. J. Cardiol. 80 (6) (1997) 779–783.
[27] P. Guzik, et al., Obstructive sleep apnea and heart rate asymmetry microstructure during sleep, Clin. Auton. Res. 23 (2) (2013) 91–100.
[28] P. Guzik, et al., Heart rate variability by Poincaré plot and spectral analysis in young healthy subjects and patients with type 1 diabetes, Folia Cardiol. 12 (suppl D) (2005) 64–67.
[29] M.A. Woo, et al., Complex heart rate variability and serum norepinephrine levels in patients with advanced heart failure, J. Am. Coll. Cardiol. 23 (3) (1994) 565–569.
[30] J. Park, S. Lee, M. Jeon, Atrial fibrillation detection by heart rate variability in Poincare plot, Biomed. Eng. Online 8 (1) (2009) 38.
[31] P.G. Jarosław Piskorski, Filtering Poincaré plots, Comput. Methods Sci. Technol. 11 (1) (2005) 39–48.
[32] S. Schulz, A. Voss, Symbolic dynamics, Poincaré plot analysis and compression entropy estimate complexity in biological time series, in: R. Barbieri, E.P. Scilingo, G. Valenza (Eds.), Complexity and Nonlinearity in Cardiovascular Signals, Springer International Publishing, Cham, 2017, pp. 45–85.
[33] S.M. Pincus, Approximate entropy as a measure of system complexity, Proc. Natl. Acad. Sci. Unit. States Am. 88 (6) (1991) 2297–2301.
[34] A.L. Goldberger, et al., Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) e215–e220.
[35] A. Voss, et al., Short-term heart rate variability—age dependence in healthy subjects, Physiol. Meas. 33 (8) (2012) 1289.
[36] P. Li, et al., Testing pattern synchronization in coupled systems through different entropy-based measures, Med. Biol. Eng. Comput. 51 (5) (2013) 581–591.
[37] P. Li, et al., Age related changes in variability of short-term heart rate and diastolic period, Computing in Cardiology Conference (CinC), 2013 IEEE, 2013.
[38] P. Li, et al., A low-complexity data-adaptive approach for premature ventricular contraction recognition, Signal Image Video Process. 8 (1) (2014) 111–120.
[39] P. Li, et al., Assessing the complexity of short-term heartbeat interval series by distribution entropy, Med. Biol. Eng. Comput. 53 (1) (2015) 77–87.
[40] P. Li, et al., Detection of coupling in short physiological series by a joint distribution entropy method, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 63 (11) (2016) 2231–2242.
[41] J.P. Royston, An extension of Shapiro and wilk's W test for normality to large samples, J. R. Stat. Soc. Ser. C Appl. Stat. 31 (2) (1982) 115–124.
[42] J. Cohen, Statistical Power Analysis for the Behavioral Sciences, Lawrence Earlbaum Associates, Hillsdale, NJ, 1988, pp. 20–26.
[43] S.S. Sawilowsky, New effect size rules of thumb, J. Mod. Appl. Stat. Methods 8 (2) (2009) 467–474.
[44] L.A. Lipsitz, A.L. Goldberger, Loss of complexity and aging: potential applications of fractals and chaos theory to senescence, J. Am. Med. Assoc. 267 (13) (1992) 1806–1809.
[45] N. Iyengar, et al., Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics, Am. J. Physiol. Regul. Integr. Comp. Physiol. 271 (4) (1996) R1078–R1084.
[46] A.L. Goldberger, C.K. Peng, L.A. Lipsitz, What is physiologic complexity and how does it change with aging and disease? Neurobiol. Aging 23 (1) (2002) 23–26.
[47] S. Hillebrand, et al., Heart rate variability and first cardiovascular event in populations without known cardiovascular disease: meta-analysis and dose–response meta-regression, EP Europace 15 (5) (2013) 742–749.
[48] J. Wang, X. Ning, Y. Chen, Multifractal analysis of electronic cardiogram taken from healthy and unhealthy adult subjects, Phys. Stat. Mech. Appl. 323 (2003) 561–568.
[49] J. Eckmann, S.O. Kamphorst, D. Ruelle, Recurrence plots of dynamical systems, World Sci. Ser. Nonlinear Sci. Ser. 16 (1995) 441–446.
[50] N. Marwan, et al., Recurrence plots for the analysis of complex systems, Phys. Rep. 438 (5) (2007) 237–329.