

# A Stochastic Nonlinear Autoregressive Algorithm Reflects Nonlinear Dynamics of Heart-Rate Fluctuations

ANTONIS A. ARMOUNDAS,<sup>1</sup> KIHWAN JU,<sup>2</sup> NIKHIL IYENGAR,<sup>1</sup> JORGEN K. KANTERS,<sup>3</sup> PHILIP J. SAUL,<sup>4</sup>  
RICHARD J. COHEN,<sup>1</sup> AND KI H. CHON<sup>2</sup>

<sup>1</sup>Division of Health Sciences and Technology, Harvard University–Massachusetts Institute of Technology, Cambridge, MA; <sup>2</sup>Dept. of Electrical Engineering, City College of the City University of New York, New York, NY; <sup>3</sup>Department of Medical Physiology, University of Copenhagen, Copenhagen, Denmark and Department of Internal Medicine, CCU, Elsinore Hospital, Elsinore, Denmark; and <sup>4</sup>Department of Pediatric Cardiology, Medical University of South Carolina, Charleston, SC

(Received 6 March 2001; accepted 8 November 2001)

**Abstract**—Current methods for detecting nonlinear determinism in a time series require long and stationary data records, as most of them assume that the observed dynamics arise only from the internal, deterministic workings of the system, and the stochastic portion of the signal (the noise component) is assumed to be negligible. To explicitly account for the stochastic portion of the data we recently developed a method based on a stochastic nonlinear autoregressive (SNAR) algorithm. The method iteratively estimates nonlinear autoregressive models for both the deterministic and stochastic portions of the signal. Subsequently, the Lyapunov exponents (LE) are calculated for the estimated models in order to examine if nonlinear determinism is present in the deterministic portion of the fitted model. To determine if nonlinear dynamic analysis of heart-rate fluctuations can be used to assess arrhythmia susceptibility by predicting the outcome of invasive cardiac electrophysiologic study (EPS), we applied the SNAR algorithm to noninvasively measured resting sinus-rhythm heart-rate signals obtained from 16 patients. Our analysis revealed that a positive LE was highly correlated to a patient with a positive outcome of EPS. We found that the statistical accuracy of the SNAR algorithm in predicting the outcome of EPS was 88% (sensitivity=100%, specificity=75%, positive predictive value=80%, negative predictive value=100%,  $p=0.0019$ ). Our results suggest that the SNAR algorithm may serve as a noninvasive probe for screening high-risk populations for malignant cardiac arrhythmias.  
© 2002 Biomedical Engineering Society.  
[DOI: 10.1114/1.1451074]

**Keywords**—Stochastic nonlinear autoregressive model, Lyapunov exponent, Nonlinear dynamics, Heart-rate variability.

## INTRODUCTION

From the time heart-rate variability (HRV) was first appreciated as a harbinger of sudden cardiac death in postmyocardial infarction patients by Wolf *et al.*,<sup>38</sup> many studies have established a significant relationship be-

tween decreased HRV and susceptibility to lethal ventricular arrhythmias<sup>17,29</sup> or sudden cardiac death.<sup>6,21</sup> A major issue has been how to describe HRV mathematically. The phenomenon of fluctuations in the interval between consecutive heart beats has been the subject of investigations using a wide range of methodologies including time-domain,<sup>2,12</sup> frequency-domain,<sup>7,18</sup> geometric,<sup>24</sup> and nonlinear<sup>19,20,28</sup> methods. Although these techniques have been shown to be potentially powerful tools for characterizing various complex physiological systems, none has been established as a reliable and accurate predictor of life threatening arrhythmias.

With the general recognition of nonlinear dynamics theory in the mid-1980s, it was proposed that HRV should be viewed as the result of nonlinear determinism in the regulatory systems governing the heart rate. Parameters indicative of possible low-dimensional nonlinear determinism include Lyapunov exponent (LE), strange attractors, and the correlation dimension.<sup>15,19,20</sup> For example, it has been suggested that the correlation dimension (CD) could be used to distinguish patients who develop ventricular fibrillation during the monitoring period from those who do not.<sup>30</sup> Obtaining a reliable CD is not an easy task, however. The data record length needed to identify a nonlinear system is estimated to increase approximately as 10 to the power of the CD of the system.<sup>32</sup> Thus, for most experimental data only low-dimensional nonlinear systems can be identified during a stationary segment of data.

Unlike the CD, LE values can be more reliably estimated using fewer data records.<sup>1</sup> For example, the present authors,<sup>9</sup> Aguirre and Billings,<sup>3</sup> and Barahona and Poon<sup>4</sup> have found that, using the nonlinear autoregressive (NAR) model, nonlinear dynamical systems can be accurately modeled even under conditions of high noise and for data record lengths as small as 1000 points in some cases. Clinical application of the NAR model has shown that nonlinear determinism is prevalent in

Address correspondence to Professor Ki Chon, City College of the City University of New York, Department of Electrical Engineering, Steinman T-677, 140th Street and Convent Avenue, New York, NY 10031. Electronic mail: chon@ccny.cuny.edu

healthy hearts, and that nonlinear determinism is decreased in patients with congestive heart failure.<sup>26</sup>

In this study, we present a new computational algorithm to quantify nonlinear heart-rate dynamics. The method relies on fitting a stochastic nonlinear autoregressive (SNAR) model to the heart-rate time series and estimating the Lyapunov exponent for the estimated models. We coined the term SNAR model to emphasize that the method models both the deterministic (system dynamics) and the stochastic (noise dynamics) components of the system.<sup>8,10</sup> The SNAR algorithm has been shown to be accurate for relatively short and noisy data records<sup>10,23</sup> and more accurate than the tradition model order search criteria such as the Akaike information criterion or minimum description length.<sup>8,22,23</sup> We have previously shown that the SNAR provides far superior parameter estimation than does the deterministic NAR model.<sup>8,10</sup>

The objective of this study is to demonstrate the applicability and reliability of the SNAR algorithm (through LE value estimation) in predicting the outcome of cardiac electrophysiologic study (EPS). Electrophysiologic study is a highly sensitive procedure widely used in clinical practice to determine the source of abnormal heart rhythms. Programed electrical stimulation, during EPS, is an invasive procedure considered to be the most reliable patient screening modality for sustained ventricular arrhythmias. However, the balance of test cost, risks, and effectiveness has led to an increasing demand for noninvasive screening modalities. Our results suggest that LE values estimated from a SNAR model, applied to noninvasive electrocardiographic recordings, may serve

as a reliable, noninvasive screening technique for malignant ventricular arrhythmias.

## METHODS

### SNAR Analysis

For the SNAR analysis, we utilize a previously developed algorithm that is accurate in determining both whether determinism is present in a time series heavily corrupted with stochastic noise, and whether this determinism has chaotic attributes, i.e., sensitivity to initial conditions.<sup>8,9</sup> The latter is indicated by a positive LE. The method is based on the fast orthogonal search (FOS) algorithm previously developed by Korenberg<sup>22,23</sup> and it relies on fitting a stochastic nonlinear autoregressive model to the time series followed by an estimation of the LE of the model over the observed distribution of states for the system. Estimation of the SNAR model is a two-step procedure in which deterministic terms in the nonlinear difference equation model are first identified, then stochastic NAR model parameters are estimated. In practice, linear and nonlinear model order selections are seldom known *a priori* and so it becomes necessary to consider how to determine the model structure. This is especially important in the case of nonlinear systems since the number of candidate model terms is very large. The SNAR algorithm, which exploits the orthogonal (i.e., uncorrelated) property, is designed to provide an optimum model term selection so as to determine which of the model terms are significant or not.

In the following section we briefly explain the SNAR algorithm. The third-order SNAR model process is defined as

$$\begin{aligned}
 y(n) = & \sum_{k=1}^K a(k)y(n-k) + \sum_{l_1=1}^{L_1} \sum_{l_2=1}^{L_2} b(l_1, l_2)y(n-l_1)y(n-l_2) + \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} \sum_{m_3=1}^{M_3} c(m_1, m_2, m_3) \\
 & \times y(n-m_1)y(n-m_2)y(n-m_3) + \sum_{p=1}^P d(p)e(n-p) + \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} f(q_1, q_2)e(n-q_1)e(n-q_2) \\
 & + \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \sum_{r_3=1}^{R_3} g(r_1, r_2, r_3)e(n-r_1)e(n-r_2)e(n-r_3) + \dots,
 \end{aligned} \tag{1}$$

where  $K$  is the maximum deterministic linear model order,  $(L_1, L_2)$  is the maximum second-order deterministic model order, and  $(M_1, M_2, M_3)$  is the maximum third-order deterministic model order. The term  $e(n)$  in Eq. (1) is considered a noise source or prediction error term where  $P$  is the maximum stochastic linear model order,  $Q$  is the maximum second-order stochastic model order, and  $R$  is the maximum third-order stochastic model or-

der. The parameters  $a(k)$ ,  $b(l_1, l_2)$ , and  $c(m_1, m_2, m_3)$  represent coefficients of the deterministic linear, second-order, and third-order autoregressive (AR) terms, respectively. The parameters  $d(p)$ ,  $f(q_1, q_2)$ , and  $g(r_1, r_2, r_3)$  represent coefficients of the stochastic linear, second-order, and third-order autoregressive terms, respectively.

The first step of the SNAR model involves estimating the deterministic linear and nonlinear AR terms accu-

rately without considering the noise source  $e(n)$ . The next step is then to estimate coefficients  $a(k)$ ,  $b(l_1, l_2)$ , and  $c(m_1, m_2, m_3)$  associated with the chosen significant model terms using the least-squares approach. Once proper linear and nonlinear deterministic AR model terms have been selected using our robust model order search criterion, the noise source or the prediction error terms are estimated by subtracting the deterministic portion of the dynamics from the experimentally obtained output response. The next step is to estimate only the significant parameters associated with the prediction error terms using the same robust model order search criterion as described above. The procedure of finding only the significant deterministic and stochastic terms is then repeated several times to increase the accuracy of the coefficient estimates, until there is little change in the model. For the present analysis, we have chosen the SNAR model order to be 20 linear, with 8 second-order, 5 third-order, and 3 error terms. From the search of 91 deterministic terms due to the selected model orders, only those significant terms are selected (normally, only 10 terms out of 91).

#### Lyapunov Exponent Estimation

Once a SNAR model is obtained, the LE of the deterministic component (includes those terms that do not include the prediction error terms,  $e_n$ ) are easily obtained by writing the deterministic part of the SNAR model as a  $j$ -dimensional system, where  $j$  is the value of the largest lag in the SNAR model. If the SNAR model is  $y_n = F(y_{n-1}, y_{n-2}, \dots, y_{n-j-1}, y_{n-j})$ , where  $F$  is some polynomial, we get

$$\begin{aligned} z_n^{(1)} &= z_{n-1}^{(2)}, \\ z_n^{(2)} &= z_{n-1}^{(3)}, \\ &\vdots \\ z_n^{(j-1)} &= z_{n-1}^{(j)}, \\ z_n^{(j)} &= F(z_{n-1}^{(j)}, z_{n-1}^{(j-1)}, \dots, z_{n-1}^{(2)}, z_{n-1}^{(1)}), \end{aligned} \quad (2)$$

from which the Jacobian  $[\mathbf{DF}(n)]$  is easily seen to be

$$\mathbf{DF}(n) = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ \frac{\partial F}{\partial z^{(1)}} & \frac{\partial F}{\partial z^{(2)}} & \frac{\partial F}{\partial z^{(3)}} & \frac{\partial F}{\partial z^{(4)}} & \dots & \frac{\partial F}{\partial z^{(j)}} \end{pmatrix}. \quad (3)$$

**TABLE 1. Patient characteristics. DC: dilated cardiomyopathy; IC: idiopathic cardiomyopathy; RVD: right ventricular dysplasia; RHD: rheumatic heart disease; VT: ventricular tachycardia; S: syncope; VF: ventricular fibrillation; NSVT: nonsustained ventricular tachycardia; EF: ejection fraction; MI: myocardial infarction; Y: yes; N: no; M: male; and F: female.**

Study	Age (yrs)/Sex	EF/MI	Heart disease	Presenting arrhythmia
1	72 M	15 N	CAD	VT
2	67 M	45 Y	CAD	VT
3	59 M	45 Y	CAD	S
4	74 M	65 Y	DC	S
5	69 M	34 Y	CAD	VT
6	74 M	30 Y	CAD	VF
7	31 M	40 N	RVD	VT
8	57 F	60 N	RHD	VT
9	79 M	42 N	IC	VT
10	61 M	18 N	CAD	VF
11	67 M	35 Y	CAD	NSVT
12	65 M	28 Y	CAD	VT
13	65 M	40 N	CAD	VT
14	57 M	35 Y	CAD	S
15	73 M	38 Y	CAD	VF
16	73 M	35 Y	CAD	VT

By evaluating the Jacobian over the time series, it is straightforward to calculate the LE of the system given in Eq. (2). In the present study we used an algorithm based on  $QR$  decomposition.<sup>11,14</sup> For further details concerning the algorithm as well as the calculation of the LE, the reader is referred to recently published papers on this subject.<sup>8,9,22,23</sup>

#### Clinical Study

We applied the SNAR algorithm to heart-rate time series data obtained from a patient population scheduled to undergo invasive cardiac EPS. The sign of the LE obtained using the SNAR algorithm was compared against the outcome of the EPS.

We studied 16 patients at the electrophysiology laboratory at New England Medical Center. All patients enrolled in the study gave informed consent for both the recording of the ECG data and the EPS. The patient clinical characteristics are presented in Table 1. Surface electrocardiographic data, in the form of three orthogonal lead projections (Frank orthogonal leads), were recorded for 4–5 min, while the patient's heart was in sinus rhythm. Electrocardiographic signals were preamplified (gain 1000 for sensitivity 1 cm/mV) and filtered (0.05–100 Hz) in the cardiograph. Analog data from each of three orthogonal ECG leads were digitized simultaneously (1000 Hz with 12 bit resolution) after additional amplification, and stored on the portable computer.

We first identified ECG complexes, then determined fiducial points by a template-matching-based QRS detection algorithm.<sup>31</sup> The annotation file was left not edited

and reflected locations of normal as well as abnormal QRS complexes (premature ventricular depolarizations, supraventricular beats, aberrantly conducted sinus beats). Once fiducial points for all complexes were identified, RR intervals for consecutive beats were calculated. Following RR interval calculation, the heart-rate signal was obtained using an efficient approach previously derived.<sup>5</sup>

#### *Electrophysiologic Study*

All patients referred for diagnostic EPS were eligible for this study. The indications for the selection of a patient for the study were broad and representative of a typical electrophysiology referral population. Programmed ventricular stimulation was performed with up to three ventricular extrastimuli, using a protocol previously described.<sup>25</sup> Patients were excluded if a permanent pacemaker had previously been implanted. Twelve patients had documented coronary disease, ten of whom had a history of myocardial infarction. Eight patients had ejection fraction lower than 40%. Five studies were conducted while the patients were on antiarrhythmic drugs.

The results of EPS were *a priori* defined by the most severe ventricular arrhythmia elicited during each study and were categorized as: either sustained ventricular tachycardia (VT), a monomorphic ventricular rhythm that persisted for at least 30 s or required active termination by either pacing or DC cardioversion due to hemodynamic collapse; or ventricular fibrillation (VF), a self-sustained polymorphic ventricular rhythm at a rate in excess of 280 bpm that required defibrillation for termination. Patients were considered to have an inducible

arrhythmia [positive (POS)] only if sustained VT or VF was elicited, otherwise they were classified as having no inducible arrhythmia [negative (NEG)].

#### *Statistics*

Chi-squared analysis was used to analyze the correlation between SNAR estimation and the outcome of EPS. All  $p$  values reflect single-sided tests;  $p$  values less than 0.05 were deemed significant. We used the Mann-Whitney test to compare the LE of patients with a negative or positive EPS.

## **RESULTS**

### *Simulation Results of Discrete and Continuous Systems Using the SNAR Model*

In this section, we illustrate the efficacy of the stochastic NAR model for detecting nonlinear determinism in short time series from discrete and continuous noise-driven nonlinear dynamic systems. Two such well-known systems, the Henon map and the Lorenz attractor, were simulated, and the performance of the stochastic NAR algorithm was compared to that of the deterministic NAR algorithm. Note that the stochastic NAR algorithm models both the system and the noise dynamics, whereas the deterministic NAR algorithm models only the system dynamics. Since the main focus of the present study is to identify deterministic chaos in noise-driven systems, noise was added to both simulations.

For the first example, we consider the Henon map with a Gaussian white-noise (GWN) source,  $e_n$ , described by the following difference equation:

$$y_n = \begin{cases} 3.168y_{n-1} + 0.3y_{n-2} - y_{n-1}^2 + ke_n, & \text{if } -0.1 < y_n < 5 \\ 0.1e_n, & \text{otherwise} \end{cases} \quad (4)$$

The second line of Eq. (4) was necessary to prevent the system from approaching infinity for large values of  $k$ . The value of  $k$  was adjusted to provide different levels of the signal-to-noise ratio (SNR), which was computed as the ratio of the variance of the signal to the variance of the noise.

To demonstrate the advantages of the stochastic NAR model algorithm in estimating the coefficients of the Henon map, we simulated Eq. (4) with 500 data points and a SNR of  $-9$  dB. The latter signifies that the variance of the noise is about 2.8 times greater than the variance of the signal, and in this case the noise is significant enough to completely obliterate the structure of the strange attractor for the Henon map. LE were

calculated using the estimated NAR models and Eqs. (2) and (3). The deterministic NAR model is the same algorithm as the stochastic NAR model, including automatic model order search, but it does not model the noise source. The true model order for the Henon map of Eq. (4) includes two linear lags, and one second-order nonlinear lag. The space of candidate terms from which the SNAR algorithm chose potential terms included five linear lags ( $y_{n-1}, \dots, y_{n-5}$ ), and three second-order lags ( $y_{n-1}y_{n-1}, \dots, y_{n-3}y_{n-3}, \dots, y_{n-2}y_{n-3}$ ), resulting in a total of 12 terms (including a constant). With the stochastic NAR model, we also selected a candidate space with three linear lags for the noise terms. For comparison, we also show the global LE (time going to infinity)

**TABLE 2. Lyapunov exponents for the Henon map with Gaussian white noise ( $N=500$  data points,  $SNR=-9$  dB).**

	Global LE for noiseless case	Local LE for noisy case	Deterministic	Stochastic
1	0.61	0.86	-0.05	0.46
2	-2.35	-2.60	-2.32	-2.74

of the corresponding noiseless Henon map ("global LE for noiseless case") together with the local LE calculated on the same 500 data points, but with no attempt to estimate their limit as time goes to infinity, using the true system equations ("local LE for noisy case"). As shown in Table 2, the deterministic NAR model incorrectly provided a negative LE whereas the stochastic NAR model correctly identified the underlying system as having a positive largest LE. Note that a positive Lyapunov exponent is an indication of a deterministic chaotic system. This example signifies that noise can cause problems in deciding whether a system is deterministically chaotic or not and that a sensitive method such as the SNAR algorithm is needed to correctly determine whether a given system is truly deterministic or not.

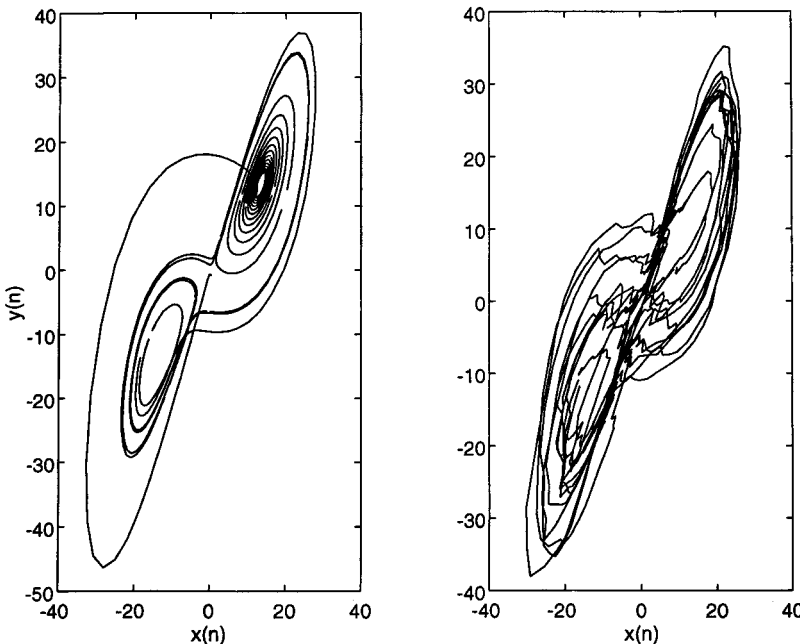
As an example of the continuous nonlinear dynamic system, the Lorenz equations were simulated as follows:<sup>37</sup>

$$\begin{aligned}
 dx/dt &= 16(y-x), \\
 dy/dt &= x(45.92-z)-y, \\
 dz/dt &= xy-4z.
 \end{aligned}
 \tag{5}$$

**TABLE 3. Lyapunov exponents for the Lorenz attractor with Gaussian white noise ( $N=1000$  data points,  $SNR=-5$  dB).**

	Global LE for noiseless case	Local LE for noisy case	Deterministic	Stochastic
1	2.16	1.30	0.12	0.83
2	0.00	-0.22	-5.11	-5.31
3	-32.40	-31.37	-69.42	-71.70
4			-69.52	-72.61
5			-69.54	-72.63
6			-69.55	

For both the deterministic and the stochastic NAR models, a search space of ten linear, five second-order, and four third-order terms was used. In addition, a search space of three error lag terms was used for the stochastic NAR model. Results based on 1000 data points are shown in Table 3. The data were sampled with a time step of 0.01, and the  $x$  variable was used for the analysis. Dynamic noise in the form of three independent realizations of GWN, was added to all three system variables in Eq. (5) to provide a SNR of  $-5$  dB. Notice that independent dynamic noise added to all three variables effectively corresponds to a nonlinear noise source, despite the fact that the individual terms are i.i.d. Gaussian random variables. Figure 1 shows the Lorenz attractor for 1000 data points without (left panel) and with (right panel) the  $-5$  dB noise. For the Lorenz system both the deterministic and the stochastic NAR models had a single positive LE regardless of whether the models were estimated on 1000 data points (Table 3). However, the stochastic NAR model provided an estimate for the larg-



**FIGURE 1. Lorenz attractor with (right panel) and without (left panel)  $-5$  dB noise ( $N=1000$  data points).**



**TABLE 4. RR interval statistics for the patient subject group (NTOB: number of thrown out beats, RMSSD RR: root-mean sum of squares of differences between adjacent RR intervals; RMSSD NN: root-mean sum of squares of differences between adjacent normal-to-normal intervals; VT: ventricular tachycardia; VF: ventricular fibrillation; NSVT: nonsustained ventricular tachycardia; NEG: negative; POS: positive, NA: not applicable; and S: significance).**

Study	NTOB	Mean $\pm$ SD RR (ms)	RMSSD RR (ms)	RMSSD NN (ms)	Induced arrhythmia	Maximum Lyapunov exponent (MLE)	Surrogate data MLE Mean $\pm$ SD, S (significance)
1	74 (18%)	643 $\pm$ 120	174.4	85.88	VT (POS)	POS (0.474)	-0.181 $\pm$ 0.162, (4.03)
2	7 (2%)	937 $\pm$ 46	76.5	40.21	VT (POS)	POS (0.561)	-0.435 $\pm$ 0.168 (5.93)
3	36 (7%)	727 $\pm$ 48	77.1	10.87	NSVT (NEG)	NEG (-0.330)	
4	17 (5%)	908 $\pm$ 77	90.3	24.99	NSVT (NEG)	POS (0.773)	-0.314 $\pm$ 0.332 (11.64)
5	290 (65%)	755 $\pm$ 166	286.2	16.11	VT (POS)	POS (0.435)	-0.364 $\pm$ 0.332 (2.40)
6	22 (4%)	864 $\pm$ 85	139.1	20.12	VT (POS)	POS (0.464)	-0.434 $\pm$ 0.254 (3.53)
7	11 (4%)	1085 $\pm$ 120	131.6	114.5	NEG	NEG (-0.168)	
8	9 (2%)	786 $\pm$ 43	51.8	17.94	NEG	POS (0.760)	-0.748 $\pm$ 0.719 (2.11)
9	2 (1%)	914 $\pm$ 41	47.9	30.93	NSVT (NEG)	NEG (-0.636)	
10	25 (5%)	638 $\pm$ 56	74.9	22.28	NEG	NEG (-0.238)	
11	153 (46%)	745 $\pm$ 196	346.0	80.63	VT (POS)	POS (0.125)	-0.141 $\pm$ 0.111 (2.39)
12	34 (6%)	695 $\pm$ 60	100.2	4.90	NSVT (NEG)	NEG (-0.382)	
13	108 (26%)	1179 $\pm$ 215	240.2	78.29	NEG	NEG (-0.054)	
14	2 (0%)	720 $\pm$ 15	22.5	9.84	VT (POS)	POS (1.231)	-0.579 $\pm$ 0.218 (8.28)
15	21 (5%)	869 $\pm$ 99	156.1	12.78	VF (POS)	POS (0.850)	-0.153 $\pm$ 0.137 (7.32)
16	2 (1%)	1487 $\pm$ 37	31.6	31.70	VT (POS)	POS (0.583)	-0.461 $\pm$ 0.211 (4.96)

est LE that was closer to the true value than did the deterministic model. Clearly, both algorithms overestimated the total model order as indicated by the large number of spurious negative LE exponents. The simulations were repeated ten times with independent realizations of the noise process. In all cases both the deterministic and the stochastic algorithm gave results similar to those in Table 3. Similar to the discrete case, because SNAR exclusively models the noise source, it provides better characterization of the deterministic system.

#### Clinical Study

We studied 16 patients [15 male and one female, from 31 to 79 years old (the average was 65.2 $\pm$ 11.2)]. In Table 4, the mean and standard deviation of the durations of the RR intervals, the square root of the mean of the sum of the squares of differences (RMSSD) between adjacent RR intervals (RMSSD RR), as well as the square root of the mean of the sum of the squares of differences between adjacent normal-to-normal intervals (RMSSD NN) of the 16 patients is presented.

#### Application of Stochastic Nonlinear Autoregressive Analysis to Predicting the Outcome of EPS

To investigate the ability of the SNAR algorithm to provide accurate detection of nonlinear determinism in heart-rate fluctuations, we applied the SNAR algorithm to heart-rate time series data obtained from a patient population scheduled to undergo invasive cardiac EPS. The subject-by-subject LE results are presented in Table

4. A positive or negative Lyapunov exponent is our prediction of correspondence to a patient positive or negative for EPS, respectively. A positive LE obtained from the SNAR identified the inducible patient population with sensitivity 100%, specificity 75%, positive predictive value 80%, negative predictive value 100%, and overall accuracy (defined as true positive and true negative over all studies) 88% ( $P < 0.002$ ). The SNAR was able to correctly predict the outcome of EPS in three cases with base-line prolongation of the unfiltered QRS complex due to bundle branch block or other intraventricular conduction. In addition, the Lyapunov exponent correctly predicted the EPS outcome in one patient who had high ventricular ectopy ( $> 10\%$ ).

In Fig. 2 we show representative Poincaré plots of the three groups of patients. The left, middle, and right panels represent VT, nonsustained VT, and noninducible arrhythmia, respectively. A study by Garfinkel *et al.* has shown that such Poincaré plots may be used to distinguish dynamics between normal and diseased conditions.<sup>13</sup> It is not clear whether differences in dynamics between groups of subjects can be found in our data. Poincaré plots require subjective interpretation rather than quantitative analysis.

To compare the Lyapunov exponents in patients who had a positive or negative EPS, we used the Mann-Whitney U test and found that the amplitude of the LE in patients with positive EPS was statistically significantly higher compared to those patients with negative EPS ( $p = 0.0063$ ).

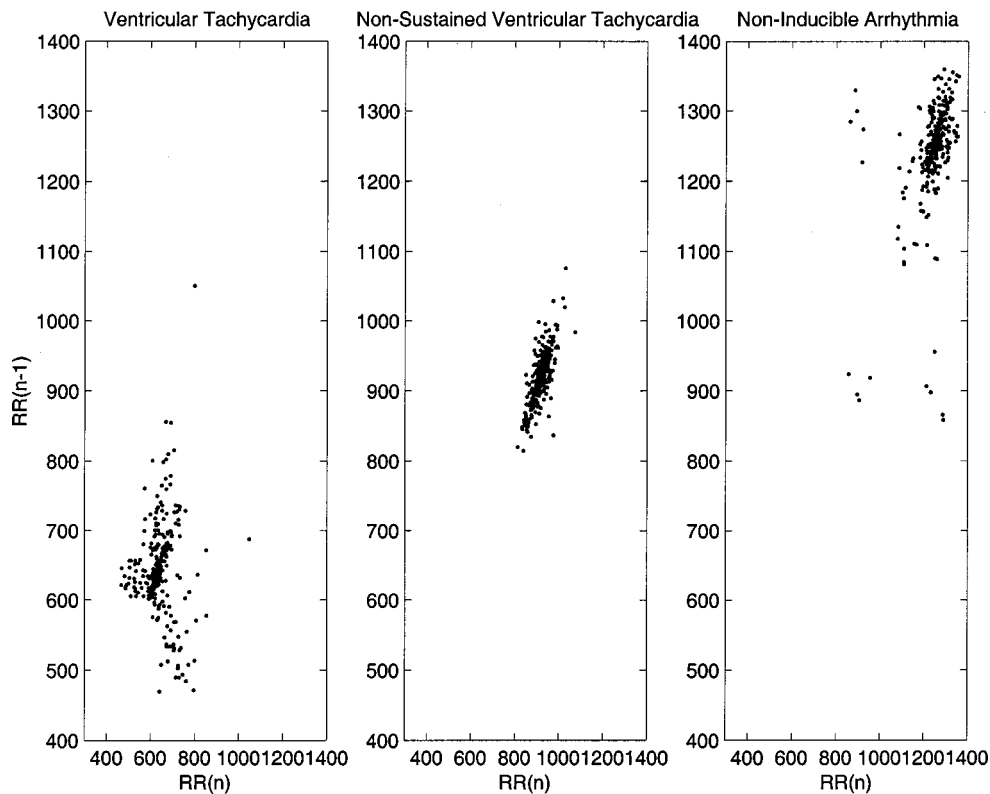


FIGURE 2. Poincaré plots of consecutive RR interval time series.

*Application of Stochastic Nonlinear Autoregressive Analysis in Predicting the Outcome of EPS in Normal-to-Normal Beat Sequences*

We also analyzed the patient data in the case that only normal-to-normal beat sequences were included. A beat was classified as good when both of the following criteria were satisfied: (i) the morphology criterion, which required the correlation coefficient between the current beat and the average beat to be greater than 0.95, and (ii) the RR criterion, which required that the current RR interval be less than 25% premature or delayed from the mean RR interval of the previous five beats. If the morphology criterion was not satisfied for a beat, then that and the next beat (and their corresponding RR intervals) were classified as bad ones. If the RR criterion was not satisfied for a beat, then only that (and its corresponding RR interval) was classified as a bad beat. After all bad beats were discarded, the good beats were shifted forward so as to form a normal-to-normal sequence of beats. In Table 4, we present the number of thrown-out beats (NTOB) and the percent of the NTOBs out of the total number of beats in each data file.

Following the above methodology, the analysis results of the SNAR algorithm predicted the outcome of EPS with sensitivity 100%, specificity 40%, positive predic-

tive value 62%, negative predictive value 100%, and overall accuracy 69% ( $p=0.05$ ).

*Confirmation of Determinism of Time Series Using Surrogate Data*

In order to ensure that the SNAR algorithm is not itself introducing deterministic nonlinearities, we performed a test on only the time series which had resulted in positive Lyapunov exponents. The time series were subjected to the surrogate data technique. The goal of the surrogate data transformation is to destroy the nonlinear dynamics in the time series. This leaves a time series with only linear properties. If the SNAR proceeds to find coefficients which lead to positive LE, then we know that SNAR introduces spurious nonlinear dynamics.

We employed the amplitude-adjusted Fourier transform algorithm to generate surrogate data from the original time series.<sup>34</sup> The surrogate data are created by rescaling the original time series followed by the Fourier transform of the rescaled time series. This manipulation leads to the surrogate data having the same Fourier spectrum. Finally, the surrogate time series is rescaled once more by applying the inverse map, so that it has the same amplitude distribution as the original time series.

To test the statistical significance of the surrogate data, we computed a dimensionless quantity,  $S$ , defined as

$$S = \frac{|x_d - \overline{x_{\text{surr}}}|}{\sigma_{\text{surr}}}, \quad (6)$$

where  $x_d$  is the LE value for the real data, and  $\overline{x_{\text{surr}}}$  and  $\sigma_{\text{surr}}$  denote the mean value and the standard deviation of the surrogates, respectively. A value of  $S$  greater than 2 implies that the null hypothesis is rejected, meaning that true nonlinear dynamics exist in the experimental data.

For each of the positive LE heart-rate time series (including both normal and non-normal beats), we generated ten surrogate data sets and applied the SNAR algorithm to estimate the corresponding LE of each. The mean and standard deviation values of the LE are provided in Table 4. We consistently found negative LE for all of the surrogate data sets, meaning that the positive LE found in the original data sets (Table 4) were due to inherently nonlinear dynamics governing the system. This is confirmed by the fact that for those with a positive LE, the significance factor  $S$ , as determined using Eq. (6), is greater than the value of two (Table 4). Therefore, the null hypothesis of the computed measure rejects linear correlation.

## DISCUSSION

In this study, we evaluated the applicability and reliability of the SNAR algorithm, as a potential noninvasive technique for screening patients for malignant ventricular arrhythmias. The SNAR algorithm employs nonlinear principles to capture heart-rate dynamics. The method presented in this article overcomes some of the deficiencies associated with the traditional methods for detecting nonlinear determinism. The stochastic NAR approach allows an efficient estimation of the NAR model coefficients as well as the error terms, and makes it possible to decouple system parameters from noise model parameters. This enables the approach to be accurate for detecting nonlinear determinism for data lengths as short as 500 data points, and for a SNR as low as  $-9$  dB, in certain circumstances. With the shorter data length requirement, stationarity requirements are more easily satisfied. Importantly, in some cases when either the data length is short or the SNR is low, the deterministic NAR model is unable to correctly identify sensitivity to initial conditions in a time series, whereas the stochastic NAR model can. The method presented is computationally fast, thus it has the potential for on-line application in clinical settings. Despite short RR interval series, the SNAR predicted the outcome of EPS with 88% accuracy.

It has been suggested that variability in the heart rate is due to deterministic chaos in the physiological control systems that govern the function of the heart.<sup>26,33</sup> However, a recent analysis based on calculation of both the correlation dimension, and the nonlinear prediction error, on experimental and surrogate data, failed to provide evidence for the presence of low-dimensional chaos in the heart.<sup>20</sup> The normal heart beat is initiated by a group of pacemaker cells in the sinus node. In isolation, the sinus node appears to be relatively stable, producing heart beats at a regular rate. This is evident in heart transplant patients where the nervous innervation to the heart is missing. In these patients the instantaneous heart rate shows much less variation than in normal subjects.<sup>27</sup> This implies that an important source of the normal variability is due to the external forcing of the sinus node by the nervous system. Thus, the normal heart-rate variability could be an example of a system where the fluctuations arise as a consequence of a nonlinear deterministic system (the sinus node) being forced by a high-dimensional input (the activity in the nerves innervating the sinus node). It has been suggested that the variability seen in the heart rate is not due to the action of the deterministic component of the system, but rather should be attributed to a high-dimensional input to the system that controls the heart.<sup>9</sup> The activity in autonomic nerves is not only determined by simple feedback from the baroreceptors (pressure sensors) in the cardiovascular system, but is also influenced by inputs from many other systems, including hormonal systems, and higher centers such as the cerebral cortex. Thus, transition from the higher-dimensional nonlinear behavior during normal physiological functioning to lower-dimensional nonlinear determinism seen with patients who recorded positive EPS is consistent with the fact that heart-rate variability is reduced in patients with cardiac disease.

The negative LE obtained for those patients with the negative EPS most likely reflects a high-dimensional system. To reliably estimate a high-dimensional system would require many data points, thus, a negative LE may have been obtained due to having only limited heart-rate data (note that even if we had had more data points, the stationarity assumption would be violated and the results would be suspect).

In this study, the SNAR algorithm applied to short time series of non-normal-to-normal beats, irrespective of the presence of bundle branch block or ventricular ectopy, was a more accurate predictor of the outcome of EPS than was application to the normal-to-normal beats, a result that is consistent with the results of Huikuri *et al.*<sup>16</sup> who recently demonstrated that short-term fractal scaling exponents analyzed without exclusion of the premature beats performed even better as a predictor of all-cause mortality rate than that after editing for premature beats. Furthermore, the SNAR algorithm correctly



predicted the outcome of EPS in three cases with bundle branch block and in one case with a high incidence of ectopic activity (>10%). This is an advantage of the SNAR algorithm compared to other proposed noninvasive screening techniques like the time-domain signal-averaged electrocardiogram and QT dispersion that cannot be applied to patients with intraventricular conduction abnormalities such as bundle branch block.<sup>35,36</sup>

A limitation of this pilot study was the small number of patients enrolled. Further investigations of the predictive value of the SNAR algorithm needs to be conducted in larger and more homogeneous patient population groups. In addition, it will be important in future studies to compare SNAR with other noninvasive measures of arrhythmic vulnerability as well as evaluate the ability of SNAR to predict arrhythmia-free survival.

In summary, in this patient population, SNAR appears to be a promising noninvasive technique, with the potential to be used for screening populations susceptible to ventricular arrhythmias.

#### ACKNOWLEDGMENT

This work was supported by the Whitaker Foundation.

#### REFERENCES

- <sup>1</sup>Abarbanel, H. D. I. *Analysis of Observed Data*. New York: Springer, 1996.
- <sup>2</sup>Adamson, P. B., M. H. Huang, E. Vanoli, R. D. Foreman, P. J. Schwartz, and S. S. Hull. Unexpected interaction between adrenergic blockade and heart-rate variability before and after myocardial infarction. A longitudinal study in dogs at high and low risk for sudden death. *Circulation* 90:976–982, 1994.
- <sup>3</sup>Aguirre, L. A., and S. A. Billings. Identification of models for chaotic systems from noisy data: Implications for performance and nonlinear filtering. *Physica D* 85:239–258, 1995.
- <sup>4</sup>Barahona, M., and C. S. Poon. Detection of nonlinear dynamics in short, noisy time series. *Nature (London)* 381:215–217, 1996.
- <sup>5</sup>Berger, R. D., S. Akselrod, D. Gordon, and R. J. Cohen. An efficient algorithm for spectral analysis of heart-rate variability. *IEEE Trans. Biomed. Eng.* 33:900–904, 1986.
- <sup>6</sup>Bigger, J. T., J. L. Fleiss, R. C. Steinman, L. M. Rolnitzky, R. E. Kleiger, and J. N. Rottman. Frequency-domain measures of heart period variability after myocardial infarction. *Circulation* 85:164–171, 1992.
- <sup>7</sup>Bigger, Jr., J. T., L. R. Rolnitzky, R. C. Steinman, and J. L. Fleiss. Predicting mortality after myocardial infarction from the response of RR variability to antiarrhythmic drug therapy. *J. Am. Coll. Cardiol.* 23:733–740, 1994.
- <sup>8</sup>Chon, K. H., K. P. Yip, B. M. Camino, D. J. Marsh, and N. H. Holstein-Rathlou. Modeling nonlinear determinism in short time series from noise driven discrete and continuous systems. *Int. J. Bifurcation Chaos Appl. Sci. Eng.* 10:2745–2766, 2000.
- <sup>9</sup>Chon, K. H., J. K. Kanters, R. J. Cohen, and N. H. Holstein-Rathlou. Detection of chaotic determinism in time series from randomly forced maps. *Physica D* 99:471–486, 1997.
- <sup>10</sup>Chon, K. H., M. J. Korenberg, and N. H. Holstein-Rathlou. Application of fast orthogonal search to linear and nonlinear stochastic systems. *Ann. Biomed. Eng.* 25:793–801, 1994.
- <sup>11</sup>Eckman, J. P., S. O. Kamphorst, D. Ruelle, and S. Ciliberto. Lyapunov exponents from time series. *Phys. Rev. A* 34:4971–4979, 1985.
- <sup>12</sup>Faber, T. S., A. Staunton, K. Hvatkova, A. J. Camm, and M. Malik. Stepwise strategy of using short- and long-term heart-rate variability for risk stratification after myocardial infarction. *PACE* 19:1845–1851, 1996.
- <sup>13</sup>Garfinkel, A., P. S. Chen, D. O. Walter, H. S. Karagueuzian, B. Kogan, S. J. Evans, M. Karpoukhin, C. Hwang, T. Uchida, M. Gotoh, O. Nwasokwa, P. Sager, and J. N. Weiss. Quasiperiodicity and chaos in cardiac fibrillation. *J. Clin. Invest.* 99:156–157, 1997.
- <sup>14</sup>Geist, K., U. Parlitz, and W. Lauterborn. Comparison of different methods for computing Lyapunov exponents. *Prog. Theor. Phys.* 83:875–893, 1990.
- <sup>15</sup>Grassberger, P., and I. Procaccia. Measuring the strangeness of strange attractors. *Physica D* 9:183–208, 1983.
- <sup>16</sup>Huikuri, N. V., T. H. Mäkikallio, C.-K. Peng, A. L. Goldberger, U. Hintze, and M. Møller. For the DIAMOND study group. Fractal correlation properties of RR interval dynamics and mortality in patients with depressed left ventricular function after an acute myocardial infarction. *Circulation* 101:47–53, 2000.
- <sup>17</sup>Huikuri, H. V., J. O. Valkama, K. E. J. Airaksinen, T. Seppänen, K. M. Kessler, J. T. Takkenen, and R. J. Myerburg. Frequency domain measures of heart-rate variability before the onset of nonsustained and sustained ventricular tachycardia in patients with coronary artery disease. *Circulation* 87:1220–1228, 1993.
- <sup>18</sup>Huikuri, H. V., K. E. J. Mäkikallio, J. Airaksinen, T. Seppänen, P. Puukka, I. J. Räihä, and L. B. Sourander. Power-law relationship of heart-rate variability as a predictor of mortality in the elderly. *Circulation* 97:2031–2036, 1998.
- <sup>19</sup>Kanters, J. K., M. V. Høgaard, E. Agner, and N. H. Holstein-Rathlou. Short- and long-term variations in nonlinear dynamics of heart-rate variability. *Cardiovasc. Res.* 31:400–409, 1996.
- <sup>20</sup>Kanters, J. K., N. H. Holstein-Rathlou, and E. Agner. Lack of evidence for low-dimensional chaos in heart-rate variability. *J. Cardiovasc. Electrophysiol.* 5:591–601, 1994.
- <sup>21</sup>Kleiger, R. E., J. P. Miller, J. T. Bigger, and A. J. Moss, and the Multicenter Post-Infarction Research Group. Decreased heart-rate variability and its association with increased mortality after acute myocardial infarction. *Am. J. Cardiol.* 59:256–282, 1987.
- <sup>22</sup>Korenberg, M. J. A robust orthogonal algorithm for system identification and time series analysis. *Biol. Cybern.* 60:267–276, 1989.
- <sup>23</sup>Korenberg, M. J., S. A. Billings, Y. P. Li, and P. J. McIlroy. Orthogonal parameter estimation algorithm for nonlinear stochastic systems. *Int. J. Control* 48:193–210, 1988.
- <sup>24</sup>Malik, M., R. Xia, O. Odemuyiwa, A. Staunton, J. Poloniecki, and A. J. Camm. Influence of the recognition artifact in the autonomic analysis of long-term electrocardiograms on time-domain measurement of heart-rate variability. *Med. Biol. Eng. Comput.* 31:539–544, 1993.
- <sup>25</sup>Manolis, A. S., and N. A. M. Estes III. Value of programmed ventricular stimulation and management of patients with nonsustained ventricular tachycardia associated with coronary artery disease. *Am. J. Cardiol.* 65:201–205, 1990.
- <sup>26</sup>Poon, C. S., and C. K. Merrill. Decrease of cardiac chaos in

- gestive heart failure. *Nature (London)* 389:492–495, 1997.
- <sup>27</sup>Sands, K. E., M. L. Appel, L. S. Lilly, F. J. Schoen, G. H. J. Mudge, and R. J. Cohen. Power spectrum analysis of heart-rate variability in human cardiac transplant recipients. *Circulation* 79:76–82, 1989.
- <sup>28</sup>Schmidt, G., and G.E. Monfill. Nonlinear methods for heart-rate variability assessment, edited by M. Malik, A. J. Camm. *Heart-rate Variability*, Armonk, NY: Futura, 1995, pp. 87–98.
- <sup>29</sup>Shusterman, V., B. Aysin, V. Gottipaty, R. Weiss, S. Brode, D. Schwartzman, and K. P. Anderson. Autonomic nervous system activity and the spontaneous initiation of ventricular tachycardia. ESVEM Investigators. Electrophysiologic study versus electrocardiographic monitoring trial. *J. Am. Coll. Cardiol.* 32:1891–1899, 1998.
- <sup>30</sup>Skinner, J. E., C. M. Pratt, and T. Vybiral. A reduction in the correlation dimension of heart beat intervals precedes imminent ventricular fibrillation in human subjects. *Am. Heart J.* 125:731–743, 1993.
- <sup>31</sup>Smith, J. M., E. A. Clancy, C. R. Valeri, J. N. Ruskin, and R. J. Cohen. Electrical alternans and cardiac electrical instability. *Circulation* 77:110–121, 1988.
- <sup>32</sup>Smith, L. A. Intrinsic limits on dimension calculations. *Phys. Lett. A* 133:283–288, 1988.
- <sup>33</sup>Sugihara, G., Ó Allan, D. Sobel, and K. D. Allan. Nonlinear control of heart-rate variability in human infants. *Proc. Natl. Acad. Sci. U.S.A.* 93:2608–2613, 1996.
- <sup>34</sup>Theiler, J., S. Eubank, A. Lontin, B. Galdrikian, and J. D. Farmer. Testing for nonlinearity in time series: The method of surrogate data. *Physica D* 58:77–94, 1992.
- <sup>35</sup>Turitto, G., E. Caref, G. Macina, J. M. Fontaine, S. N. Ursell, and N. El-Sherif. Time course of ventricular arrhythmias and the signal averaged electrocardiogram in the postinfarction period: A prospective study of correlation. *Br. Heart J.* 60:17–22, 1988.
- <sup>36</sup>Yunus, V., A. M. Gillis, H. J. Duff, D. G. Wyse, and L. B. Mitchell. Increased precordial QT<sub>c</sub> dispersion predicts ventricular fibrillation during acute myocardial infarction. *Am. J. Cardiol.* 78:706–708, 1996.
- <sup>37</sup>Wolf, A., J. B. Swift, H. L. Swinney, and J. A. Vastano. Determining Lyapunov exponents from a time series. *Physica D* 16:285–317, 1985.
- <sup>38</sup>Wolf, M. M., G. A. Vargios, D. Hunt, and J. G. Sloman. Sinus arrhythmia in acute myocardial infarction. *Med. J. Aust.* 2:52–53, 1978.