Artificial neural networks: application to electrical stimulation of the human nervous system

Richard B. North, M.D., J. Paul McNamee, M.S., Lee Wu, M.S., and Steven Piantadosi, M.D., Ph.D.

Departments of Neurosurgery and Biostatistics, School of Medicine and Applied Physics Laboratory, Johns Hopkins University, Baltimore, Maryland

Artificial neural networks are used increasingly in applications such as graphic pattern recognition, which are difficult to address with conventional statistical methods. In the management of chronic pain, graphic methods are used routinely; patients describe their patterns of pain using "pain drawings." The authors have previously reported an automated, computerized pain drawing methodology, which has been used by patients with implanted spinal cord stimulators to represent a technical goal of the procedure, the overlap of pain by stimulation paresthesias. Standard linear discriminant statistical methods have shown associations between stimulation parameters and electrode positions as independent variables and technical outcome and relief of pain as dependent variables.

The authors have applied artificial neural networks to the problem of optimizing implanted stimulator adjustment. A data set of 3000 electrode combinations obtained in 41 patients was used to develop a linear discriminant statistical model on a mainframe computer and to train artificial neural networks on a personal computer. The performance of these two systems on a new data set obtained in 10 patients was compared with that of human "experts." The best neural network model was marginally better than the linear discriminant model; the variance in patient ratings was predicted by these models to a degree that the human experts were unable to predict. The authors anticipate expanding the role of these models and incorporating them into expert systems for clinical use.

Key Words * artificial neural networks * chronic pain * electrical stimulation * pain drawing * spinal cord * low-back pain

Artificial neural networks are increasingly used to accomplish tasks such as graphic pattern recognition, which conventional statistical methods address only with difficulty. Graphic methods are used routinely in the management of chronic pain; patients describe their patterns of pain using "pain drawings." Neural network techniques and artificial intelligence techniques have been used to classify pain drawings and to diagnose patients with low-back conditions.[1,4,5]

We have previously reported a computer-based pain drawing methodology, which has been used by our patients with implanted spinal cord stimulators to represent the overlap of pain by stimulation paresthesias, a technical goal of the procedure.[7] Standard linear discriminant statistical methods have shown some associations among stimulation parameters and electrode geometries as well as technical
outcome and relief of pain.[8] The present study supplements these methods with computerized neural networks.

**CLINICAL MATERIAL AND METHODS**

**Patient Population and Data Collection**

We have developed a patient-interactive neurological stimulation system (NSS) for routine adjustment of implanted spinal cord stimulators used for pain control. The system has been described in previous published reports.[7,8] Figure 1 shows the system's patient interface: a graphics tablet and color monitor. A series of patients with chronic low-back and lower-extremity pain have used the NSS in the course of treatment, providing the data for this study.

Fig. 1. Photograph showing the graphics tablet used by patients to draw areas of pain and paresthesias, to adjust stimulation amplitude, to enter 100 mm visual analog ratings, and to answer simple yes-no questions. Graphics, instructions, and questions for the patient are displayed on a personal computer color monitor. The computer keyboard is used only by the system operator and not by the patient. A separate enclosure (not shown), connected to the computer serial port, contains a microprocessor controlling two radiofrequency transmitters.

Each patient tested the available electrode combinations. These were presented in random order by the computer at a fixed pulse repetition rate of 60 per second; the amplitude was set by the patient at a comfortable level. The patient used the system's graphics tablet to record a "pain drawing," outlining the topography of pain, and then to record corresponding drawings for the areas of stimulation paresthesias. Overlap of pain by paresthesias was rated by the patient on a scale of 0 to 100.

The patient rating of overlap was used as the "gold standard" among three dependent variables representing the outcome of each trial. Overlap was also calculated by the system as the ratio of the number of pixels common to the two drawings and the number of pixels in the pain drawing. Extraneous stimulation, outside the pain drawing, was quantified similarly. Our earlier work has shown a high correlation between patient ratings of overlap and these graphic measures.[8] Figure 2 illustrates these concepts.
Fig. 2. Illustration of pain overlap models. Overlap of pain by stimulation paresthesias is a necessary condition for pain relief by spinal cord stimulation. This overlap is calculated quantitatively by the system from patient drawings. Overlap is calculated as the number of pixels (within the body outline) common to the pain drawing and the stimulation drawing divided by the number of pixels in the pain drawing. Extraneous stimulation is calculated as the number of pixels (within the body outline) that are within the stimulation drawing but outside the pain drawing divided by the number of pixels in the stimulation drawing. Pain drawings are characterized by calculating overlap with specific body regions.

The same graphic methods of overlap calculation were used to decompose pain drawings into body areas. For each area of interest (for example, left low back), a drawing was created, and overlap by each patient's pain drawing was calculated from pixel counts as a percentage. Each pain drawing was therefore characterized by a set of 30 variables. In addition, each patient was characterized by the vertebral level of the electrode array (for example T10.5 = middle of T-10 vertebral body as seen on anteroposterior x-ray film). Each combination of contact assignments tested was characterized by the anode and cathode positions, intercontact distances, and orientations. In total, 44 independent variables characterized each of the 3000 trials in the data set. The independent and dependent variables are listed in Table 1.

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Independent Variables (44)</strong></td>
<td><strong>Dependent Variables (3)</strong></td>
</tr>
<tr>
<td>overlap of pain drawing w/ each of 30 body areas</td>
<td>patient rating of overlap of pain by stimulation paresthesias</td>
</tr>
<tr>
<td>vertebral level of electrode</td>
<td>overlap calculated from drawings of pain &amp; stimulation</td>
</tr>
<tr>
<td>amplitude (in volts) of stimulation anode &amp; cathode assignments</td>
<td>extraneous stimulation outside pain drawing</td>
</tr>
<tr>
<td>dipole length &amp; orientation</td>
<td></td>
</tr>
</tbody>
</table>

**Study Design**
The study objective was a priori prediction of the rank order of each patient's ratings of the 50 available contact assignments, given the patient's pain drawing and the vertebral level of the electrode array. The human experts worked from hard copies of the patients' computerized drawings and relied on their clinical experience to predict the rank order. The statistical model and the neural network models were developed from a data set of 3000 combinations obtained in 41 patients. A separate set of 500 combinations obtained in 10 other patients was used as the test set. This allowed us to assess the accuracy of the clinicians and the two computer models in predicting the patients' ratings.

The predictions from the linear regression model were compared with predictions from the neural networks and from the clinicians; rank order lists of electrode combinations were generated for each patient by each method. Predicted rankings were compared with the rankings of the actual patient ratings, using Spearman correlation coefficients.

**Statistical Methods**

*Linear Discriminant Model.* The patient rating was considered as the dependent variable and the pain drawing areas and electrode characteristics as the independent variables in the analysis. Independent variables were screened for associations using ordinary regression methods to select a smaller set of variables that might be useful for predicting the response to a particular electrode combination. Duplicate entries (caused by a patient rating the same combination twice) were eliminated, and thus only 2447 of the 3000 entries were used for this analysis.

In the second step, the significant factors from the screening regressions were used to predict patient responses. Because each patient contributed approximately 50 or 60 observations to the data file, we used a generalized estimating approach and a robust estimate of the covariate matrix for correlated regressions. The size of the covariate matrix exceeded the capacity of a desktop computer; therefore a mainframe computer was used. A parsimonious prediction model was generated in this fashion.

In the third step, data obtained in 10 patients were used in a validation study to predict patient ratings from the linear regression model.

*Neural network models.* Neural network models were developed on a desktop personal computer. All of the networks had 44 numeric inputs and three outputs. Ninety percent of the training set data was used to train the networks and the other 10% was used as a benchmark to measure network generalization. Based on these benchmark results, four different neural network models were chosen from an initial field of nine models with different architectures. These completed, trained models were then used on the test set of 10 new patients. The following network architectures were selected. 1) A backpropagation algorithm with 18 hidden nodes. The learning rate was 0.5 and a momentum term of 0.3 was used (to smooth the gradient descent over the error surface). The training set was presented to the network 127 times as the network updated its connections between nodes. The connections were updated using batch (or epoch) training, whereby the errors in the prediction of several (20-30) combinations were summed and then used to update the weights. 2) A backpropagation algorithm with six hidden nodes. The learning rate was 0.5 and a momentum term of 0.3 was used. The training set was presented to the network 51 times. 3) A backpropagation network using software by NeuralWare with 10 hidden nodes. The learning rate and momentum term were dynamically adjusted based on system defaults. The data set was presented to the network 29 times. 4) A backpropagation network using software by NeuralWare with 10 hidden nodes. The network was presented the training set eight times.
Each neural network was "trained" on the 3-MB data set described above, to a specific criterion, that is, 85% accuracy in predicting patient rating within 15 points. Each training session would run for a few hours or sometimes overnight. Inputs were presented to the neural networks in randomized order so that data from an individual patient would not be presented all at one time.

RESULTS

Rank order lists of the available electrode combinations were developed for each patient by the methods previously described, using the linear discriminant statistical model and the artificial neural networks on a personal computer. The performance of these systems was compared with rank order lists developed by human "experts." The best neural network model was marginally better than the linear discriminant model; both explained a small percentage of the variance in outcome, which the human experts were unable to predict. Spearman correlation coefficients were 0.130 for the best neural network, 0.113 for the linear regression model, and 0.006 for the best clinician.

DISCUSSION

Neural Networks

Contemporary "artificial neural networks" are processing devices that range from software algorithms running on general purpose computers (as used in this study) to specialized hardware. Artificial neural network designs are inspired by naturally occurring biological systems. They consist of simple, multiple processors ("units" or "neurodes," analogous to neurons) and unidirectional communication channels ("connections," analogous to synapses). These connections may be arranged to excite or inhibit and to propagate backward or forward between units. Units commonly are arranged in layers, like their biological counterparts. The units have input layers, intermediate or "hidden" layers, and output layers. Figure 3 shows the network design used in this study: a three-layer "backpropagation" network.

Fig. 3. Illustration of a neural network. The backpropagation neural networks used here have three interconnected layers of "neurodes." The number of neurodes in the input and output layers is determined by the study data set; the number in the hidden layer may be varied to modify the network's performance in generalized learning or memorizing data sets.

The response of a neural network to any set of inputs is determined not only by its network topography, but also by the strength of its connections, which may be modified individually. These modifications are
based on the difference between observed and expected output ("error signal") for a given set of inputs. Neural networks thus "learn," like their biological counterparts. After repeated presentations of data, learned algorithms minimize the average error signal. Backpropagation networks use an iterative gradient descent procedure. A momentum term smooths the gradient descent over the error surface, increasing the likelihood of finding the global minimum rather than a local minimum in the surface.

"Learning" generalized rules is to be distinguished from learning by rote, that is, memorizing a data set. In the present study, networks "learned" with a high degree of accuracy how the patients in the training set rated contact combinations. For example, 85% of the predictions were within 15% of observed values; however, the measure of clinically useful performance involved a separate test set. Accordingly, network architectures were chosen and training runs were conducted to minimize memorization. There is a trade off between the size of the hidden layer and the ability of the network to generalize its learning. Larger hidden layers can solve problems of greater complexity; however, they are also susceptible to memorizing their data. Smaller layers trained with fewer passes of the training set can be used to attempt greater generalization.

Neural networks are applied increasingly to medical problems, particularly pattern recognition, such as the diagnosis of low-back pain and sciatica and even the interpretation of pain drawings.[1,3-5] Discriminant analysis and neural networks (multilayer perceptrons) perform similarly and may be considered equivalent when the latter are linear.[2] To the extent that the neural networks outperformed the linear discriminant model in this study, we may infer nonlinearity in our patient data.

Clinical Application

Patients with persistent, intractable pain, exemplified by low-back and lower-extremity pain after multiple surgeries, are managed in several ways. Anatomical procedures may be used to correct an abnormality thought to cause the pain. Ablative procedures may be performed to destroy neural pathways that transmit pain signals. Augmentative procedures that superimpose electrical stimulation or medications that act on the intact nervous system may also be administered. Augmentative procedures, which are reversible and amenable to straightforward trial before devices are permanently implanted, are used increasingly because of their obvious advantages.

Spinal cord stimulation is a neuroaugmentative procedure: electrodes are implanted over the dorsal surface of the spinal cord and electrical pulses are delivered from an implantable generator. At stimulus amplitudes adequate to relieve pain, this produces a paresthesia. Overlap of the topography of a patient's pain by paresthesias is a necessary condition for pain relief, and there is a highly significant association between overlap and patient ratings of pain relief.[6] The position of the implanted electrode(s) largely determines the location of stimulation paresthesias. To maximize the likelihood of achieving the desired overlap and to minimize the need for repeated surgical revisions of electrode position, manufacturers have developed arrays of multiple contacts.

As the number of contacts has increased, however, the task of adjustment (thoroughly testing the available anode and cathode assignments) has grown enormously. For a simple contact pair, there are only two assignments to test; for four contacts, there are 50; for eight contacts, there are 6050. The time required of the patient, the physician, and the staff to test these combinations is substantial. If adjustments are made manually, representing the data quantitatively for comparisons and retesting is cumbersome and becomes prohibitively time consuming. This has motivated the development of computerized methods.[7,8]
Even using computerized methods that require minimal patient supervision after initial training, the task of adjustment can become prohibitively time consuming if all possible combinations are to be tested exhaustively. For an eight-electrode array, even at a rate of 50 combinations per hour, it would take 121 hours to test all the possibilities. The performance of different electrode combinations may be assumed to be interdependent, and thus representative subsets may be defined and search strategies developed so that not all combinations need to be tested. Validating this assumption may be difficult for large numbers of contacts and for complex (two-dimensional) geometries. In the present study, the available 50 combinations for an array of four electrodes is a manageable number; they have been tested exhaustively, allowing validation of our models.

To date linear discriminant methods have yielded the finding that a particular electrode configuration ("guarded cathode" or "split anode," that is, cathode[s] flanked by anode[s] above and below) is preferred by our patients to a statistically significant degree.[8] In this prior work, as in the present study, the importance of overlap of pain by stimulation paresthesias in achieving the clinical goal of pain relief is assumed; this has been established before and is beyond the scope of the present study.[6]

**CONCLUSIONS**

This study addresses only a limited problem: a priori ranking of electrode combinations based on pain drawing and electrode level. In clinical practice, a human expert chooses and tests combinations serially and may base each choice on the results of the last test. The procedures developed here need to be supplemented by a set of rules (expert system) and implemented iteratively in real time. Such a system might use the output of a neural net, along with incoming data, as its inputs. A corollary problem is to define a preliminary test set to guide the adjustment process. Rather than attempting to find the best combination first, a more effective strategy may be to find the combinations that give the most information so that the adjustment process is ultimately more expeditious.

---

**References**


Manuscript received November 19, 1996.

Accepted in final form December 18, 1996.

Address reprint requests to: Richard B. North, M.D., Department of Neurosurgery, Johns Hopkins University School of Medicine, 600 North Wolfe Street, Meyer 7-113, Baltimore, Maryland 21287-7713.