ORIGINAL RESEARCH

Modeling short-term antidepressant responsiveness with artificial neural networks

Eugene Lin[1]
Po See Chen[2,6]
I Hui Lee[2]
Hui Hua Chang[3]
Po-Wu Gean[4]
Yen Kuang Yang[2]
Ru-Band Lu[2,5]

[1]Vita Genomics, Inc.,
Taipei, Taiwan; [2]Department
of Psychiatry, [3]Institute of
Biopharmaceutical Sciences,
[4]Department of Pharmacology,
[5]Institute of Behavioral Medicine,
[6]Department of Psychiatry, National
Cheng Kung University Hospital,
Dou-Liou Branch, Taiwan

Correspondence: Po See Chen
Department of Psychiatry, Hospital and
College of Medicine, National Cheng
Kung University, Tainan, Taiwan
Tel + 886 6-235-3535 ext 5190
Fax + 886 6-275-9259
Email chenps@mail.ncku.edu.tw

Abstract: Due to the varying nature of patient response to different types and even dosages of the same antidepressant, doctors currently prescribe antidepressants on a trial and error basis. Therefore, it is highly desirable, both clinically and economically, to establish tools that distinguish responders from non-responders and to predict possible outcomes of the antidepressant treatments. The overall effectiveness of treatment using antidepressants may thus be optimized. Common genetic polymorphisms, such as single nucleotide polymorphisms (SNPs) can be used in clinical association studies to determine the contribution of genes to drug efficacy. In this work we developed a prediction model resulting from the analysis of clinical factors such as SNPs, age, baseline Hamilton Rating Scale for Depression (HAM-D) score, antidepressant groups, and gender of depression patients. We used it to predict the responsiveness of antidepressant treatment. By using candidate genes reported in the literature, we selected four SNPs that were strongly relevant to antidepressant efficacy. Our study population consisted of Taiwanese patients with major depression recruited from the National Cheng Kung University Hospital. The genotyping data was generated in the high-throughput genomics lab of Vita Genomics, Inc. With the wrapper-based feature selection approach, we employed multilayer feedforward neural network (MFNN) and logistic regression as a basis for comparisons. Our data revealed that the MFNN models were superior to the logistic regression model. The MFNN approach provides an efficient way to develop a tool for distinguishing responders from nonresponders prior to treatments. Our preliminary results showed that the MFNN algorithm is effective for deriving models for pharmacogenomics studies and for providing the link from clinical factors such as SNPs to the responsiveness of antidepressants in clinical association studies.

Keywords: antidepressants, artificial neural networks, major depressive disorder, pharmacogenomics, single nucleotide polymorphisms

Introduction

Major depressive disorder (MDD) is a serious mental health concern that affects nearly 1 in 5 people worldwide.[1,2] Evidence from neuroscience, genetics, and clinical investigation suggests that MDD is caused by the cumulative impact of genetics, adverse psychosocial events in childhood, and ongoing or recent psychosocial stress.[3] There are a variety of antidepressants that can be used to treat MDD. However, there is little guidance on selecting which antidepressant to be most effective in a certain individual.[4,5] Moreover, there is no way to predict the response to a specific antidepressant. Therefore, models that distinguish responders from non-responders and are able to predict the possible outcome of the antidepressant treatment would be highly desirable.[4]

The efficacy of antidepressants is likely influenced by the combined effects of a number of genetic variants.[6] Accumulating evidence reveals that single nucleotide polymorphisms (SNPs) could be used as genetic markers to predict antidepressant drug treatment outcome in MDD.[7–10] Several studies in different populations[9,10] support the implication that the effects of antidepressants are associated with serotonin-related genes including guanine nucleotide binding protein beta polypeptide 3 (GNB3), 5-hydroxytryptamine receptor 1A (HTR1A; serotonin receptor 1A), 5-hydroxytryptamine receptor 2A (HTR2A; serotonin receptor 2A), and solute carrier family 6 member 4 (SLC6A4; serotonin neurotransmitter transporter). A recent study tested whether the GNB3, HTR1A, HTR2A, and SLC6A4 genes could contribute to the pharmacogenomics of short-term antidepressant response independently and/or through complex interactions in a Taiwanese population with MDD.[11] Their results support the hypothesis that GNB3, HTR2A, and SLC6A4 may play a role in short-term antidepressant treatment outcome with MDD in an interactive manner.[11]

Artificial neural network (ANN) algorithms are generally adopted for complex classification applications owing to the advantages of ANN algorithms, such as nonlinearity, fault tolerance, universality, and real-time operation.[12,13] ANN algorithms have been used to build a prediction model for the drug efficacy of interferon in chronic hepatitis C patients based on SNPs and other clinical factors.[14,15] We aim to do the same with antidepressants in patients with MDD. The possible nonlinear relationships between genetic variants and antidepressant response have already been explored using ANN algorithms with two genetic polymorphisms in a pharmacogenetic study.[16]

The previous findings[11] mainly reported significant gene–gene interactions among GNB3, HTR1A, HTR2A, and SLC6A4 in affecting short-term antidepressant treatment response. In this work, we extended the previous research to test the hypothesis that ANN algorithms with a feature selection approach could be applied to predict short-term antidepressant treatment outcomes with genetic and demographic factors.

Materials and methods
Patients
MDD patients were original to the previous study by Lin and colleagues[11] and are described in detail elsewhere.[11] All patients enrolled in this study were from the National Cheng-Kung University Hospital, Tainan, Taiwan. Blood samples were collected from 101 Taiwanese patients with MDD during the years 2005–2008. Eligibility was limited to patients with a diagnosis of MDD and the 17-item Hamilton Rating Scale for Depression (HAM-D) score ≥15.[7,9,11] Informed consent was obtained from each patient.

The study period covered the first 2 weeks of antidepressant therapy.[11] Administered drugs were venlafaxine and selective serotonin reuptake inhibitors (SSRIs). To be consistent with the previous studies,[7,9,11] patients that showed a minimum of a 50% decrease in HAM-D score after 2 weeks of treatment were defined as rapid responders to this treatment. There were 35 rapid responders and 66 non-responders.[11]

We further converted the gender data into numerical forms, that is, 1 for "male" and 0 for "female". Two antidepressant groups were formed according to types of antidepressant medication, including 1 for "SSRIs" and 0 for "venlafaxine". The baseline HAM-D score and age data were normalized between 0 and 1. In addition, we converted the clinical diagnostic data into numerical forms, that is, 1 for "rapid responder" and 0 for "nonresponder".

Genotyping
Genomic DNAs were extracted from each of the blood samples by using QIAamp DNA Blood kit according to the manufacture's instructions as described in detail elsewhere.[11]

Furthermore, genomic DNA was amplified using a commercially available AmSure SNP detection kit (Vita Genomics, Inc., Taiwan) according to the manufacture's instruction as described in detail elsewhere.[11]

Candidate genes
In the present study, we only focused on the 4 SNPs as described in Table 1.[11] As shown in Table 1, there were 4 candidate genes including the GNB3, HTR1A, HTR2A, and SLC6A4 genes. The rationale of selecting these SNPs is described in detail elsewhere.[11] Characteristics of the

Table 1 Single nucleotide polymorphisms (SNP) and gene characteristics for the four candidate genes: guanine nucleotide binding protein beta polypeptide 3 (GNB3); 5-hydroxytryptamine receptor 1A (HTR1A; serotonin receptor 1A); 5-hydroxytryptamine receptor 2A (HTR2A; serotonin receptor 2A), and solute carrier family 6 member 4 (SLC6A4; serotonin neurotransmitter transporter)

Gene	SNP	Gene characteristics
GNB3	rs5443	G-protein beta3 subunit
HTR1A	rs6295	serotonin neurotransmitter receptor
HTR2A	rs6313	serotonin neurotransmitter receptor
SLC6A4	rs25533	serotonin neurotransmitter transporter

candidate genes are related to protein coding for G-protein beta3 subunit, the serotonin receptor, and the serotonin transporter. The SNPs genetic markers of the subjects were generated at the high-throughput genomics lab of Vita Genomics, Inc.

Because there are three genotypes per locus, each SNP was coded as 0 for homozygote of the major allele, 1 for heterozygote, and 2 for homozygote of the minor allele, respectively.

Artificial neural network algorithms

In this study, we used two families of classification algorithms, including multilayer feedforward neural network (MFNN) and logistic regression as a basis for comparisons. An MFNN is one type of ANN models where connections between the units do not form a directed cycle.[17] These classifiers were performed using the Waikato Environment for Knowledge Analysis (WEKA®) software.[18] The publicly available WEKA® software is a popular suite of machine learning software that contains a collection of pattern recognition algorithms and visualization tools for data mining.[18]

From a structural point of view, an MFNN is a spatial and iterative neural network which has several layers of hidden neuron units between the input and output neuron layers.[14,15] The basis function of each neuron is the linear basis function, and a non-decreasing and differentiable sigmoid function models the activation.[19] In our approach, we employed an MFNN for modeling the responsiveness to antidepressants. Inputs contain the information about clinical factors such as SNPs, age, baseline HAM-D score, antidepressant groups, and gender for the depression patients. Outputs contain the information about the responsiveness of antidepressants.

From an algorithmic point of view, the underlying process of this MFNN can be divided into the retrieving and learning phases.[14,15] Let us assume an L-layer feedforward neural network (with N_l units at the l-th layer). In the retrieving phase, the MFNN iterates through all the layers to produce the retrieval response $\{a_i(\text{L}), i = 1, \ldots, N_\text{L}\}$ at the output layer based on the inputs of test patterns $\{a_i(0), i = 1, \ldots, N_0\}$, the known weights w_{ij} of the network, and the nonlinear activation function f_i (for example, sigmoid function). In the learning phase of this MFNN, the back-propagation algorithm[19] is employed for the learning scheme. The back-propagation algorithm is a simple gradient descent approach. The weight updating process adopts the mechanism of back-propagated corrective signals from the output layer for the hidden layers. The goal is to iteratively select a set of weights $w_{ij}(l)$ for all layers such that the squared error function E can be minimized by giving a pair of input training patterns $\{a_i(0), i = 1, \ldots, N_0\}$ and target training patterns $\{t_j, j = 1, \ldots, N_L\}$.

Mathematically, the iterative gradient descent formulation for updating each specific weight $w_{ij}(l)$ can be expressed as the following equation:

$$w_{ij}(l) \Leftarrow w_{ij}(l) - \eta \frac{\partial \text{E}}{\partial w_{ij}(l)} \tag{1}$$

where η is the learning rate and $\partial E / \partial w_{ij}(l)$ can be effectively calculated through a numerical chain rule by back-propagating the error signal from the output layer to the input layer.[14]

In summary, the MFNN is trained first by repeatedly providing input-output training pairs and executing the back-propagation learning algorithm.[14,15] After this training process, the MFNN is tested by giving the inputs of testing data (that is, clinical factors) to the network. The forward propagation of the MFNN furnishes us with the responsiveness of antidepressants for a particular patient, indicating a means of inference from cause to effect. Here, we used WEKA's default parameters, such as the learning rate = 0.3 and the momentum variable = 0.2.

Feature selection

To identify a subset of clinical factors that maximizes the performance of the prediction model, we employed the wrapper-based feature selection approach, where the feature selection algorithm acts as a wrapper around the classification algorithm.[20] In the wrapper-based feature selection approach, there is no need of knowledge of the classification algorithm, such as MFNN or logistic regression.[20] By utilizing the classification algorithm itself as part of the evaluation function for choosing feature subsets, the wrapper-based approach performs best-first search for a good subset.[21] More specifically, best-first search starts with an empty set of clinical factors and searches forward to add possible subsets of clinical factors by greedy hill-climbing augmented with a backtracking technique.[18] We applied MFNN and logistic regression with the wrapper-based approach, respectively.

Evaluation of the predictive performance

To investigate the generalization of the prediction models produced by the above algorithms, we utilized leave-one-out cross-validation.[17] One of the 101 cases was removed and the rest were used to build the model. Then, the resulting model was used to make prediction on the record set aside. This was undertaken for each of the 101 records.

To measure the performance of prediction models, we defined the accuracy as the proportion of true predicted subjects of all tested subjects.[14,15] In addition, we calculated sensitivity, the proportion of true predicted responders of all tested responders, and specificity, the proportion of true predicted nonresponders of all the tested nonresponders.

Results

The demographic and clinical characteristics of the study population were reported in our prior study (data not shown).[11] Similarly, the genotype and allele distributions of the selected four SNPs were reported in our prior study (data not shown).[11] As reported in our prior study,[11] a significant association was detected between short-term antidepressant response and the SNP rs5443 in the GNB3 gene (data not shown).

Table 2 summarizes the results of leave-one-out cross-validation experiments using the MFNN algorithm and logistic regression with the wrapper-based feature selection method. First, the input–output training data pairs were used to train the MFNN models. There were eight genetic and demographic factors, including four SNPs, age, baseline HAM-D score, antidepressant groups, and gender. Using this information, the MFNN models were trained with one to four hidden layers using the wrapper-based feature selection method. These trained MFNNs approximate the model of the responsiveness of antidepressant among depression patients. After the networks were trained, we used the trained networks to find the responsiveness condition corresponding to the testing set with the leave-one-out cross-validation method. We calculated sensitivity, specificity, and accuracy for the leave-one-out cross-validation experiments. As indicated in Table 2, the average values of accuracy for the MFNN prediction models with one to four layers were 71.3%, 70.3%, 69.3%, and 69.3%, respectively. Of all the MFNN models, MFNN with one layer performed best, outperforming the other three MFNN models. For the MFNN models with the wrapper-based approach, only 2 factors out of 8 were identified, including the SNP rs25533 in the SLC6A4 gene and baseline HAM-D score.

Next, we employed logistic regression with the wrapper-based approach for comparisons. As shown in Table 2, the average value of accuracy for the logistic regression prediction model with the wrapper-based approach was 64.4%. Among all five predictive models, the MFNN models were superior to the logistic regression model in terms of accuracy and specificity. However, the sensitivity of the logistic regression model was better than that of the MFNN models. Moreover, logistic regression with the wrapper-based approach selected 2 out of 8 factors, including the SNP rs25533 in the SLC6A4 gene and the SNP rs6313 in the HTR2A gene.

Finally, the performance of logistic regression with the same two factors as the selected MFNN (that is, the SNP rs25533 in the SLC6A4 gene and baseline HAM-D score) was an accuracy of 62.4%, a sensitivity of 0.61, and a specificity of 0.66, respectively.

Discussion

To the best of our knowledge, this is the first study that proposes to use MFNN and logistic regression with the wrapper-based feature selection method to model the drug responding status in depression patients using genetic and demographic factors. We developed a pharmacogenomics methodology to predict the drug efficacy of antidepressants in depression patients based on clinical factors such as SNPs, age, baseline HAM-D score, antidepressant groups, and gender. Our results demonstrated that a trained MFNN model is a promising method for providing the inference from clinical factors, such as SNPs and baseline HAM-D score, to the responsiveness of antidepressants. Our findings suggest that our tool may provide the medical reference prior to treatment based on the information of clinical factors such as SNP genotypes and baseline HAM-D score.

In this study, our MFNN model achieved a higher successful rate of prediction than the traditional logistic regression model. The MFNN and logistic regression models are currently the most widely used pattern recognition techniques. Unlike logistic regression, MFNN has the ability

Table 2 The result of a leave-one-out cross-validation experiment using multilayer feedforward neural network (MFNN) and logistic regression with the wrapper-based feature selection method

Algorithm	Accuracy (%)	95% Confidence interval (%)	Sensitivity	Specificity	Number of factors
MFNN with 1 hidden layer	71.3	69.4, 73.1	0.59	0.77	2
MFNN with 2 hidden layers	70.3	68.4, 72.2	0.59	0.76	2
MFNN with 3 hidden layers	69.3	67.4, 71.2	0.59	0.75	2
MFNN with 4 hidden layers	69.3	67.4, 71.2	0.59	0.75	2
Logistic regression	64.4	62.4, 66.3	0.64	0.65	2

to model the multidimensional and nonlinear relationships between the variables as found in complex medical applications.[22–24] Furthermore, the MFNN algorithms show robust performance in dealing with noisy or incomplete data.[22–24] The clinical utility of the MFNN may be restricted due to the fact that it is difficult to interpret individual variables generated by the MFNN, while logistic regression analysis provides insightful information for the interpretation of model parameters.[17,23] Thus, logistic regression can be used as a complementary method to the MFNN approach by using the same predictors.[22] In future work, we will consider alternative methods, such as decision trees, which may provide similar outcomes but in a way that is more intuitive to healthcare professionals.

In the wrapper-based approach, no knowledge of the classification algorithm is needed for the feature selection process, which finds optimal features by using the classification algorithm as part of the evaluation function.[20,21] Furthermore, the wrapper-based method has the advantage that it includes the interaction between feature subset search and the classification model.[20] However, the wrapper-based method may have a risk of over-fitting.[20,25] In a recent study, Huang and colleagues applied three classification algorithms including naive Bayes, the support vector machine algorithm, and the C4.5 decision tree algorithm with two feature selection methods to identify a subset of influential SNPs.[20] They utilized the wrapper-based feature selection method and the hybrid feature selection approach combining the chi-squared and information-gain methods. Their results suggested that the naive Bayes model with the wrapper-based approach performed maximally among predictive models to infer the disease susceptibility dealing with the complex relationship between chronic fatigue syndrome and SNPs.[20]

A similar study by Serretti and Smeraldi[16] has reported to utilize the ANN algorithms to evaluate the possible nonlinear interactions between antidepressant response and two polymorphisms in the SLC6A4 and tryptophan hydroxylase 1 (TPH1) genes. In their study, one hundred and twenty one patients with depression were treated with fluvoxamine. They reported that an ANN network with one hidden layer correctly classified 77.5% of responders and 51.2% of nonresponders (accuracy = 68.2%; sensitivity = 0.775; specificity = 0.512).[16] The specificity in their study was low because the analyzed genetic factors were not significantly associated with outcome, and further markers should also be considered in order to obtain much higher specificities.[16] The difference between our study and theirs was that our study used four SNPs and four other clinical factors such as age,

baseline HAM-D score, antidepressant groups, and gender instead of only two polymorphisms. The selected SNPs were also different in both studies. Moreover, their study did not employ the wrapper-based feature selection method. As shown in our simulation results, our MFNN prediction model performed slightly better in terms of accuracy. These preliminary results suggested that an MFNN model may be considered a good method to deal with the complex nonlinear relationship between clinical factors and the responsiveness of antidepressants.

It has been shown that the MFNN with one hidden-layer only should be adequate as universal approximations of any nonlinear functions, implying that the FFNN with one hidden-layer is always enough.[14,26] In our simulation results, we validated this implication and found that the MFNN models with two to four layers performed slightly better than the MFNN with one hidden layer in terms of accuracy. When an approximation with one hidden layer would require an impractically large number of hidden units in solving some complex real world problems, multiple hidden layers may become necessary.[14,27,28]

There were several limitations to this study as follows. First, the small size of the sample does not allow drawing definite conclusions. Second, the contributions of other genetic markers and demographic factors should be further examined. Third, the performance of our MFNN models had low sensitivities. In addition, the models are not presented in the WEKA software. Assessing the bivariate associations between each independent variable and treatment response would also be informative. Furthermore, the final models contained different dependent variables (that is, SNP rs25533 and baseline HAM-D in the MFNN model versus SNP rs25533 and SNP rs6313 in the logistic regression model). Therefore, the difference in model performance may be attributable to the fact that the wrapper-based selection technique works better for the MFNN model than it does for logistic regression. In future work, large prospective clinical trials are necessary in order to answer whether these candidate genes and clinical factors are reproducibly associated with antidepressant treatment response.

Conclusions

In this study, we developed an ANN methodology with the wrapper-based feature selection method to predict the drug efficacy of antidepressants in depression patients based on clinical factors such as SNPs, age, baseline HAM-D score, antidepressant groups, and gender. There were four SNPs selected from four candidate genes. We demonstrated that a

trained MFNN model is a promising method for providing the inference from genetic and demographic factors to the responsiveness of antidepressants.

Our findings suggested that our tool may allow patients and doctors to make more informed decisions based on clinical factors such as SNP genotyping data. Over the next few years, genetic tests for the pretreatment prediction may become a reality in patient care after prospective large clinical trials to validate clinical factors and genetic markers. It may also provide potential drug targets for the development of alternative therapeutic agents to treat depression patients, especially for those nonresponders.

Acknowledgments

Disclosure

The authors report no conflict of interest in this work.

References

1. Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA*. 2003;289(23):3095–3105.
2. Hasin DS, Goodwin RD, Stinson FS, et al. Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. *Arch Gen Psychiatry*. 2005;62(10):1097–1106.
3. aan het Rot M, Mathew SJ, Charney DS. Neurobiological mechanisms in major depressive disorder. *CMAJ*. 2009;180(3):305–313.
4. Lin E, Hwang Y, Tzeng CM. A case study of the utility of the HapMap database for pharmacogenomic haplotype analysis in the Taiwanese population. *Mol Diagn Ther*. 2006;10:367–370.
5. Lekman M, Paddock S, McMahon FJ. Pharmacogenetics of major depression: insights from level 1 of the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial. *Mol Diagn Ther*. 2008; 12(5):321–330.
6. Ising M, Lucae S, Binder EB, et al. A genomewide association study points to multiple loci that predict antidepressant drug treatment outcome in depression. *Arch Gen Psychiatry*. 2009;66(9):966–975.
7. Lin E, Chen PS, Huang LC, Hsu SY. Association study of a brain-derived neurotrophic-factor polymorphism and short-term antidepressant response in major depressive disorders. *Pharmacogenomics and Personalized Medicine*. 2008;1:1–6.
8. Lin E, Hsu SY. Gender differences and pharmacogenomics with antidepressants in depression. In: Hernandez P, Alonso S, editors. *Women and Depression*. New York, NY: Nova Science Publishers; 2009.
9. Lin E, Chen PS. Pharmacogenomics with antidepressants in the STAR*D study. *Pharmacogenomics*. 2008;9:935–946.
10. Horstmann S, Binder EB. Pharmacogenomics of antidepressant drugs. *Pharmacol Ther*. 2009;124(1):57–73.
11. Lin E, Chen PS, Chang HH, et al. Interaction of serotonin-related genes affects short-term antidepressant response in major depressive disorder. *Prog Neuropsychopharmacol Biol Psychiatry*. 2009; 33(7):1167–1172.
12. Kung SY, Hwang JN. Neural networks for intelligent multimedia processing. *Proceedings of the IEEE*. 1998;86:1244–1272.
13. Erb RJ. Introduction to backpropagation neural network computation. *Pharm Res*. 1993;10:165–170.
14. Lin E, Hwang Y, Wang SC, Gu ZJ, Chen EY. An artificial neural network approach to the drug efficacy of interferon treatments. *Pharmacogenomics*. 2006;7(7):1017–1024.
15. Lin E, Hwang Y, Chen EY. Gene-gene and gene-environment interactions in interferon therapy for chronic hepatitis C. *Pharmacogenomics*. 2007;8:1327–1335.
16. Serretti A, Smeraldi E. Neural network analysis in pharmacogenetics of mood disorders. *BMC Medical Genetics*. 2004;5:27.
17. Bishop CM. *Neural Networks for Pattern Recognition*. Oxford, UK: Clarendon Press; 1995.
18. Witten IH, Frank E. *Data Mining: Practical Machine Learning Tools and Techniques*. San Francisco, CA: Morgan Kaufmann Publishers; 2005.
19. Rumelhart DE, Hinton GE, William RJ. Learning internal representation by error propagation. Parallel Distributed Processing: Explorations. In: *The Micro-Structure of Cognition, vol.1, Foundations*. Cambridge, MA: MIT Press; 1996.
20. Huang LC, Hsu SY, Lin E. A comparison of classification methods for predicting Chronic Fatigue Syndrome based on genetic data. *J Transl Med*. 2009;7(1):81.
21. Kohavi R, John GH. Wrappers for feature subset selection. *Artificial Intelligence*. 1997;97:273–324.
22. Sargent DJ. Comparison of artificial neural networks with other statistical approaches. *Cancer*. 2001;91(8):1636–1642.
23. Eftekhar B, Mohammad K, Ardebili HE, Ghodsi M, Ketabchi E. Comparison of artificial neural network and logistic regression models for prediction of mortality in head trauma based on initial clinical data. *BMC Med Inform Deci Mak*. 2005;5:3.
24. Lin E, Hwang Y, Liang KH, Chen EY. Pattern-recognition techniques with haplotype analysis in pharmacogenomics. *Pharmacogenomics*. 2007;8(1):75–83.
25. Saeys Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics*. 2007;23:2507–2517.
26. White H. Connectionist nonparametric regression: Multilayer feedforward networks can learn arbitrary mappings. *Neural Netw*. 1990;3:535–549.
27. Sontag ED. Feedback stabilization using two-hidden-layers nets. *IEEE Trans Neural Netw*. 1992;3(6):981–990.
28. Lin CT, Lee G. *Neural Fuzzy Systems*. Upper Saddle River, NJ: Prentice-Hall; 1996.