

# A New Entropy-Based Heart Failure Detector

Jinle Xiong, Xueyu Liang, Chengyu Liu

The State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering,  
Southeast University, Nanjing, 210096, China

## Abstract

Heart failure is a common type of clinical cardiovascular disease and has a high prevalence, disability and fatality rate. Entropy measures, typically as sample entropy (SampEn), has been used in clinic for detecting heart failure. However, SampEn values are sensitive to the selection of threshold  $r$ , resulting in a difficulty in the clinical interpretation. This study proposed a new entropy measure named sample difference entropy (SampDEn), for distinguishing congestive heart failure (CHF) patients from normal sinus rhythm (NSR) subjects. Unlike SampEn, the new SampDEn calculated the entropy value by comparing the information increase rate at two threshold settings of  $r_{max}$  and  $r_{min}$ , to reduce the statistical instability of SampEn due to the single threshold decision. The new algorithm was tested on the MIT-BIH RR Interval Databases. For 300 RR interval time series, the new proposed SampDEn reported an accuracy of 70.46% while SampEn reported 69.18%, 60.47%, 51.21% and 43.27% for  $r=0.1, 0.15, 0.2$  and  $0.25$  respectively. For 1,000 RR interval time series, SampDEn reported an accuracy of 75.33% while SampEn reported 68.34%, 56.09%, 44.85% and 36.76% for  $r=0.1, 0.15, 0.2$  and  $0.25$  respectively. The results suggested that the new SampDEn method is more effective for identifying CHF and NSR subjects than the traditional SampEn.

## 1. Introduction

Heart failure has been proved to increase muscular sympathetic activity, thus influences the non-linear heart rhythms in patients. Entropy methods, typically as approximate entropy (ApEn) [1] and sample entropy (SampEn), have been proven that can characterize the non-linear property of cardiovascular system, and thus can be used for heart failure assessment.

Based on ApEn proposed by Pincus, Richman and Moorman developed SampEn method to solve the shortcomings of bias and relative inconsistency in ApEn [2]. SampEn has been applied for distinguishing congestive heart failure (CHF) from normal sinus rhythm (NSR) [3,4]. Herein, three parameters need to be determined when using SampEn: embedding dimension  $m$ , tolerance threshold  $r$  and time series length  $N$ . Moreover, it is found that

SampEn values are quite sensitive to threshold  $r$  [5]. For clinical applications, recommended  $r$  for SampEn is usually between 0.10 and 0.25 times the standard deviation (SD) of RR interval time series [6]. However, when  $r$  increased from 0.10 to 0.25, SampEn values of NSR group were first higher than those of CHF group, but then became lower [7]. The inconsistency of SampEn values makes it hard to efficiently detect CHF patients from NSR subjects in clinical applications.

This study proposed a new entropy method named sample difference entropy (SampDEn) for distinguishing CHF and NSR subjects, aiming to enhance the detection accuracy between these two groups.

## 2. Methods

### 2.1 Sample entropy

SampEn was taken as baseline algorithms in this study. The calculation process of SampEn was summarized as follows [2,8]:

For RR segment  $x(i)$  ( $1 \leq i \leq N$ ), given the parameters  $m$  and  $r$ , the vector sequences  $X_i^m$  ( $1 \leq i \leq N - m$ ) can be formulated as:

$$X_i^m = \{x(i), x(i+1), \dots, x(i+m-1)\} \quad (1)$$

The vector  $X_i^m$  represents  $m$  consecutive  $x(i)$  values. Then the distance between  $X_i^m$  and  $X_j^m$  based on the maximum absolute difference is defined as:

$$d_{i,j}^m = d[X_i^m, X_j^m] = \max_{0 \leq k \leq m-1} |x(i+k) - x(j+k)| \quad (2)$$

For each  $X_i^m$ , we denote  $B_i^m(r)$  as  $(N-m)^{-1}$  times the number of  $X_j^m$  ( $1 \leq j \leq N-m$ ) that meets  $d_{i,j}^m \leq r$ . Similarly, we set  $A_i^m(r)$  as  $(N-m)^{-1}$  times the number of  $X_j^{m+1}$  that meets  $d_{i,j}^{m+1} \leq r$  for all  $1 \leq j \leq N-m$ . Then SampEn is defined by

$$\text{SampEn} = -\ln\left(\frac{\sum_{i=1}^{N-m} A_i^m(r)}{\sum_{i=1}^{N-m} B_i^m(r)}\right) \quad (3)$$

Herein, we pre-define two parameters in the calculation of entropy metrics: embedding dimension  $m=2$  and tolerance threshold  $r=0.10, 0.15, 0.20$  and  $0.25$  times the SD of RR interval time series. Since  $m$  is suggested to deal with the time series with a length of  $10^m$  to  $10^{m+1}$ , a

relatively large  $m$  may lead to inefficient entropy results, thus we use  $m=2$  in our study. Likewise, the values of  $r$  we chose is verified to provide stable outputs for certain RR interval time series. We also select the time series length  $N$  to be 300 and 1000 to check the influence of various-size RR interval segments [9].

## 2.2 Sample difference entropy

Basing on the shortcomings of the current SampEn, we proposed a new method named SampDEn. The calculation process for SampDEn are summarized below.

For RR segment  $x(i)$  ( $1 \leq i \leq N$ ), given the same parameter  $m$  and certain  $r$ , the vector sequences  $X_i^m$  ( $1 \leq i \leq N - m$ ) can be formulated as:

$$X_i^m = \{x(i), x(i+1), \dots, x(i+m-1)\} \quad (4)$$

The vector  $X_i^m$  represents  $m$  consecutive  $x(i)$  values. Then the distance between  $X_i^m$  and  $X_j^m$  based on the maximum absolute difference is defined as:

$$d_{i,j}^m = d[X_i^m, X_j^m] = \max_{0 \leq k \leq m-1} |x(i+k) - x(j+k)| \quad (5)$$

For each distance  $d_{i,j}^m$ , it is compared with two different thresholds, which are named as  $r_{max}$  and  $r_{min}$ . Therefore, two groups of outcomes will be obtained using different  $r$  values. We denote  $B_i^m(r_{max})$  as  $(N-m)^{-1}$  times the number of  $X_j^m$  ( $1 \leq j \leq N-m$ ) that meets  $d_{i,j}^m \leq r_{max}$ , and  $B_i^m(r_{min})$  as  $(N-m)^{-1}$  times the number of  $X_j^m$  ( $1 \leq j \leq N-m$ ) that meets  $d_{i,j}^m \leq r_{min}$ . Similarly, we set  $A_i^m(r_{max})$  as  $(N-m)^{-1}$  times the number of  $X_j^{m+1}$  that meets  $d_{i,j}^{m+1} \leq r_{max}$  for all  $1 \leq j \leq N-m$ , and  $A_i^m(r_{min})$  as  $(N-m)^{-1}$  times the number of  $X_j^{m+1}$  that meets  $d_{i,j}^{m+1} \leq r_{min}$  for all  $1 \leq j \leq N-m$ . Then SampDEn is defined by

$$\begin{aligned} SampDEn = & \ln\left(\sum_{i=1}^{N-m} A_i^m(r_{max}) / \sum_{i=1}^{N-m} B_i^m(r_{max})\right) \\ & - \ln\left(\sum_{i=1}^{N-m} A_i^m(r_{min}) / \sum_{i=1}^{N-m} B_i^m(r_{min})\right) \end{aligned} \quad (6)$$

Like SampEn, the entropy results of SampDEn is based on four parameters: embedding dimension  $m$ , RR segment length  $N$  and two tolerance thresholds  $r_{max}$  and  $r_{min}$ .

## 2.3 Experiment design

All data used in this study were from the RR Interval Databases from <http://www.physionet.org>, a free-access, on-line archive of physiological signals [10]. The NSR RR Interval Database was used as the non-pathological and control group data. This database included 54 long-term RR interval recordings of subjects in normal sinus rhythm

aged 29 to 76. The CHF RR Interval Database was used as the pathological group data. This database included 29 long-term RR interval recordings of subjects aged 34 to 79, with congestive heart failure (NYHA classes I, II, and III). Each of the long-term RR interval recordings is a 24-hour recording including both day-time and night-time. The original ECG signals were digitized at 128 Hz, and the beat annotations were obtained by automated analysis with a manual review and correction.

The procession of data consisted of three major steps. (1) Pre-processing and segmenting for each RR interval recording; (2) entropy calculation for each RR segment with different combinations of parameters; (3) comparison between NSR and CHF groups to determine whether SampDEn is better than SampEn.

In step (1), the RR intervals greater than 2 s were firstly removed from the raw RR interval recordings to ignore the influence from the artefacts. For each beat in the raw ECG signals, it was annotated as a normal (denoted as 'N') or abnormal heartbeat. The RR intervals formed from abnormal heartbeats were removed from the RR interval recordings in case of confounding the entropy analysis of HRV [11]. After that, we used two different length windows  $N$  to segment the long-term RR interval recordings to form the RR segments for the entropy calculation. In this study, we set  $N=300$  and  $N=1,000$  respectively to observe the performances of entropy measures for different length of RR segments. We did not consider the overlapping operation between adjacent  $N$ -length windows since the previous study reported that overlapping between adjacent  $N$ -length windows did not improve atrial fibrillation organization estimation with respect to the analysis of non-overlapping windows [12]. For each RR segment, we finally removed the RR intervals without 99% confidence interval (CI), (i.e.,  $\pm 3 \times SD$ ).

In step (2), SampEn and SampDEn were used to calculate the entropy values for each RR segment under the different parameter settings: embedding dimension  $m$  was set as 2, and tolerance threshold  $r$  was set as 0.10, 0.15, 0.20 and 0.25 respectively for SampEn. As for SampDEn method, previous study has found that the selection of  $r$  from 0.10 to 0.25 turned out to yield a relative fine classification results for the NSR and CHF groups, so we set 0.10 to be the minimum value and 0.25 to be the maximum value for SampDEn calculation here.

In step (3), the entropy results were compared between the NSR and CHF groups under the different combinations of parameters  $m$ ,  $r$  and  $N$ , aiming to explore whether SampDEn is superior to SampEn in distinguish the CHF patients from the NSR subjects.

## 2.4 Evaluation methods

First, the overall mean and SD values of SampEn and SampDEn were calculated across all RR interval recordings, separately for the NSR and CHF groups.

Student's  $t$ -test was used to test the statistical difference between the two groups. All statistical analyses were performed using the MATLAB software (Version R2017a, The MathWorks, Natick, USA). Statistical significance was reported with  $P < 0.05$ .

Then, classifier accuracy was calculated via the following performance metrics:

$$Acc = (TP + TN) / (TP + FN + TN) \quad (7)$$

where TP, TN, FP and FN are the numbers of true positives (CHF segment were corrected classified as CHF), true negatives (NSR segment were corrected classified as NSR), false positives (NSR segment were incorrected classified as CHF) and false negatives (CHF segment were incorrected classified as NSR). Cross validation method was used to calculate  $Acc$  via the equation above.

### 3. Results

Figure 1 gives an overview of SampEn and SampDn results for the two groups from the different combinations of  $(m, r, N)$ . All SampDn values in the CHF group were significantly lower than the NSR group ( $P < 0.01$ ), while the

SampEn values only had statistical significances for part of parameter combinations. Moreover,  $Acc$  values of SampEn ranged from 43.27% to 69.18% when  $N=300$ , which were lower than the 70.46% of SampDn. The difference in  $Acc$  results was more significant when  $N=1000$ , as the accuracy of SampDn was 75.33%. Meanwhile,  $Acc$  values of SampEn only ranged from 36.76% to 68.34%.

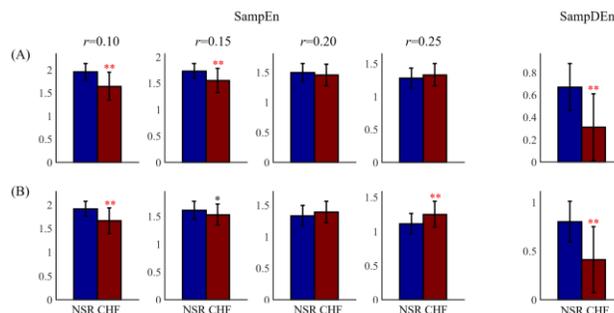


Figure 1. Results of SampEn and SampDn between NSR and CHF groups when (A)  $N=300$  and (B)  $N=1000$ . The symbol '\*' means statistical significance  $P < 0.05$  and '\*\*' means statistical significance  $P < 0.01$ .

Table 1. Results of SampEn and SampDn when applied on two length of RR time series:  $N=300$  and  $N=1,000$ . The embedding dimension  $m$  was set as 2.  $P$ -value measured the statistical significance between two groups. Data are expressed as number or mean  $\pm$  standard deviation (SD). '\*': statistical significance  $P < 0.05$ , '\*\*': statistical significance  $P < 0.01$ .

Length of RR time series	Method	Parameter $r$	NSR	CHF	$P$ -value	$Acc$
300	SampEn	0.10	1.8 $\pm$ 0.2	1.5 $\pm$ 0.3	7 $\times$ 10 <sup>-8</sup> **	69.18%
		0.15	1.6 $\pm$ 0.1	1.4 $\pm$ 0.3	5 $\times$ 10 <sup>-6</sup> **	60.47%
		0.20	1.4 $\pm$ 0.1	1.3 $\pm$ 0.2	0.10	51.21%
		0.25	1.2 $\pm$ 0.1	1.2 $\pm$ 0.2	0.35	43.27%
	SampDn	--	0.7 $\pm$ 0.2	0.3 $\pm$ 0.3	2 $\times$ 10 <sup>-8</sup> **	70.46%
1000	SampEn	0.10	1.8 $\pm$ 0.2	1.5 $\pm$ 0.3	3 $\times$ 10 <sup>-7</sup> **	68.34%
		0.15	1.5 $\pm$ 0.2	1.4 $\pm$ 0.2	0.013*	56.09%
		0.20	1.2 $\pm$ 0.2	1.3 $\pm$ 0.2	0.31	44.85%
		0.25	1.0 $\pm$ 0.1	1.1 $\pm$ 0.2	0.003**	36.76%
	SampDn	--	0.8 $\pm$ 0.1	0.4 $\pm$ 0.3	2 $\times$ 10 <sup>-10</sup> **	75.33%

### 4. Discussions

When using SampDn, the entropy values of CHF group were constantly lower than those of NSR group, thus it overcame the defect of SampEn. The better statistical significance and accuracy in discriminating CHF subjects from NSR subjects also confirms that SampDn is the improvement of SampEn.

The impact of threshold  $r$  on entropy values was studied before. Although the  $r$  used for ApEn is usually between 0.10 and 0.25 times the SD of time series, researches have revealed that this range might lead to incorrect conclusions when the dynamic performance of the time series becomes faster [13]. Since SampEn is a modification of ApEn, this shortcoming still remains. Other studies also pointed out that, higher SampEn values are not always associated to high complexity [14]. SampEn is prone to assign higher

entropy to the randomized surrogate time series as well as to certain pathological time series, which is a misleading observation [15]. This could be attributed to the fact that threshold  $r$  is based on long term SD of the original time series, hence unable to explore the real complexity by accounting the beat-to-beat variations inherited in a signal. In addition, the setting of constant  $r$  also results in the poor stability for analysis of physiological signals [16]. Therefore, the usage of  $r$  is controversial and SampEn might be unable to reflect the real complexity associated with a physiological time series.

The current study supports that the proposed SampDn is more adaptive to shorter time series than SampEn, when testing on the setting of  $N=300$  and  $N=1000$ . Regardless of its good performance on the tested data, there is still some uncertainty about SampDn. First, SampDn was tested at  $m=2$  merely. Since embedding dimension under 4 are commonly used, more  $m$  values need to be examined.

Second, as the length of time series usually varies largely, from dozens such as 75 points, to up to thousands of points, it is necessary to test SampEn and SampDn respectively with other  $N$  values [17,18]. Moreover, due to the special property of SampDn, any change in the two tolerance thresholds  $r_{max}$  and  $r_{min}$  might cause its advantages over SampEn to decrease. In this study, we set 0.10 to be the  $r_{min}$  and 0.25 to be the  $r_{max}$  for SampDn calculation according to expert experience. The large difference between these two thresholds contributed to superior entropy results for SampDn. Since there are plenty of other threshold values between 0.10 and 0.25, the variation of  $r_{max}$  and  $r_{min}$  might lead to different outcomes. Further work will therefore be focused on the various threshold values.

## 5. Conclusion

This study proposes a SampDn method to avoid the inconsistency of SampEn in previous applications for distinguishing the differences between the NSR and CHF groups more significantly. The better performance of SampDn indicates that it could discriminate CHF subjects more precisely and present better adaptability to relative short time series. Therefore, the proposed SampDn would be more advantageous in CHF detection.

## Acknowledgement

The project was partly supported by the National Natural Science Foundation of China (81871444 and 61571113), the Natural Science Foundation of Jiangsu Province (BE2017735). The authors thank the support from the Southeast-Lenovo Wearable Heart-Sleep-Emotion Intelligent monitoring Lab.

## References

- [1] Pincus S M. Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences of the United States of America*. 1991; 88: 2297-2301.
- [2] Richman J S, Moorman J R. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol*. 2000; 278: H2039-2049.
- [3] Ho K K, Moody G B, Peng C K, Mietus J E, Larson M G, Levy D, ., Goldberger A L. Predicting survival in heart failure case and control subjects by use of fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics. *Circulation*. 1997; 96: 842-848.
- [4] Liu C Y, Gao R. Multiscale entropy analysis of the differential RR interval time series signal and its application in detecting congestive heart failure. *Entropy*. 2017; 19: 251-.
- [5] Castiglioni P, Di Rienzo M. How the threshold “r” influences approximate entropy analysis of heart-rate variability. *Computing in Cardiology*. 2008; 35: 561-564.
- [6] Pincus S M. Assessing serial irregularity and its implications for health. *Ann NY Acad Sci*. 2001; 954: 245-267.
- [7] Zhao L N, Wei S S, Zhang C Q, Zhang Y T, Jiang X E, Liu F, Liu C Y. Determination of sample entropy and fuzzy measure entropy parameters for distinguishing congestive heart failure from normal sinus rhythm subjects. *Entropy*. 2015; 17: 6270-6288.
- [8] Zhang T, Yang Z, Coote J H. Cross-sample entropy statistic as a measure of complexity and regularity of renal sympathetic nerve activity in the rat. *Experimental Physiology* 2010; 92: 659-669.
- [9] Pincus S M, Huang W M. Approximate entropy: Statistical properties and applications. *Commun Stat Theory Methods*. 1992; 21: 3061-3077.
- [10] Goldberger A L, Peng C K, Lipsitz L A. What is physiologic complexity and how does it change with aging and disease? *Neurobiol Aging*. 2002; 23: 23-26.
- [11] Liu C Y, Li L P, Zhao L N, Zheng D C, Li P, Liu C C. A combination method of improved impulse rejection filter and template matching for identification of anomalous intervals in electrocardiographic RR sequences. *J Med Biol Eng*. 2012; 32: 245-250.
- [12] Alcaraz R, Abásolo D, Hornero R, Rieta J J. Optimal parameters study for sample entropy-based atrial fibrillation organization analysis. *Comput Methods Programs Biomed*. 2010; 99: 124-132.
- [13] Liu C Y, Liu C C, Shao P, Li L P, Sun X, Wang X P, Liu F. Comparison of different threshold values r for approximate entropy: application to investigate the heart rate variability between heart failure and healthy control groups. *Physiol Meas*. 2011; 32: 167-180.
- [14] Luiz E V S, Otavio Murta L. Evaluation of physiologic complexity in time series using generalized sample entropy and surrogate data analysis. *Chaos An Interdisciplinary Journal of Nonlinear Science* 2012; 22: 479-487.
- [15] Marwaha P, Sunkaria R K. Complexity quantification of cardiac variability time series using improved sample entropy (I-SampEn). *Australasian Physical & Engineering Sciences in Medicine*. 2016; 39: 755-763.
- [16] Liu C Y, Zhao L N. Using fuzzy measure entropy to improve the stability of traditional entropy measures. *Computing in Cardiology*. 2011; 38: 681-684.
- [17] Aktaruzaman M, Sassi R. Parametric estimation of sample entropy in heart rate variability analysis. *Biomed Signal Process Control*. 2014; 14: 141-147.
- [18] Mayer C C, Bachler M, Hörtenhuber M, Stocker C, Holzinger A, Wassertheurer S. Selection of entropy-measure parameters for knowledge discovery in heart rate variability data. *BMC Bioinformatics*. 2014; 15: S2.

Address for correspondence.  
 Chengyu Liu, PhD  
 School of Instrument Science and Engineering  
 Southeast University, Nanjing, China  
 E-mail: [chengyu@seu.edu.cn](mailto:chengyu@seu.edu.cn)