The Use of Polynomial Neural Networks for Mortality Prediction in Uncontrolled Venous and Arterial Hemorrhage

David A. Roberts, MS, John B. Holcomb, MD, FACS, B. Eugene Parker, Jr., PhD, Jill L. Sondeen, PhD, Anthony E. Pusateri, PhD, William J. Brady, Jr., MD, FACEP, David E. Sweenor, and Jeffrey S. Young, MD, FACS

Background: **The ability to rapidly and accurately triage, evacuate, and utilize appropriate interventions can be problematic in the early decision-making process of trauma care. With current methods of prehospital data collection and analysis, decisions are often based upon single data points. This information may be insufficient for reliable decision-making. To date, no studies have attempted to utilize data at multiple time points for purposes of enhancing prediction, nor have studies attempted to synthesize prediction models with data reflecting both large-vessel venous and arterial injuries. Therefore, we performed a retrospective study to examine the potential utility of dynamic neural networks in predicting mortality using highly discretized uncontrolled hemorrhagic shock data.**

Methods: **One hundred forty-three swine with either grade V liver injuries or 2.8-mm aortotomies had hemodynamic data collected every minute throughout injury and resuscitation. The independent variables used as inputs to the polynomial neural networks (PNNs) included systolic blood pressure and mean arterial pressure (MAP). These inputs were used to predict mortality in individual swine 1 hour after injury using data up to 20 minutes after injury. Survival models were compared based on discrimination power (DP), i.e., where specificity equals sensitivity, and area under the receiver operating characteristic (ROC) curve (*c*-statistic). The Hosmer-Lemeshow (H-L) statistic was used to measure model calibration.**

Results: **The best PNN model predicted mortality at 60 minutes utilizing data from injury to 20 minutes after injury. This model produced a ROC area of 0.919, a DP of 0.857, and a H-L value of 16.47. A DP of 0.857 means that 85.7% of the survivors are correctly predicted to survive, and 85.7% of the nonsurvivors are predicted to die. MAP of survivors and nonsurvivors were graphed for comparative purposes. As this graph illustrates, the use of MAP alone cannot discriminate survivors from nonsurvivors.**

Conclusion: **This study demonstrates that PNN models can effectively harness the dynamic nature of uncontrolled hemorrhagic shock data, despite utilizing data from large-vessel arterial and venous injuries. Utilizing the dynamic nature of hemorrhagic shock data in PNNs may ultimately allow the development of novel decision assist devices.**

Key Words: **Hemorrhagic shock, Polynomial neural network, Injury, Triage, Mortality, ROC area, Discrimination power**

J Trauma. 2002;52:130–135.

With the invention of the sphygmomanometer in 1891 came the first noninvasive physiologic sensing device. Now, over 100 years later, there is a major effort by both civilian companies and the military to develop new noninvasive sensor technologies and improve the accuracy of existing sensors. Military projects focusing on prehospital frequently sampled digital data collection have spearheaded the development of advanced noninvasive sensors, including pulse oxymetry, capnography, and a durable digitized electrocardiogram.[1]

Equipment in these sophisticated monitoring systems collect highly sampled discretized data, mimicking the trend analysis common in the intensive care environment. If collected in the prehospital environment, similar digitized data could greatly change the relatively crude decision-making algorithms currently utilized in prehospital trauma triage and injury severity assessment. Standard trauma triage algorithms currently used in prehospital and hospital settings, such as the Revised Trauma Score (RTS) and Trauma and Injury Severity Scoring (TRISS), use only physiologic data at one time point.[2–7] When these and other algorithms were developed, collection of highly sampled data for real-time triage and assessment was not an option due to the lack of sensor technology and computing power. With these technological advances in hardware, sophisticated algorithms, such as neural networks, can analyze and integrate hemodynamic discrete-time data and provide the near instantaneous results desired by clinicians.

Submitted for publication December 22, 1999.

Accepted for publication September 24, 2001.

From Barron Associates, Inc. (D.A.R., B.E.P.), and Departments of Emergency Medicine (W.J.B.) and Surgery (J.S.Y.), University of Virginia School of Medicine, Charlottesville, Virginia, Department of Trauma and Critical Care (J.B.H.), UT-Houston Medical School, Houston, U.S. Army Institute of Surgical Research (J.L.S., A.E.P.), Ft. Sam Houston, Texas, and Department of Applied Physics and Astronomy (D.E.S.), Rensselaer Polytechnic Institute, Troy, New York.

These SBIR data and results are furnished by Barron Associates, Inc., with SBIR rights under contract No. DAMD17-96-C-6022 P70001 titled “Neural Network Medical Decision Algorithms for Pre-Hospital Injury Severity and Risk Assessment.”

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision unless so designated by other documentation.

Address for reprints: David A. Roberts, Barron Associates, Inc., 1160 Pepsi Place, Suite 300, Charlottesville, VA 22901-0807; email: roberts@bainet.com.

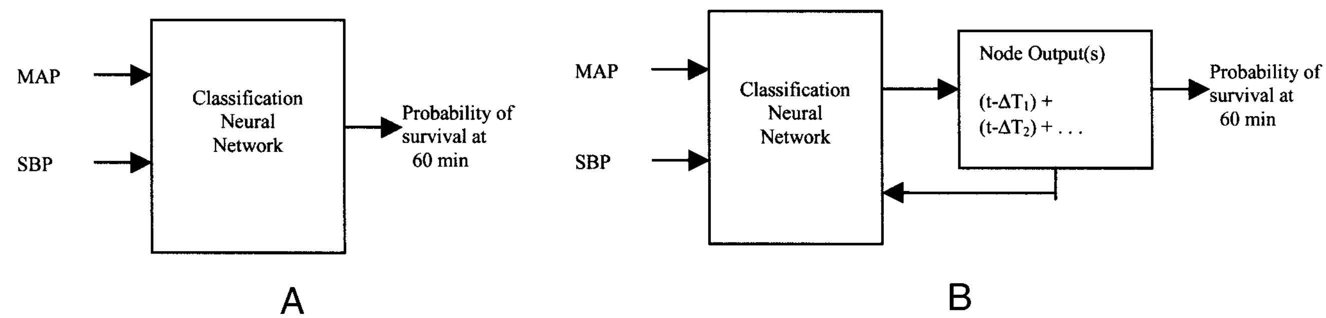

Fig. 1. *(A) Static PNN. (B) Dynamic PNN.*

The purpose of this study is to examine the utility of polynomial neural networks (PNNs) in predicating mortality during uncontrolled hemorrhagic shock secondary to severe liver or aorta injury. We hypothesized that the PNN models would be able to learn the dynamic characteristics of the venous and arterial uncontrolled hemorrhagic shock data for use in mortality prediction soon after the injury.

MATERIALS AND METHODS

Swine Hemorrhagic Shock Database

The database used for this retrospective analysis was compiled from several different experiments examining resuscitation techniques on swine in hemorrhagic shock. The hemorrhage was induced in 79 pigs by a grade V liver injury (Liver Injury Scale of the American Association for the Surgery of Trauma).[8] Resuscitation in the different experiments included the application of a variety of hemorrhage control methods, as well as fluid administration, with a goal of rapid return to preinjury mean arterial pressure (MAP).[8,9] Hemorrhage was induced in 64 pigs by an aortotomy experiment which examined the optimal timing and rate of administration of resuscitation fluids in uncontrolled hemorrhage. One hundred forty-three commercial swine weighing 40 $\pm$ 5 kg were used in this study. All animals were maintained in a facility accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International. The protocol was approved by the Animal Care and Use Committees of the Institute of Surgical Research, Ft. Sam Houston, Texas, and William Beaumont Army Medical Center, El Paso, Texas. All animals received care in strict compliance with the *Guide for the Care and Use of Laboratory Animals* (National Research Council, 1996).

Splenectomies were performed on each pig before injury so that the physiologic reaction to the hemorrhage would be more comparable to humans; the pig spleen is contractile and can autotransfuse a significant volume of blood. Data variables were measured from preinjury until 1 hour or until the animal died. These included MAP and systolic blood pressure (SBP) sampled every minute. Missing data entries at several time iterations and entries that were deemed physiologically impossible were computed using linear interpolation of the surrounding data points. Similar techniques have been employed in previous studies to overcome the problem of missing data entries.[10]

Network Synthesis

Neural network algorithms are mathematical constructs for determining the functional relationships between two sets of data (inputs and outputs). In this context, they represent an extension of traditional regression analysis.[11] The commercial neural network development software we used for this experiment synthesizes neural networks to solve multivariate estimation and classification problems.[12] Proceeding from a numerical database of known input-output relationships, the software synthesizes a PNN from zero connections to an optimum level of complexity. As part of this process, the software then selects those variables that are most useful in modeling from the list of available candidate inputs. Linear and nonlinear algebraic nodal functions are selected automatically and interconnected by the software into networks that become global models of the functional relationships in the data.[13] The software then produces C-language code that may readily be compiled on essentially any personal computer.

In our experiment, two types of networks were synthesized by this software: a static PNN model that uses physiologic data at one time point, and a dynamic PNN model, which uses the historical data over the course of the hemorrhage. The static model, illustrated in Figure 1*A*, was trained to predict swine survival at time 60 minutes using MAP and SBP at 20 minutes into hemorrhage. The 20-minute time cutoff was chosen because it had the greatest number of pigs still alive within the 20- to 30-minute window of interest. Two additional static (without feedback) networks were trained to predict survival at 60 minutes using these variables at 16 and 18 minutes after injury for comparative purposes. Similarly, the dynamic model, illustrated in Figure 1*B*, used the independent variables MAP and SBP at 1-minute intervals over the first 20 minutes of the uncontrolled hemorrhage as inputs into the neural network to predict survival at 60 minutes. Again, similar dynamic models were trained using variables at 16 and 18 minutes for comparison to the model synthesized at 20 minutes.

Due to the limited number of pig specimens, model efficacy was determined via an *N*-fold cross-validation procedure. Under this paradigm, each pig was, in turn, withheld from use in model synthesis and then used to evaluate only that model. This process was repeated for each pig and the accumulated results were used to construct a receiver operating characteristic (ROC) curve, the area under which is known as the *c*-statistic.[14] The networks synthesized utilized linear polynomials and a logistic distortion function.

Statistical Analysis

The term "discrimination power" (DP) is defined as the value where the specificity, the percent of actual survivors correctly identified as such, and sensitivity, the percent of nonsurvivors correctly identified as such, are equal. The use of this measure avoids reporting a high value of specificity that corresponds to a poor value of sensitivity, as this can often be misleading. ROC curves plot 1 − specificity on the abscissa vs. sensitivity on the ordinate at varying decision thresholds. The area under the ROC curve is a measure that indicates how close the predictive outcomes are from random guessing, where an area of "0.5" indicates no discrimination and an area of "1.0" indicates perfect discrimination. The Hosmer-Lemeshow (H-L) goodness-of-fit statistic was computed as an indication of the quality of model calibration. This statistic compares true and model estimated outcome over a decile of risk.[15,16]

Calibration is a measure of the accuracy of a measuring instrument; in this case, a prediction model. For mortality prediction models which provide a probability for a binary outcome, the calibration would determine how close the probabilities generated by the model are to the actual outcome: a 1.0 for survivors and a 0.0 for nonsurvivors. Although the models were able to perform very well in discriminating survivors from nonsurvivors, four out of the six models had H-L goodness-of-fit statistics that were higher than the 15.5 threshold value that indicates a good fit. These high values are due, in part, to the relatively small number of swine available for training the PNN models. The use of the statistical analysis of calibration as a determination of model efficacy, however, is still under debate.[17] This is because models can discriminate 100% of the time while having poor calibration using the H-L statistic. Bars indicating 1 SD above and below mean were used in Figure 2 to illustrate that mortality discrimination is not feasible using just MAP values alone, as there is a large inconsistency in pressures of individual swine over the course of the hemorrhage.

RESULTS

Of the 143 swine in the database, 87 (60.8%) lived past the observational period, and 56 (39.2%) died before the end of the experiment. Tables 1 and 2 summarize the means and standard deviations of the physiologic inputs for the survivors and the nonsurvivors, respectively. In Table 1, we see that the two physiologic variables of the survivors revert back to baseline after an initial fluctuation at the beginning of the hemorrhage. In Table 2, we see that both the mean pressures of the nonsurvivors decrease steadily over the course of the hemorrhage. The changes reflect either effective hemostasis (Table 1) or continued bleeding (Table 2), combined with resuscitation.

Results of the static PNNs trained using 142 of the 143 animals are seen in Table 3 and the corresponding ROC curves are provided in Figure 3. From these results we see discrimination improves over the course of the hemorrhage.

Results of the dynamic PNNs trained with 142 animals are given in Table 4 and the corresponding ROC curves provided in Figure 4. In Table 4, we see that the model at time

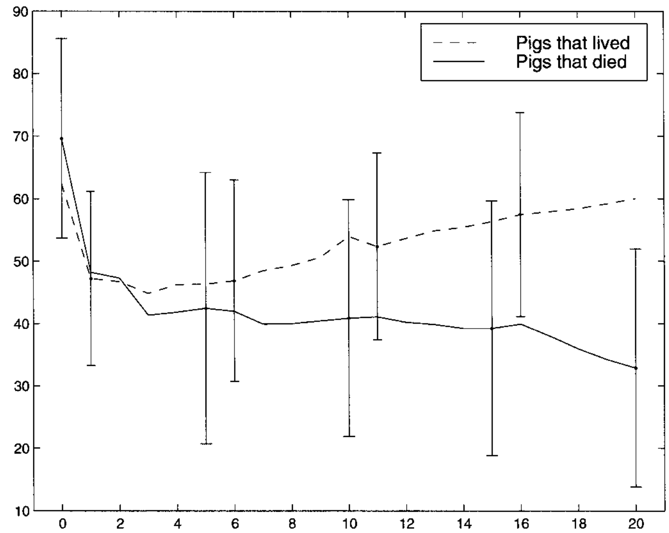

Fig. 2. *Mean MAP (mm Hg) over time (minutes) of pigs who lived up to 60 minutes and those who died before 60 minutes. The error bars represent 1 SD above and below the mean.*

Table 1 Statistical Summary of the Physiological Inputs at Postinjury Times for the Pigs Who Survived to 60 Minutes

Time (min)	Mean SBP (mm Hg) (SD)	Mean MAP (mm Hg) (SD)
1	81.22 (17.16)	62.37 (15.16)
4	63.12 (22.76)	44.84 (17.66)
16	78.60 (18.86)	56.34 (15.35)
18	80.48 (18.84)	57.92 (15.53)
20	83.23 (18.49)	60.02 (15.20)

Table 2 Statistical Summary of the Physiological Inputs at Postinjury Times for the Nonsurvivors

Time (min)	Mean SBP (mm Hg) (SD)	Mean MAP (mm Hg) (SD)
1	88.33 (17.36)	69.62 (15.96)
4	61.68 (25.55)	41.37 (21.26)
16	59.50 (21.77)	39.25 (20.41)
18	57.71 (23.71)	38.05 (21.10)
20	54.36 (21.47)	34.24 (19.07)

Table 3 Results of Static PNN Models Trained on 142 Pigs

Model at Time (min)	Discrimination Power	Area under ROC Curve	Goodness-of-Fit Statistic
16	0.732	0.774	24.115
18	0.816	0.847	15.054
20	0.857	0.886	4.81

20 minutes produces the best overall results in terms of discrimination.

Figure 2 illustrates the mean MAP of the survivors vs. nonsurvivors over time. The clear divergence seen at 3 minutes postinjury appears to suggest easy discrimination between the pigs that live and those that die. However, upon examining the standard deviations from these means, represented by the bars on the means, there is a large overlap in the MAP values of the survivors and nonsurvivors. This indicates the clinically recognized difficulty in early discrimination between survivors and nonsurvivors and highlights the need for highly sampled discrete data collection, integration, and data analysis, such as the neural network techniques described herein.

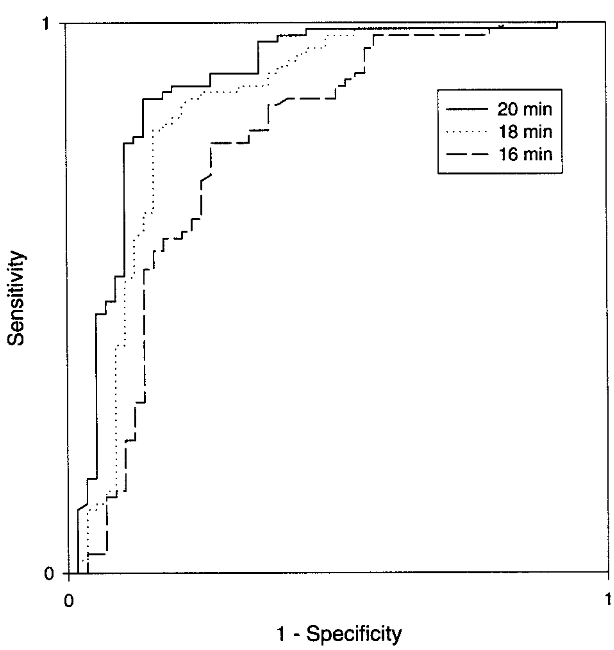

Fig. 3. *ROC curve of static PNN trained on all 142 pigs.*

Table 4 Results of Dynamic PNN Models Trained on 142 Pigs

Time (min)	Discrimination Power	Area under ROC Curve	Hosmer-Lemeshow
16	0.803	0.827	59.63
18	0.839	0.866	16.47
20	0.857	0.919	14.74

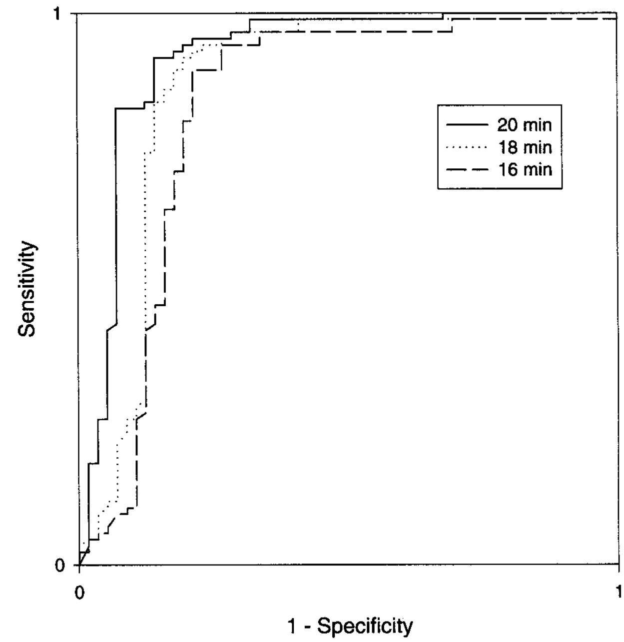

Fig. 4. *ROC curves of dynamic PNN.*

DISCUSSION

The goal of this present study was to examine the ability of PNN-based algorithms to utilize the large amounts of digitized data collected during an uncontrolled hemorrhage experiment to accurately predict mortality at the earliest possible time after injury. From our analyses, we showed that the models were successful in discriminating between those pigs that lived from those that died, at 18 and 20 minutes postinjury, where use of mean MAP could not discriminate.

The cutoff time of 20 minutes was chosen for several reasons. First, 20–30 minutes after injury is the time window within which army medics and civilian prehospital personnel must make triage and evacuation decisions.[18] In the civilian setting, 70% of all trauma deaths occur in the prehospital environment.[19] On the battlefield, roughly 90% of deaths occur before injured soldiers can be transported to the field hospital.[20,21] After arriving at the field hospital, mortality rates fall to near 3%; therefore, to significantly decrease mortality, gains are needed in the prehospital setting.[22] We feel that there is potential for improving upon the current single time point data collection in the prehospital arena by utilizing data from multiple time points, possibly decreasing the traditionally high prehospital death rate.[23] Utilizing PNNs may help facilitate triage decisions, resulting in improved utilization of resources and evacuation priorities, and leading possibly to overall lower mortality.

Limitations

There are a variety of limitations with the present study. The purpose of this retrospective study was to examine the utility of PNNs to predict mortality given time-stamped physiologic data. Although this was accomplished, there were no standard algorithms with which to compare these results. This is due to the lack of applicable data from these animal studies such as Glasgow Coma Scale score, used in most prehospital scoring systems, and Injury Severity Score, used in retrospective outcome algorithms, such as TRISS. To have some gauge of comparison, PNN models at a variety of times were synthesized. The ability of the PNN algorithm to discriminate well over the different time provides evidence that good discrimination at 20 minutes was not due to chance. The improvement in predictability of the models over 16, 18, and 20 minutes supports the hypothesis that PNNs can effectively model the physiologic dynamics of hemorrhage.

Another limitation with our results lies in the data used. The data analyzed in this retrospective study were obtained from two experiments specifically designed to assess various types of hemorrhage control and resuscitation methods, and thus the swine involved were homogenous across breed, weight, and two types of severe injury. To be useful for field triage, a model should be able to take into account more than two etiologies of life-threatening hemorrhage. Also, although statistically valid conclusions were made on models synthesized with data from 143 animals, a model synthesized with a larger training data set is needed to be useful in possibly providing an earlier mortality discrimination. Additionally, training on human data, not pig data, is needed for models to be directly useful in the clinical setting.

Creating a human prehospital database for modeling, however, would be a large undertaking, due, in part, to the amount of patients needed. For example, the RTS algorithm was constructed using a 2,166-patient database, and evaluated using a 26,000-patient database.[2,5] Collecting the necessary physiologic data needed for modeling would also be difficult due to the sophisticated sensors and storage hardware this would require. Data collected in the civilian prehospital setting, unlike data in this study, would have a significant lag from the time of injury to the time of data collection due to the response time of medics.

Another potential limitation that needs to be considered is the effect of various resuscitation endpoints, and different methods of hemorrhage control on our analysis. Although standard protocols were followed in the administration of lactated Ringer's solution for all pigs, the application of different hemostatic techniques among the liver hemorrhage pigs could have affected model discrimination.

Clinical Utility

The PNN models synthesized in this experiment indicated an ability to produce an accurate probability of survival prediction using limited data from two uncontrolled hemorrhage swine models. These models provide an option not currently available in present prehospital scoring systems to effectively harness valuable physiologic variables taken at multiple times. Unlike RTS and TRISS, which produce a probability of survival using physiologic data from one point in time, the PNN-based models combine data from multiple times and thus are able to utilize valuable information on hemorrhage dynamics to improve outcome classification.

While it is premature to discuss applications of this algorithm in civilian trauma, due to our use of data which would be difficult to collect in that setting, the idea of utilizing similar PNN-based models may soon be possible in combat trauma where efforts are being made to collect patient data immediately upon injury. This concept of data analysis at multiple times could also fuel research in the emergency department and in-patient critical care areas.

REFERENCES

1. DeMeis R. Stretching the chances of battlefield survival. *Aerospace Am.* 1995;33:22–23.
2. Champion HR, Sacco WJ, Copes WS, Gann DS, Gennarelli TA, Flanagan ME. A revised trauma score. *J Trauma.* 1989;29:623–629.
3. Roorda J, van Beek EF, Stapert JW, ten Wolde W. Evaluating performance of the revised trauma score as a triage instrument in the prehospital setting. *Injury.* 1996;27:756–757.
4. Hedges JR, Feero S, Moore B, Haver DW, Shultz B. Comparison of prehospital trauma triage instruments in a semirural population. *J Emerg Med.* 1987;5:197–207.
5. Rutledge R, Osler T, Emery S, Emery S, Kromhout-Schiro S. The end of the Injury Severity Score (ISS) and the Trauma and Injury Severity Score (TRISS): ICISS, an international classification of diseases, ninth revision-based prediction tool, outperforms both ISS and TRISS as predictors of trauma patient survival, hospital charges and hospital length of stay. *J Trauma.* 1998;44:41–49.
6. Champion HR, Sacco WJ, Hunt TK. Trauma severity scoring to predict mortality. *World J Surg.* 1983;7:4–11.
7. Wong DT, Barrow PM, Gomez M, McGuire GP. A comparison of the Acute Physiology and Chronic Health Evaluation (APACHE) II score and the Trauma-Injury Severity Score (TRISS) for outcome assessment in intensive care unit trauma patients. *Crit Care Med.* 1996;24:1642–1648.
8. Moore EE, Cogbill TH, Jurkovich GJ, Shackford SR, Malangoni MA, Champion HR. Organ injury scaling: spleen and liver (1994 revision). *J Trauma.* 1995;38:323–324.
9. Holcomb JB, Pusateri AE, Harris RA, Charles NC, Gomez RR, et al. Effect of dry fibrin sealant dressings vs gauze packing on blood loss in grade V liver injuries in resuscitated swine. *J Trauma.* 1999;46:46–54.
10. Meredith W, Rutledge R, Fakry SM, Emery S, Kromhout-Schiro S. The conundrum of the Glasgow Coma Score in intubated patients: a linear regression prediction of the Glasgow verbal score from the Glasgow eye and motor scores. *J Trauma.* 1998;44:839–844.
11. Ivaknenko AG. Polynomial theory of complex systems. *IEEE Trans Syst Man Cybernetics.* 1971;SMC-1:364–378.
12. *GNOSIS, Version 3.0*. Charlottesville, VA: Barron Associates, Inc; 1998.
13. Tenorio MF, Lee WT. *Advances in Neural Information Processing Systems*. San Mateo, CA: Morgan Kauffman; 1989.
14. Iezzoni LI. *Risk Adjustment of Measuring Healthcare Outcomes*. Chicago, IL: Health Administration Press; 1997.
15. Lemeshow S, Hosmer DW. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol.* 1982;115:92–106.
16. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. New York: John Wiley & Sons; 1989.
17. Hunt JP, Meyer AA. Predicting survival in the intensive care unit. *Curr Probl Surg.* 1997;34:529–599.

18. Reilly PM, Schwab CW, Zonies DH, McDonald G, Kauder DR, et al. Urban firearm deaths (UFD) revisited: trends over a decade [abstract]. *J Trauma.* 1998;45:1120.
19. Bellamy R, Safar P, Tisherman SA, Basford R, Bruttig SP, et al. Suspended animation for delayed resuscitation. *Crit Care Med.* 1996; 24:S24–S47.
20. Shoemaker WC, Peitzman AB, Bellamy RF, Bellomo R, Bruttig SP, et al. Resuscitation from severe hemorrhage. *Crit Care Med.* 1996; 24:S12–S22.
21. Bellamy RF. The causes of death in conventional land warfare: implication for combat casualty care research. *Mil Med.* 1984; 149:55–62.
22. Bzik KD, Bellamy RF. A note on combat casualty statistics. *Mil Med.* 1984;149:229–230.
23. Roberts DA, Brady WJ, Parker BE, Clark BR, Young JS. The use of neural networks in predicting outcome in trauma patients: a comparison of static and dynamic techniques. *Acad Emerg Med.* 1998;5:537–538.