0031-3998/04/5601-0006
PEDIATRIC RESEARCH
Vol. 56, No. 1, 2004

DOI: 10.1203/01.PDR.0000129654.02381.B9

COMMENTARY

Are Artificial Neural Networks "Ready to Use" for Decision Making in the Neonatal Intensive Care Unit?

Commentary on the article by Mueller *et al.* and page 11

CYRIL ROBIN WALKER AND MONIQUE FRIZE

Department of Paediatrics, University of Ottawa, Ottawa, Ontario, Canada K1H 8L1 [C.R.W.]; and Department of Systems and Computer Engineering, Carlton University, Ottawa, Ontario, Canada K1S 5B6 [M.F.]

An artificial neural network (ANN) is a processor of experimental knowledge used particularly in analyzing outcome estimations. ANNs are widely used outside medicine in applications for analysis (*e.g.* voice recognition), control (industrial applications), and forecasting (*e.g.* market and weather). In a process analogous to information processing by the human brain, the ANN uses interconnected pathways (Fig. 1), often including one or more "hidden layers" to allow more complex interactions to be explored, to acquire knowledge through a learning process and store the acquired information using interneuron connection strengths (synaptic weights; Fig. 2). The ANN thus is able to develop a set of outputs based on a system of input conditions. Once sufficient training experiments have been performed to minimize the error or to maximize the correct classification rate, the ANN can model the system automatically.

In this respect, the ANN serves a similar function to common statistical techniques. However, compared with a standard regression model, some users have suggested that ANNs may perform better with nonlinear relationships between input and output variables (1). Also, ANNs require no previous knowledge of the relationships between the factors under investigation. However, a review of studies on adult data sets in which ANNs were compared with statistical approaches used on the same medical data suggested that although ANNs sometimes outperform regression, both approaches have merit and probably should be explored as complementary rather than competing approaches to medical data (2) .

As in common statistical approaches, one data set is used to train the ANN, and then the model is tested on a new data set to verify its performance. An ANN can be trained to model any number of outputs from a database and can create a different algorithm appropriate to each output. Moreover, the ANN can be retrained at any time as more data become available or if the relationship between inputs and outputs changes (*e.g.* introduction of new therapies or procedures). Once trained to model a specific outcome, an ANN can predict that outcome on a "new" case within seconds of new data being entered, thus allowing a clinician immediate access to potentially valuable information. With on-line acquisition of data now available (*e.g.* physiologic data from monitors, investigation data from laboratory, PACS (picture archiving and communication system), and other hospital systems), future ANN-based systems will be able to process data in "real time" so that outcome prediction will be immediate and updated continuously. Information and even alerts will be able to be sent to clinical staff immediately as the clinical situation requires.

The literature on adult uses of ANNs in outcome prediction is now large and rapidly growing. A Medline search using only the single MeSH term "neural networks (computer)" recorded ~2000 "hits" in March 2003 but 5759 as of March 2004. There are, however, only a few previous studies using ANNs in prediction of outcomes for preterm newborns. In these, ANNs were used mainly to predict mortality and length of stay in preterm infants (3–5). Whereas length-of-stay prediction was believed to be of sufficient accuracy to use in clinical situations, the two studies of mortality prediction confirmed the need for ongoing work to improve the performance of the systems before they could be used for individual treatment decisions. All of these studies compared ANNs with regression models, and, in each, the ANN approach outperformed the statistical method but usually by only a small margin. Thus, it was again suggested that these approaches may be complementary. More recently, ANNs have been used to predict the duration of assisted ventilation in preterm infants with good results (6). However, there has been no previous attempt to use ANNs in predicting the likely success of a therapeutic intervention in newborns.

Received March 16, 2004; accepted March 29, 2004.
Correspondence: Monique Frize, Ph.D., Carleton University, Department of Systems and Computer Engineering, 1125 Colonel By Drive, Ottawa ON K1S 5B6 Canada; e-mail: mfrize@connect.carleton.ca
Natural Sciences and Engineering Research Council of Canada, Canada Foundation for Innovation, Physician Services Incorporated Foundation

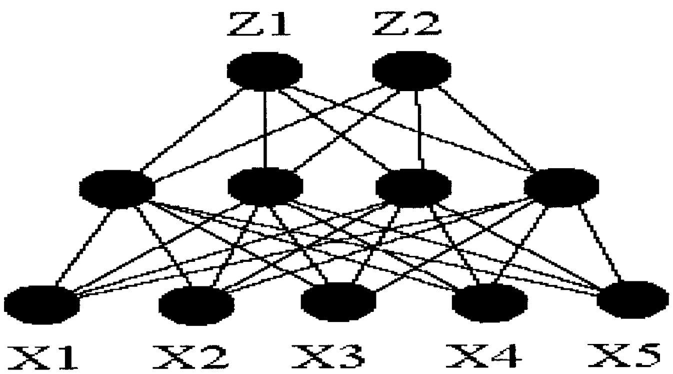

Figure 1. Architecture of a typical artificial neural network with five inputs, two outputs, and a single "hidden layer."

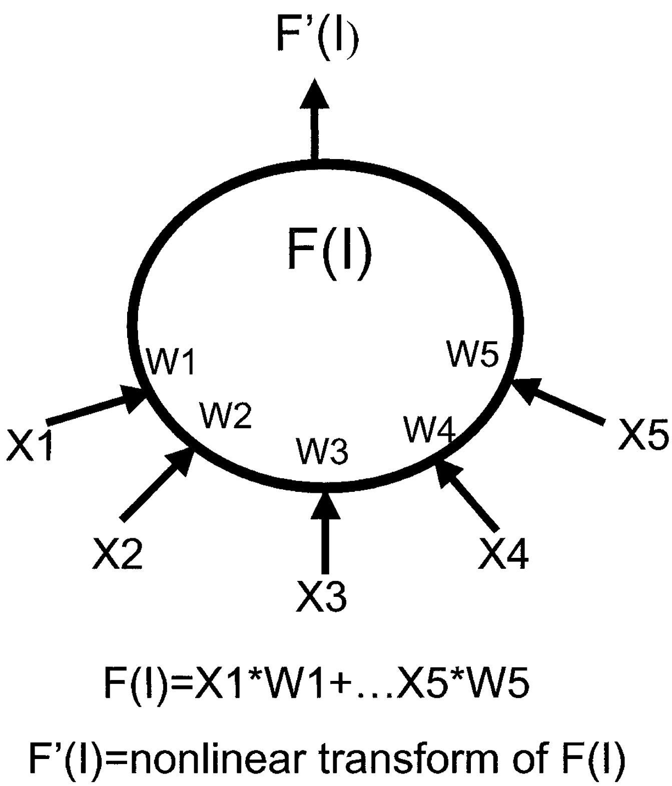

Figure 2. A typical node ("synapse") of an ANN.

Mueller *et al.* compared the performance of an ANN with a multiple logistic regression (MLR) model in predicting extubation outcome in newborns who weigh 900–1500 g (7). Mueller's approach uses ANNs to estimate outcomes for a "new" patient, based on the experience acquired with a large database of a similar population of previous patients. They also compared both systems with clinical predictions of extubation failure or success from four neonatologists who were provided with the same data sets used to develop the mathematical models. In this study, there was little difference in the ability to predict extubation success (although the ANN had marginally the best performance: 86% ANN *versus* 84% MLR and clinicians), but for prediction of extubation failure, the ANN significantly outperformed both the MLR model (86 *versus* 56%) and the clinicians (86 *versus* 41%). There was some evidence that clinical predictions were more accurate from more experienced clinicians. Several potential problems in using data for such predictions are addressed, including imputation of missing data in medical databases, a problem for both statistical and ANN approaches (8). Also, the former description of ANNs as "black boxes" is no longer valid as weights at the nodes can be extracted to allow estimates of the contribution of input variables to the final model (9,10).

"Knowledge-based" systems, including ANNs, have been studied as part of clinical decision-support systems (11–13). In some studies, such systems have improved clinician performance and patient outcomes in clinical settings through uses such as quality assurance for active medical care; however, the need for well-designed studies to assess the effects and cost-effectiveness of clinical decision-support systems has been noted, especially when attempting to affect outcomes (14). Moreover, decision-support systems (statistical or artificial intelligence based) are often poorly taken up in routine practice because they model clinical factors with little impact on treatment decisions, use model structures that lack credibility, violate well-established clinical precepts of cause-and-effect pathways, or are insufficiently validated (15). In response to this problem, Lisboa (16) surveyed randomized and nonrandomized clinical trials of ANNs in the domains of oncology, critical care, and cardiovascular medicine, including four studies in perinatal or neonatal care. This review noted the potential for extensive benefit but criticized poor methods and exaggerated claims in many studies, concluding with a blueprint for the design of complex decision systems to improve the clinical use of intelligent systems.

The potential of these systems to support or even improve decision making by the health care team (or in pediatrics perhaps also by parents) is obvious and exciting. However, although the provision of better "evidence" to support clinical, "ethical," and resource decisions seems likely to be a valuable contribution to care and decision making, without clinical trials of such systems, it cannot yet be said that this information will always lead to appropriate use or that knowledge gained through such trials will lead to beneficial clinical application. Mueller's study breaks new ground by applying ANN technology to therapeutic decision making but still focuses mainly on the performance of the system in predicting an outcome. Future work should assess not only the performance of the systems themselves but also their use in and impact on clinical practice.

REFERENCES

1. Livingstone DH, Manallack DT, Tetko IV 1997 Data modelling with neural networks: advantages and limitations. J Comput Aided Mol Des 11:135–142
2. Sargent DJ 2001 Comparison of artificial neural networks with other statistical approaches—results from medical data sets. Cancer 91(suppl S):1636–1642
3. Ambalavanan N, Carlo WA 2001 Comparison of the prediction of extremely low birth weight neonatal mortality by regression analysis and by neural networks. Early Hum Dev 65:123–137
4. Zernikow B, Holtmannspoetter K, Michel E, Pielemeier W, Hornschuh F, Westermann A, Hennecke KH 1998 Artificial neural network for risk assessment in preterm neonates. Arch Dis Child Fetal Neonatal Ed 79:F129–F134
5. Zernikow B, Holtmannspötter K, Michel E, Hornschuh F, Groote K, Hennecke KH 1999 Predicting length-of-stay in preterm neonates. Eur J Pediatr 158:59–62
6. Tong Y, Frize M, Walker R 2002 Extending successful predictions of ventilation needs using artificial neural networks from adult to neonatal ICU patients. IEEE Trans Inf Technol Biomed 6:188–191

7. Mueller M, Wagner CL, Anibale DJ, Hulsey TC, Knapp RG, Almeida JS 2004 Predicting extubation outcomes in preterm newborns: a comparison of neural networks with clinical expertise and statistical modeling. Pediatr Res (in press)
8. Ennett CM, Frize M, Walker CR 2001 Influence of missing values on artificial neural network performance. Medinfo 10(suppl):449–453
9. Walker RC, Ennett CM, Frize M 2003 Use of artificial neural networks to identify the contribution of different physiologic variables to neonatal mortality in newborn infants admitted to neonatal intensive care units. Pediatr Res 53:444A
10. Frize M, Ennett CM, Hebert P 2001 Improving the efficiency and effectiveness of artificial neural networks as decision-support systems. Medinfo 10(suppl):574
11. Yu VL, Fagan LM, Bennett SW, Clancy WJ, Scott AC, Hannigan JF, Blum RL, Buchanan BG, Cohen SN 1979 Anti-microbial selection by a computer; a blinded evaluation by infectious disease specialists. JAMA 242:1279–1282
12. Dawant BM, Uckun S, Manders EJ, Lindstrom DP 1993 The SIMON project; model-based signal acquisition, analysis and interpretation in intelligent patient monitoring. IEEE Eng Med Biol 12:82–91
13. Frize M, Ennett CM, Stevenson M, Trigg HCE 2001 Clinical decision-support systems for intensive care units using artificial neural networks. Med Eng Physics 23:217–225
14. Johnston ME, Langton KB, Haynes RB, Mathieu A 1994 Effects of computer-based clinical decision-support systems on clinician performance and patient outcome. Ann Int Med 120:135–142
15. Wyatt JC, Altman DG 1995 Commentary: prognostic models; clinically useful or quickly forgotten? BMJ 311:1539–1541
16. Lisboa PJG 2002 A review of health benefit from artificial neural networks in medical intervention. Neural Netw 15:11–39