Preliminary Report of Detecting Microembolic Signals in Transcranial Doppler Time Series With Nonlinear Forecasting

R.W.M. Keunen, MD, PhD; C.J. Stam, MD, PhD; D.L.J. Tavy, MD; W.H. Mess, MD; B.M. Titulaer, PhD; R.G.A. Ackerstaff, MD, PhD

Background and Purpose—Most algorithms used for automatic detection of microembolic signals (MES) are based on power spectral analysis of the Doppler shift. However, controversies exist as to whether these algorithms can replace the human expert. Therefore, a different algorithm was applied that takes advantage of the periodicity of the MES. This so-called nonlinear forecasting (NLF) is able to detect periodicity in a time series, and it is hypothesized that this technique has the potential to detect MES. Moreover, because of the lack of prominent periodicity in both the normal Doppler signals (DS) and movement artifacts (MA), the NLF has a potential to differentiate MES from normal blood flow variations and MA.

Methods—Twenty single MES and 100 MA were selected by 2 human experts. NLF was applied to MES and MA and compared with 200 randomly chosen DS. NLF resulted in a so-called prediction value that ranges from 1 in signals with prominent periodicity to 0 in signals that lack periodicity.

Results—NLF revealed that MES are more predictable than the normal Doppler signals (prediction [MES] = 0.829 ± 0.084 versus prediction [DS] = −0.060 ± 0.228; P < 0.0001). Moreover, MES are more predictable than the MA (prediction [MA] = −0.034 ± 0.223; P < 0.0001). No difference in prediction could be found between DS and MA.

Conclusions—This preliminary report shows that MES can be separated from DS and MA by NLF. Research is needed as to whether this technology can be further developed for automatic detection of MES. (Stroke. 1998;29:1638-1643.)

Key Words: emboli ■ ultrasonics ■ cerebrovascular diseases ■ nonlinear analysis

Since Spencer and colleagues described the on-line detection of cerebral embolism by transcranial Doppler (TCD) during carotid surgery, a vast number of reports have been published to describe the Doppler phenomena of these emboli. Emboli that pass through a sample volume of the Doppler beam result in a so-called microembolic signal (MES). MES are characterized by a unidirectional transient increase of the power in the Doppler spectrum and a typical musical sound. This musical sound is the result of the harmonic character of the MES seen in the Doppler signal (Figure 1).

For clinical application, it is vital that MES can be separated from intensity fluctuations of the Doppler signal that are caused by normal blood flow and artifacts. One of the most relevant artifacts, which may lead to erroneous embolus classification, arises from subtle probe movements. These so-called movement artifacts (MA) normally give bidirectional intensity increases in the lower frequency range of the velocity spectrum, but they may resemble a MES. On the other hand, small solid particles that result in only a minor increase in intensity may erroneously be classified as random fluctuations. It might be that these low-intensity signals are forerunners of larger solid emboli that could harm brain function.

Most algorithms used for automatic detection of MES are based on a fast Fourier transform (FFT) of the Doppler signal. Criteria used in these algorithms are often based on duration of the signal, intensity increase, and frequency distribution within the spectrum. Unfortunately, these algorithms are often not powerful enough to discriminate between MES and MA, which restricts the application of automatic emboli detection in a clinical setting. Many positive counts are caused by false interpretation of artifacts, whereas many small emboli, which do not reach the critical detection intensity threshold set by the automatic device, are not detected at all.

Therefore, we sought an alternative approach. First, time-domain data of the Doppler signal were analyzed instead of the FFT used in most automatic emboli detection devices. Second, we did not choose to detect intensity or duration of the time-domain data but took advantage of the harmonic character of the MES. To classify the harmonic character, we...
used an algorithm developed by Sugihara and May. They developed their so-called nonlinear forecasting (NLF) as a way to detect structure in short time series. NLF is an algorithm similar to the FFT, but it is not developed to reveal a spectral analysis but it is used to quantify the presence of periodicity within a time series. The basic idea is that if periodicity occurs in a signal, it should be easy to forecast the dynamics of that particular signal if one has a sample of that signal in which the periodicity occurs. On the contrary, if a signal lacks any periodicity, the dynamics will be difficult to forecast. For instance, a pure sine wave is perfectly predictable, whereas white noise is not predictable at all. Most biological signals fall somewhat between these two extremes. On visual inspection, MES have a much more prominent periodicity than the normal fluctuations of the Doppler signal (Figure 1). MA lack periodicity on visual inspection (Figure 2), and, on the basis of these observations, we thought it should be possible to distinguish MES and MA in Doppler time series. The purpose of this preliminary report is to assess whether NLF can be used to distinguish MES from normal Doppler signals or MA.

Materials and Methods

Twenty single MES and 100 MA were selected by 2 human experts (R.G.A.A. and W.H.M.) during carotid surgery. It was uncertain whether the MES were reflections from gaseous or solid emboli. MES fulfilled the criteria published by the Consensus Committee of the Ninth International Cerebral Hemodynamic Symposium. MES were identified on the basis of their short duration (<300 ms), their typical musical sound, a unidirectional appearance in the Doppler velocity spectrum, a random occurrence in the cardiac cycle, and an amplitude exceeding the background signal by at least 3 dB. The movement artifacts were recorded during artificial probe movements. They were characterized by the absence of the high frequency sound, a duration of >300 ms, and a bidirectional appearance in the Doppler velocity spectrum.

TCD monitoring was performed by means of a PIONEER with a 2-MHz monitoring transducer. The MCA was insonated just lateral to the terminal internal carotid artery. The pulse repetition frequency was adequately chosen to record the maximum blood flow velocity in the MCA. The time-domain data of the Doppler signal were sampled at the highest frequency rate and stored in an ASCII file for off-line analysis.

Software for NLF calculations was developed at the department of Clinical Neurology and Neurophysiology of the Leyenburg Hospital and written in Borland Pascal 7.0 for Windows by one of the authors (C.J.S.). Figures display on their axis intensities and time scales in

![Figure 1](http://stroke.ahajournals.org/)

**Figure 1.** MES observed during carotid surgery. Embolic signal characterized by a gradual increase in intensity and a more prominent periodicity compared with background Doppler signal (on vertical axis an arbitrary scale of the Doppler signal, on horizontal axis an arbitrary time scale).

![Figure 2](http://stroke.ahajournals.org/)

**Figure 2.** Movement artifact observed during artificial probe movement during insonation of middle cerebral artery. Movement artifact is characterized by sudden increase in intensity compared with background Doppler signal. Signal lacks prominent periodicity (on vertical axis an arbitrary intensity scale of the Doppler signal, on horizontal axis an arbitrary time scale).
arbitrary units. Absolute values are irrelevant for NLF calculations; the method quantifies changes over time independent of the absolute value. Statistical analysis was done by the statistical package Systat (version 5.1 for Windows). Student independent t tests were applied for discriminating between-group effects. P values were set at 0.05 for statistical significance.

Nonlinear Forecasting

Nonlinear forecasting explores the functional relation between future and previous data in a time series. After visual inspection of the Doppler signal, a time window was placed across the MES or MA (Figure 3). NLF was performed on the time series within the window. The time window always included a part of the MES in which the periodicity was most prominent on visual inspection. Each MES and MA was compared with 10 randomly chosen normal TCD signals of similar length (Figure 3).

The time series within the window was split in 2 equal parts: the first part of the time series (t1) and the following or second part (t2). The goal was to analyze t1 to predict the time series t2. Once the time series t2 is predicted, one can compare the actual t2 with the predicted t2 and calculate the difference between these two time series. If both time series completely match, the prediction is 100%, but if both time series completely mismatched, the signal is in fact unpredictable. For periodic signals the prediction (p) is 1.0, which means that the signal is 100% predictable. For uncorrelated white noise, the prediction will be 0; if the prediction is exactly the opposite of the original curve (a hypothetical situation) then P = -1.0. The prediction therefore varies between +1 and -1.

To perform this “prediction” calculus, one has to know the dynamics of t1. To study the dynamics of a signal, mathematicians have developed a 2-step procedure. The first step is called “reconstruction of trajectories in phase space”; the second step involves characterization of the reconstructed dynamics. The reconstruction of trajectories in phase space is a mathematical procedure by which the time domain data are reconstructed in at least a 2-dimensional figure (the phase space) to study the coherence of all data points over time of the signal in a visually instructive manner, which in addition facilitates the calculation of the dynamics of such a signal. The reconstruction is done by a procedure of “time-delay embedding,” a procedure explained in the Appendix. Calculus behind the NLF is also explained.

To make the reader familiar with trajectories in phase space, we show in Figure 4 an example of trajectories of a sine wave with some additive noise in phase space. Notice that the trajectories form a closed loop and the noise component results in a relative broad banding of the trajectories. Figure 5 shows the phase space of the MES signal given in Figure 1. It shows a cluster of trajectories in the center of the phase space with a number of trajectories at the periphery of the phase space. The center trajectories relate to the background TCD signal, whereas the trajectories at the periphery relate to the MES. The trajectories at the periphery appear to have less complex dynamics than the trajectories in the center of the phase space, which suggests that MES have a more prominent periodicity than background TCD.

Results

Figure 6 shows the predictability for MES, MA, and background Doppler signals. MES show the highest prediction values, which range from 0.97 to 0.67. The prediction of the MA and background Doppler signals are similar and ranged between 0.45 and -0.80 for MA and

Figure 4. Phase space of sine wave with some additive noise. x-axis shows value of signal at a certain moment X(t); y-axis shows value of signal at fixed interval ahead in time [X(t+ Δt)]. Trajectories form closed circles and because of additive noise trajectories show relatively broad banding.
from 0.44 to -0.69 for Doppler signals. Statistical analysis reveals that MES (n=20) are significantly better predictable than the background Doppler signals (n=200) (prediction [MES]=0.829±0.084 versus prediction [Doppler signals]= -0.060±0.228; P<0.0001). Moreover, MES are also significantly better predictable than the MA (n=100) (prediction [MES]=0.829±0.084 versus prediction [MA]= -0.034±0.223; P<0.0001). No significant difference in prediction is found between background Doppler signals and MA.

### Discussion

This study shows that MES can be separated from both the background TCD signal and MA by NLF. This opens the possibility to design algorithms that are based on the NLF principles for automatic embolus detection in a clinical setting. Emboli detection is often a time-consuming and mentally strenuous procedure. The large number of patients that are candidates for this examination will rapidly grow, and automatic software devices are required to replace the human expert. However, Van Zuilen and coworkers showed that current automatic software devices for embolus detection are not yet capable of being used as a “stand-alone system.” The interobserver agreement of human experts is still much higher than the agreement that can be reached by the use of automatic software devices. Early automatic detection devices relied on a sudden increase in intensity of the returned signal. More sophisticated devices recognize the bell-shaped increase in the relative power occurring with an embolus. Some include a so-called artifact rejection algorithm, which looks for the occurrence of a bidirectional power increase. The most promising approach seems to be embolus detection by neural networks, but one still must realize that a neural network can only classify embolic events with a high accuracy when the training of the network allows such an identification. Therefore it is of the utmost importance that the input of the neural networks is based on information that characterizes the embolic events. We think that NLF as used in this study will be very useful as an input variable for future neural networks designed for emboli detection.

The reason that nonlinear analysis is so powerful in detection of emboli lies in the fact that it quantifies an essential character of emboli: the prominent periodicity of the MES. This periodicity of MES is related to the fact that the Doppler signal contains information of a relatively constant velocity: the velocity at which the embolus travels. However, the velocity of the embolus is not completely stable. During the travel through the sample volume it may accelerate, decelerate, or change its direction. All these possibilities result in a frequency modulation. The MES shown in Figure 1 shows such a characteristic frequency modulation. Initially the periodicity has a low intensity and a relatively high frequency when the embolus enters the sample volume. When the embolus is in the middle of the sample volume the intensity reaches its maximum value, and when the embolus leaves the sample volume both the frequency and intensity of the signal decreases. Especially air emboli appear to produce this frequency modulation compared with the relative stable frequency modulations observed in particulate emboli. Nevertheless, although the frequency of the MES modulates when the embolus travels through the sample volume, the periodicity is strong enough to result in a marked NLF of the actual signal compared with the background Doppler signal. Therefore we strongly support the idea that the prominent periodicity of the embolic signal should be included in the definition of embolic events.

Another important remark is that this technique has a potential to detect low intensity emboli. Although we did not analyze the relation between signal intensity and prediction in this particular study, we noticed, for example, in Figure 3 that both MES (A and B) did have the same prediction values although both MES differ in their intensities by visual inspection. In future studies we will focus on this important relation between intensity and prediction of MES in a more systematic way.
One remark must be made about NLF and MA. MA can have a typical low frequency sound that theoretically could lead to some degree of predictability beyond the level of the background Doppler signal. The reason that NLF did not reveal an increased predictability in MA is because these Doppler signals were examined on a relatively short time scale. During these short observations the MES show, in contrast to MA, prominent periodicity. Thus MA may exhibit a certain periodicity albeit on a much longer time scale than applied in this study.

As mentioned by Markus and Harrison, it is vital that a consensus on the optimal definition of emboli is developed and validated in subsequent studies. We think that neither the duration nor the maximum intensity itself are the essentials of MES. It is the periodicity that describes the core characteristics of the MES. A number of nonlinear algorithms are available for signal analysis. Originally, these algorithms were designed to detect the underlying dynamics of the systems that generate the signal. Most of them are very complex and time-consuming calculations. For detailed information of these algorithms, the reader is referred to Kaplan and Glass, who clearly explain the background of nonlinear analysis. The reason to choose the NLF is the fact that this method is a relatively simple and fast procedure (capable for both off-line and on-line analysis) with a strong capacity to distinguish signals with different underlying dynamics. If we were unaware of the system that generates the signal (eg, the interaction of the embolus and the ultrasound beam), we should conclude from this analysis that the underlying system of the MES has a high predictability, which means that the degrees of freedom of such a system probably will have only 1 variable. Obviously, the relatively constant frequency of the Doppler signal, being a result of the actual embolus velocity, is the most important variable that leads to a prominent periodicity.

Appendix

Time Delay Embedding

The recorded TCD is a discrete time series \( V_t \), \( t=1,2,3,\ldots,N \). From this discrete time series, vectors \( X_i \) in a \( m \)-dimensional embedding dimension were obtained with the time-delay procedure:

\[
\begin{align*}
X_i &= (V_{t-iL}, V_{t-(i+1)L}, V_{t+(i+2)L}, \ldots, V_{t+(m-1)L})
\end{align*}
\]

where \( L \) is the time delay and \( m \) is the embedding dimension. We used a combination of the procedure of Rosenstein to choose \( L \) and the procedure of Kennel to choose \( m \). Briefly, this method works as follows: We started with an embedding dimension of 1 and a lag of 1. We then calculated the expansion (“unfolding”) of the attractor from the main diagonal in state space. Subsequently, the lag was increased in steps of 1 until the expansion of the attractor no longer increased. The percentage of false nearest neighbors was then calculated. False nearest neighbors are vectors that lie close together in the reconstructed state space because of insufficient unfolding of the attractor and not because of dynamic correlations. Following Kennel et al (1992), 2 vectors \( X_i \) and \( X_j \) were considered false nearest neighbors when

\[
|V_{i+(m-1)L} - V_{i+(m-2)L}| > 15 \quad \text{for} \quad R_m(i,j) > 0
\]

where the vertical bars denote the absolute value and \( R_m(i,j) \) is the euclidian distance between the two vectors \( X_i \) and \( X_j \).

Next, the embedding dimension \( m \) was increased with 1, and the whole procedure (determining the optimum lag and then the percentage false nearest neighbors) was repeated.

This procedure was continued until either (1) The percentage of false nearest neighbors dropped under 0.05 or (2) The percentage false nearest neighbors no longer decreased with further increments of the embedding dimension. In the last case, we used the value of \( m \), which gives the lowest percentage of false nearest neighbors.

Nonlinear Forecasting

We used the algorithm described by Sugihara and May for nonlinear forecasting. Given a starting point in the time series \( V_t \), we would like to predict \( V_{t+m} \), \( V_{t+2m}, V_{t+3m}, \) and so on, a number of steps ahead, and compare the predictions, which we will designate \( P_{t+m} \) with the actual time series. First we used the time-delay procedure described above to obtain \( m \)-dimensional vectors \( X_t \) from the time series \( V_t \). However, for ease of reference, the time index of the vector \( X_t \) will now correspond with the last coordinate \( i=t+(m-1)L \). Now for each vector \( X_t \) we located the \( m+1 \) nearest neighbors in the \( m \)-dimensional state space. We will designate the \( k \) nearest neighbor vectors of \( X_t \) as \( NN_{k,j} \). The \( k \) index indicates the number from 1 to \( m+1 \) of the nearest neighbor (the \( k,j \) index is its time index in the original time series).

We excluded nearest neighbors with time indexes \( k,j \) when \( j-k \leq 3 \times r \) (\( r \) is the time after which the autocorrelation function drops to 1/e of its original value). This procedure is called “within-sample” prediction. Sugihara and May used “out-of-sample” prediction. For a time series of length \( N \), out of sample prediction requires \( i>0.5 \times N \) and \( j<0.5 \times N \) constant.

There are no fundamental differences between both procedures, only within-sample prediction may be more suitable for short data sets. Now the predicted value for \( n \) steps ahead prediction was given by

\[
P_{t+(n+1)L} = \sum_{k=1}^{m+1} V_{(k+1)L} \times W_k
\]

where \( W_k \) is the weight assigned to the \( k \)th nearest neighbor according to

\[
W_k = \frac{1}{\sum_{k=1}^{m+1} |X_{t+L} - NN_{k,j}|^2}
\]

Predictions for \( n \) ranging from 1 to 20 were made for a number of 100 different \( V_t \) evenly distributed along the time series. Next, the correlation coefficients \( r \) between the actual time series \( V_{t+n} \) and the predicted values \( P_{t+n} \) were calculated and plotted as a function of \( n \).

References


Preliminary Report of Detecting Microembolic Signals in Transcranial Doppler Time Series With Nonlinear Forecasting


Stroke. 1998;29:1638-1643
doi: 10.1161/01.STR.29.8.1638

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/29/8/1638