Am. J. Hum. Genet. 62:941–949, 1998

Using Neural Networks as an Aid in the Determination of Disease Status: Comparison of Clinical Diagnosis to Neural-Network Predictions in a Pedigree with Autosomal Dominant Limb-Girdle Muscular Dystrophy

Catherine T. Falk,[1] James M. Gilchrist,[2] Margaret A. Pericak-Vance,[3] and Marcy C. Speer[3]

[1]The Lindsley F. Kimball Research Institute of The New York Blood Center, New York; [2]Rhode Island Hospital, Providence; and [3]Duke University Medical Center, Durham, NC

Summary

Studies of the genetics of certain inherited diseases require expertise in the determination of disease status even for single-locus traits. For example, in the diagnosis of autosomal dominant limb-girdle muscular dystrophy (LGMD1A), it is not always possible to make a clear-cut determination of disease, because of variability in the diagnostic criteria, age at onset, and differential presentation of disease. Mapping such diseases is greatly simplified if the data present a homogeneous genetic trait and if disease status can be reliably determined. Here, we present an approach to determination of disease status, using methods of artificial neural-network analysis. The method entails "training" an artificial neural network, with input facts (based on diagnostic criteria) and related results (based on disease diagnosis). The network contains weight factors connecting input "neurons" to output "neurons," and these connections are adjusted until the network can reliably produce the appropriate outputs for the given input facts. The trained network can be "tested" with a second set of facts, in which the outcomes are known but not provided to the network, to see how well the training has worked. The method was applied to members of a pedigree with LGMD1A, now mapped to chromosome 5q. We used diagnostic criteria and disease status to train a neural network to classify individuals as "affected" or "not affected." The trained network reproduced the disease diagnosis of all individuals of known phenotype, with 98% reliability. This approach defined an appropriate choice of clinical factors for determination of disease status. Additionally, it provided insight into disease classification of those considered to have an "unknown" phenotype on the basis of standard clinical diagnostic methods.

Received July 1, 1997; accepted for publication February 6, 1998; electronically published March 27, 1998.

Address for correspondence and reprints: Dr. Catherine T. Falk, The New York Blood Center, 310 East 67th Street, New York, NY 10021. E-mail: cathy@server.nybc.org

0002-9297/98/6204-0027$02.00

Introduction

Study of the genetics of inherited diseases usually requires expertise in the determination of disease status, even if the disease is genetically determined by alleles at a single locus. Such is the case with autosomal dominant limb-girdle muscular dystrophy (LGMD), in which many clinical features help determine disease status (Gilchrist et al. 1988). It is not always possible to make a clear determination of affected status, because of variability in the diagnostic criteria, variable age at onset, and different presentations of the disease. To map the loci of such diseases, it is important to ensure that there is a high probability that one is dealing with a homogeneous genetic trait and that disease status can be determined with a high level of reliability. One way to ensure a homogeneous genetic trait is to identify large pedigrees with many living relatives. Gilchrist et al. (1988) identified such a large, multigenerational pedigree with apparent autosomal dominant inheritance of LGMD. This family presents a rare form of LGMD, since, in most cases, the disease occurs either sporadically or as an autosomal recessive trait. The original study of this family included clinical and genetic information on 144 individuals, 115 of whom were alive. At that time 61 individuals had been examined, resulting in the diagnosis of 16 affected individuals; there were also 8 individuals who were probably affected but who had not yet been examined. Linkage analysis by Speer et al. (1992) mapped an LGMD gene, designated "LGMD1A" (MIM 159000), to chromosome 5q. In that study, the total number of individuals who had been examined had increased to 218. There were still a number of individuals for whom disease status could not be determined. An attempt was made (Speer 1993) to use pedigree-discriminant analysis (Goldin et al. 1980) to determine the relative importance of diagnostic criteria in determination of disease status, in the hope that such an analysis might aid in resolving the status of those considered to have an "unknown" phenotype; but the results were unclear.

We have now looked at an alternative approach to classification of disease status, by application of methods

of artificial neural-network analysis to the data available on the pedigree. A neural network can be thought of as a highly interconnected set of nodes or neurons that process information in parallel. Artificial neural networks are based on the current understanding of biological neural networks. In artificial networks, all "neurons" at a particular level in the network interact with all neurons at the subsequent level. By "training" such a network with a given input set of facts and related, specified "ideal" output values, the strengths (weights) of connections between neurons are adjusted until the network can, with some reliability, produce output values that are acceptably close to the ideal output values for a set of input facts. The network can then be "tested" with a second input set of facts, in which the ideal output values are again known but are not provided to the network. By comparing the generated output values with the ideal output values in this second set, we can evaluate how well the training has worked. When such a network is acceptably trained and tested, it can be used to provide output values for input sets when the "ideal" output values are not known. The basic idea is that pattern matching can be "learned" and then used to provide a reasonable output value when "similar" patterns are encountered. A more detailed presentation of neural-network structure and theory, as used in this article, can be found in the Appendix.

We have applied methods of neural networks to information on the biologically related members of the LGMD1A (MIM 159000 [http://www.ncbi.nlm.nih.gov/omim]) pedigree mentioned above for whom diagnostic criteria and disease status were available. The goal was to "train" a network to classify, with a high level of reliability, individuals as "affected" or "not affected." With a reliably trained neural network, we are able to gain insight into disease classification of those designated as having an unknown disease phenotype, as well as to offer a confirmation of the ability of a set of clinical factors to "objectively" determine disease status.

On the basis of the strategies described below, we were able to correctly classify $\geqslant 98\%$ of the individuals of "known" phenotypes, using a set of trained neural networks. The trained networks were also used to classify individuals who have an unknown disease phenotype, and those results are discussed in relation to the current linkage information in the pedigree.

Subjects, Material, and Methods

Family Data

The family data used come from the LGMD1A family (family 39) described by Gilchrist et al. (1988), Speer et al. (1992), and Yamaoka et al. (1994). A figure showing the pedigree of this family can be found in the work of Speer et al. (1992). Clinical diagnosis was made by trained clinicians. The method for determination of disease diagnosis was described by Gilchrist et al. (1988). Information was available for analysis on a total of 257 members of the pedigree, with 50 unrelated spouses (used only as a control group), 65 clinically affected individuals, 116 clinically asymptomatic individuals, and 26 individuals for whom disease diagnosis was not clear enough to make a definite determination. Eleven input facts, based on diagnostic criteria, were used in the current analysis (table 1). The diagnostic criteria, their mean values, and their ranges are given in table 1, for the individuals classified as affected, as asymptomatic ("normal"), or as having an unknown disease phenotype. The values for the 50 spouses were close to those of the asymptomatic group and are not included.

DNA haplotypes for markers flanking the LGMD1A locus (Yamaoka et al. 1994), markers D5S178 and IL9, were constructed to determine which asymptomatic individuals or individuals who have an unknown disease phenotype were likely to carry the LGMD1A gene. The defined interval spans 7 cM, and thus the probability for misclassification of disease-gene carrier state is, at most, .005, which is the probability of a double recombination event in that interval. In cases in which one or both of these markers were uninformative, markers subsequently identified that more closely define the region (e.g., markers D5S414, D5S399, and D5S2116) were used.

Neural-Network Training Strategies

The neural-network design consisted of a three-layer network: an input layer, with 11 units containing the information on the diagnostic criteria; a hidden layer, with 4 units; and an output layer, with a single numeric code indicating disease diagnosis (Falk et al. 1996). A FORTRAN program, NNTRAIN (C. T. Falk, unpublished data), accepts as input the clinical data described above, for all individuals with known disease phenotype and, together with information about disease status, trains a neural network to "learn" the patterns of input facts associated with diagnoses of affected (coded 0) and normal (asymptomatic) (coded 1). To train the network, 50 individuals were randomly selected from the set of 181 with known disease status. The training process continued until the differences between the network classifications and the clinical diagnoses became acceptably small, that is, until $\xi < \alpha$, where ξ is the mean squared error, as defined in the Appendix, and α is a predetermined, small, positive value. Once the network was trained, the remaining 131 individuals were "tested," by means of the trained network. The values generated by the neural network ranged from 0 to 1. On the basis of this value, individuals were assigned to the affected or

Table 1

Average Values, SDs, and Ranges of 11 Diagnostic Criteria Tested in the LGMD Pedigree

Disease Status and Diagnostic Criterion	Average (SD) [Range]
Affected:	
Age at onset (years)	39.108 (14.245) [10.00–75.00]
Creatine Kinase	462.538 (444.205) [38.00–2,860.00]
Dysarthria	.554 (.501) [.00–.1.00]
Heel-cord contractures	.754 (.434) [.00–1.00]
Upper-muscle strength[a]	4.588 (.523) [2.50–5.00]
Lower-muscle strength[b]	3.735 (1.222) [.20–5.00]
Triceps deep-tendon reflexes	1.185 (.813) [.00–3.00]
Biceps deep-tendon reflexes	1.415 (.855) [.00–3.00]
Brachioradialis deep-tendon reflexes	1.308 (.828) [.00–3.00]
Knee-jerk deep-tendon reflexes	.831 (.949) [.00–3.00]
Ankle-jerk deep-tendon reflexes	.246 (.626) [.00–2.50]
Asymptomatic:	
Age at last diagnosis (years)	28.759 (16.194) [6.00–74.00]
Creatine Kinase	73.224 (38.327) [30.00–318.00]
Dysarthria	.000 (.000) [.00–.00]
Heel-cord contractures	.009 (.093) [.00–1.00]
Upper-muscle strength[a]	4.991 (.036) [4.70–5.00]
Lower-muscle strength[b]	4.998 (.074) [4.30–5.00]
Triceps deep-tendon reflexes	1.737 (.538) [.50–3.00]
Biceps deep-tendon reflexes	1.772 (.576) [.00–3.00]
Brachioradialis deep-tendon reflexes	1.716 (.576) [.50–3.00]
Knee-jerk deep-tendon reflexes	1.879 (.518) [.50–3.00]
Ankle-jerk deep-tendon reflexes	1.806 (.600) [.00–3.00]
Unknown:	
Age at last diagnosis (years)	25.846 (13.016) [8.00–65.00]
Creatine Kinase	318.077 (427.987) [51.00–2,105.00]
Dysarthria	.000 (.000) [.00–.00]
Heel-cord contractures	.077 (.272) [.00–1.00]
Upper-muscle strength[a]	4.919 (.147) [4.40–5.00]
Lower-muscle strength[b]	4.908 (.200) [4.40–5.00]
Triceps deep-tendon reflexes	1.769 (.452) [.50–2.00]
Biceps deep-tendon reflexes	1.885 (.496) [.50–3.00]
Brachioradialis deep-tendon reflexes	1.846 (.485) [.50–3.00]
Knee-jerk deep-tendon reflexes	1.788 (.493) [.00–2.00]
Ankle-jerk deep-tendon reflexes	1.538 (.720) [.00–2.00]

[a] Includes scores for facial muscles, neck flexors, infraspinatus, deltoids, biceps, triceps, and wrist extensors. Muscle strength was assessed by manual muscle testing using the Medical Research Council (MRC) scale (Brooke 1986), in which "5" represents normal strength and in which "0" indicates no movement.

[b] Includes scores for iliacus, quadriceps, hamstring, tibialis anterior, and gastrocnemius. Muscle strength was assessed by manual muscle testing using the MRC scale (Brooke 1986), in which "5" represents normal strength and in which "0" indicates no movement.

normal group. Individuals whose final output values were $\leqslant .5$ were assigned to the affected group, and those whose output values were $>.5$ were assigned to the normal group. The neural-network classifications were then compared with the known clinical diagnoses, to see whether the network was able to classify disease status reliably. To account for random fluctuations in the selection of a set of 50 individuals, the process was replicated three times, with different randomly selected training sets. The results were averaged over the three replicates, and an individual's final classification was based on the averaged result. Once a network had been trained to an acceptable level of reliability, it was then used to provide estimated classifications for those of uncertain disease phenotype. The 26 individuals in the family who have an unknown disease phenotype were classified by means of the three trained networks, and the results were averaged to provide a neural-network classification for each individual.

If the data allow for reliable training of a network, it may be informative to use an alternative training strategy, in which the network is trained by use of *all* individuals with *known* phenotype. This tests the ability of the trained network to reliably separate the data into

the two classes, affected and normal. Such a trained network can no longer be "tested" on external data with known output values, but it can still be used to classify the set of "*unknowns.*" Both of the described approaches were taken, and the results are compared below.

Results

Known Disease Phenotypes

Neural-network training was performed as described above, on the basis of diagnostic criteria and disease status for the 181 individuals of known disease phenotype. In the three networks trained with 50 randomly selected individuals, the success rate in classification of the other 131 individuals of known phenotype was 91%–97%; that is, 3–11 individuals were misclassified in each separate run. When the results were averaged over the three replicates, there were only 3 misclassifications among the 181 individuals, a success rate of ~98% (fig. 1*a*). Most individuals have values close to either 0 or 1, but some are in the intermediate region, where the classification is uncertain.

If all 181 individuals are used to train the network and then are reclassified by that network, one can see whether the data set is truly separable into two classes. A reliable separation suggests that the training facts (diagnostic criteria) are, indeed, relevant to disease classification. The neural network trained in this fashion correctly classified 180 of 181 individuals (fig. 1*b*). Only three individuals, including the one who was misclassified, are in the intermediate region. The others are widely separated, clustered around the values of 0 and 1. The network has clearly separated almost all of the data points unequivocally and correctly.

By way of a control, the 50 unrelated spouses were classified by the trained neural networks. In all cases and for both training strategies, all spouses were classified as normal (data not shown).

Unknown Disease Phenotypes

The 26 individuals who could not be phenotyped by the clinicians were classified by the three trained networks by means of the first strategy described above. The three resulting values were averaged to give an average classification for each individual (table 2). For 25 of the 26 individuals who have an unknown disease phenotype, the haplotype, defined by markers flanking the LGMD1A gene, is listed in table 2; DNA was not available for the remaining individual. Those cases in which there is disagreement between the neural-network classification and the disease-haplotype designation have been indicated in a footnote to table 2; these disagreements could be either the result of a misclassification of the individual or an indication of a recombinational event. Figure 2*a* plots the average value generated by the neural network, for the 25 individuals who have an unknown disease phenotype but for whom haplotype information is available. Here, as expected, the results are not as clear-cut as they are for those of known disease status. There are six disagreements, and there are many classification values in the central region, in the .3–.7 range. This is not surprising, considering that the clinicians were unable to make clear determinations as to disease status. The classification values generated by the neural network trained on all 181 individuals of known

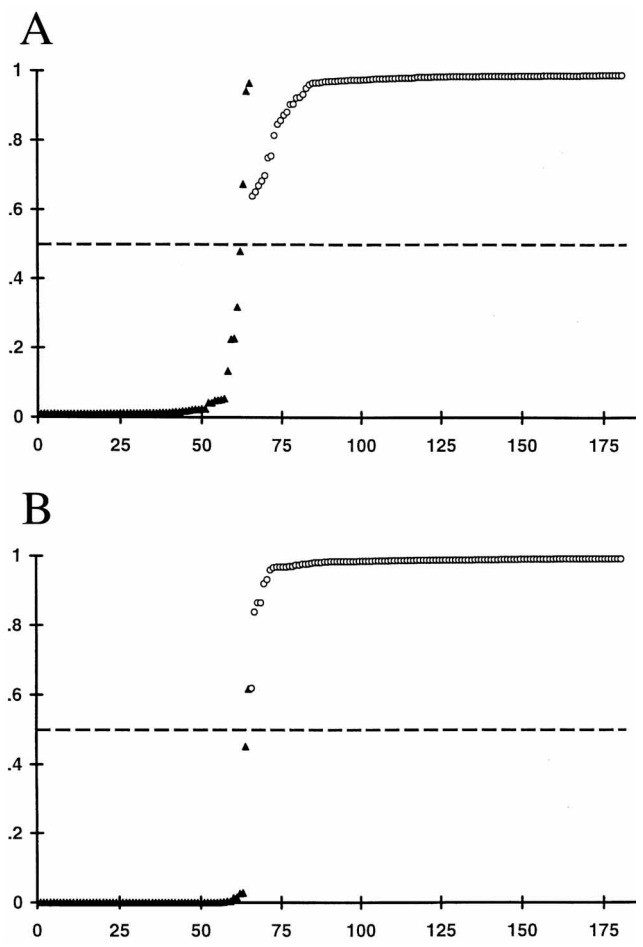

Figure 1 *A,* Neural-network classification of disease status for 181 individuals (*x*-axis) who have a known disease phenotype. Values (*y*-axis) are averaged over three trained networks and are 0–1; values ⩽.5 reflect a neural-network classification of "affected," and values >.5 reflect a neural-network classification of "normal" (asymptomatic). Individuals denoted by a blackened triangle (▲) are clinically affected, and individuals denoted by a circle (○) are clinically asymptomatic. *B,* Neural-network classification of disease status for 181 individuals who have a known disease phenotype. Values are based on a single neural network trained on all 181 individuals.

Table 2

Average Neural-Network Classification Values, Disease Designation, and Haplotype, for 26 Individuals Who Have an Unknown Disease Phenotype

Patient	Average Classification Value	Neural-Network Designation	Haplotype Designation
3	.033	Affected	Disease
5	.269	Affected	Disease
9	.342	Affected	Disease
11	.492	Affected	Disease
19	.164	Affected	Disease
20	.334	Affected	Disease
26	.325	Affected	Disease
8	.423	Affected	Normal[a]
12	.328	Affected	Normal[a]
6	.979	Normal	Unknown[b]
18	.864	Normal	Disease[a]
21	.670	Normal	Disease[a]
22	.972	Normal	Disease[a]
25	.972	Normal	Disease[a]
1	.983	Normal	Normal
2	.971	Normal	Normal
4	.952	Normal	Normal
7	.971	Normal	Normal
10	.958	Normal	Normal
13	.832	Normal	Normal
14	.536	Normal	Normal
15	.977	Normal	Normal
16	.925	Normal	Normal
17	.974	Normal	Normal
23	.911	Normal	Normal
24	.970	Normal	Normal

[a] Disagreement between neural-net designation and haplotype designation.
[b] No DNA available.

phenotype give a similar picture (fig. 2*b*). There are six disagreements between the neural-network classification and the haplotype information. The six individuals involved are the same, although their relative positions are not always the same. There are fewer intermediate values than before, and there is somewhat better separation between the classifications.

It is of interest to look at some of the characteristics of the six discrepant individuals, denoted as "8," "12," "18," "21," "22," and "25" in table 2. Of the six, four inherited the disease haplotype but were classified as normal by the neural network. In all four cases, the creatine kinase (CK) levels were slightly elevated, but all other measurements were normal. Additionally, all four are relatively young, 14–18 years of age. The average age at onset of disease in this pedigree is 39.1 years (SD = 14.2 years) (see table 1). It is likely that the disease has not yet manifested itself in these individuals and that, over time, the diagnostic features will change. In the other two cases, the individuals inherited the normal haplotype but were classified as affected by the neural network. In one case, individual 12, the CK level was 609 U/l, which is well above the range for the asymptomatic class. All other measurements were normal. Since CK levels are known to be highly influenced by certain outside factors, such as exercise, this may have been an aberrant measurement and will be retested. In the other case, individual 8, the CK level was slightly elevated, and the measurement obtained for one of the muscle tests was abnormally low. The classification of this individual is not robust. The actual numerical values obtained by use of the two neural-network strategies (i.e., averaging over three replicates and training on all individuals whose status is known) are quite different. Under the first strategy, the numerical value is .423, close

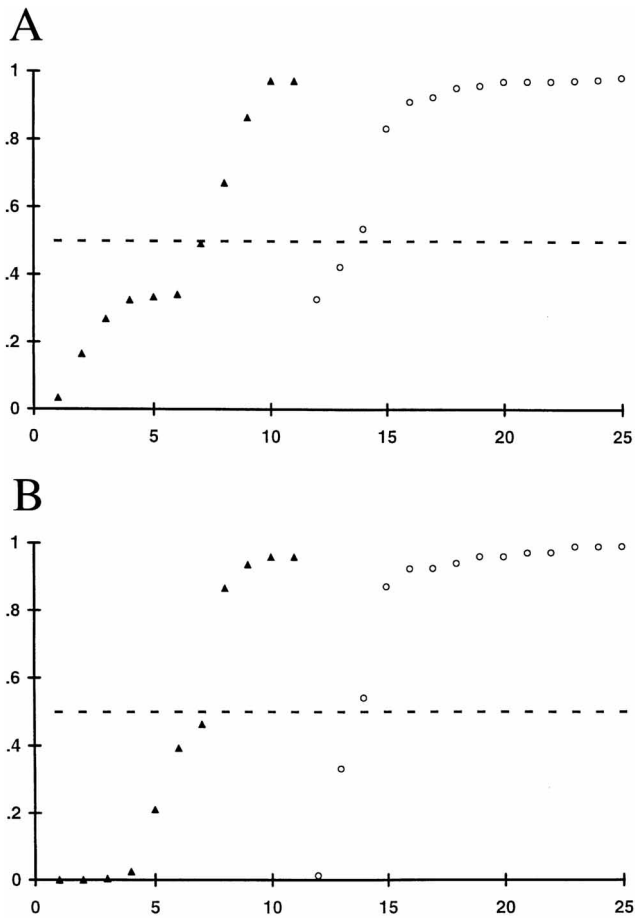

Figure 2 *A,* Neural-network classification of disease status for 25 individuals who have an unknown disease phenotype. Values are as described in the legend to figure 1. Individuals denoted by a blackened triangle (▲) inherited the disease haplotype, and individuals denoted by a circle (○) inherited the normal haplotype. *B,* Neural-network classification of disease status for 25 individuals who have an unknown disease phenotype. Values (*x*-axis) are based on a single neural network trained on all 181 individuals who have a known disease phenotype.

to the dividing point (see table 2); under the second strategy, the numerical value is .012. The results for all six of these individuals who have unknown disease status suggest that retesting of the relevant diagnostic criteria would be worthwhile.

Discussion

The determination of disease phenotype in LGMD is one example that illustrates some of the difficulties in understanding the fundamental characteristics of complex biological systems. In this case, one is presented with a set of observable characteristics that seem important in the classification of individuals into those who have the disease and those who do not. The classifying characteristics are such that it is not always possible to say with certainty whether an individual falls into one class or the other. In the case of LGMD1A, the disease phenotype is determined by a single autosomal locus, and that locus will eventually be found, thereby providing a reliable indicator of disease status. However, at this point the diagnosis of affected status is not straightforward and complicates the search for the locus. The beginning of that search, by means of linkage analysis, requires that we know who has the disease phenotype and who does not. Without this information, we must define some individuals as having an unknown disease phenotype, thus reducing the power of the analysis.

Often, one or more statistical methods are employed to help determine which characteristics are most important (and reliable) in the determination of disease phenotype. One such method, pedigree-discriminant analysis (Goldin et al. 1980), involves application of principal-component analysis to a series of clinical variables involved in disease diagnosis. The method attempts to identify a new quantitative "trait" that captures much of the variance/covariance structure of the variables. The new trait is then used in segregation analysis to search for evidence of single-locus segregation. The approach was applied to the LGMD1A data (Speer 1993), and no evidence for single-locus segregation was found, although the new quantitative trait accounted for a high proportion of the variance.

The concept of a neural network clearly has broad applications. For example, in the context of genetics, one could envision a neural network trained to recognize and discriminate between patterns of genetic transmission in a set of pedigrees. However, the task of devising and coding the appropriate set of input parameters for such a network would appear to be substantially more demanding than the task dealt with in this article. Nevertheless, such applications should become more feasible as both our knowledge of neural networks and our computational powers increase.

Neural-network methods offer a different way to use clinical information on diagnostic criteria, to aid in the reliable determination of disease phenotype. Neural networks learn to identify patterns that are important in the determination of a given output value based on a set of input parameters. In the setting up and training of a neural network, it is not necessary to identify, a priori, which of the many diagnostic criteria or input parameters are more or less important in the definition of a given output value. During the process of training, the network develops "strong" or "weak" connections that implicitly indicate which factors are important. The network does not, in its usual formulation, identify explicitly which factors are, in fact, *most* important. There are methods, such as pruning (e.g., see Halkjaer 1996; Ripley 1996), that make it possible to determine and eliminate factors that do not contribute to reliable network training. However, such methods might be of limited use if there are complex interactions between some of the factors. In the present study we chose not to attempt to determine the relative importance of the available diagnostic criteria; instead, we wanted to determine whether a neural network reliably separated individuals in the pedigree into the two classes, affected and normal (asymptomatic). Using two different strategies to train and test, we demonstrated that a neural network could be trained to perform this task. The results are useful for several reasons. First of all, they indicate that some or all of the diagnostic criteria thought to be important in determination of disease phenotype are, indeed, able to separate the two groups. Additionally, they give us a method to study the group of individuals who have been designated as having an unknown disease phenotype. The results of the neural-network classifications of these individuals help identify those cases in which additional information or reexamination may be worthwhile.

Linkage analysis has now been performed in this pedigree, and, for many individuals, it is possible to compare the neural-network classifications with the presence or absence of the haplotype thought to be carrying the disease allele. This information can be used to identify known recombinants and, thus, to map the disease locus more precisely. In the case of the 26 individuals who have an unknown disease phenotype, we cannot know for certain whether disagreement between the neural-network classification and disease-haplotype scoring results from a recombination or from an incorrect classification. However, the results could be used to give bounds to the likely location of the disease locus, as long as the uncertainty of the neural-network classification is kept in mind.

Since a major goal in the study of a genetic disease, such as LGMD1A, is to map the locus and to clone the gene, it is important to have for analysis as much reliable information as possible. It is therefore not clear exactly how the information obtained on the individuals who

have an unknown disease phenotype should be used in a linkage study, if at all. It is well known that misclassification of only a small number of individuals in a linkage study can have a major effect on the outcome of the linkage analysis (e.g., see Kelsoe et al. 1989). However, once a chromosomal region has been identified as the site for a locus, cosegregation information on closely linked markers may make it possible to look more closely at those individuals who have an unknown disease phenotype, and they may in fact become useful when one looks at the fine structure of the region. In the pedigree used in this study, for example, information is now available that has made it possible to map the disease locus and to identify two flanking markers. Those markers have localized the disease gene to a 7-cM region on chromosome 5q31-q33 (Yamaoka et al. 1994). It is therefore possible to examine in detail the characteristics of the cosegregation of alleles in the set of individuals who have an unknown disease phenotype. Comparison of information obtained from the neural-network analysis with information on the segregation of the probable disease haplotype in those individuals can help confirm disease phenotype, point to possible crossover events, and identify individuals whose clinical information may be unreliable, suggesting that they should probably be retested for the clinical diagnostic criteria.

Neural-network techniques have been applied to a variety of biological problems, with the aim of answering very different questions. These include, for example, the prediction of a drug's mechanism of action on cancer cell lines, on the basis of the drug's pattern of activity against a panel of malignant cell lines (Weinstein et al. 1992), identification of coding regions in genomic DNA sequences (Snyder and Stormo 1993), classification of human chromosomes on the basis of size, shape, and banding patterns (Sweeney et al. 1994), and diagnosis of Alzheimer disease or HIV on the basis of measurements of cerebral blood flow or activity (Halkjaer 1996). The results of these studies illustrate that neural networks can be added to the list of computational methods that may provide answers to some questions about complex biological processes. In problems in which pattern recognition and data classification can be utilized, neural-network analysis should be considered along with other methods.

Acknowledgments

C.T.F. would like to express thanks to Dr. Sara Solla, for a very enlightening conversation on many aspects of neural networks, and to Dr. H. Falk, for many helpful comments. This work was supported by NIH grants GM29177 (to C.T.F.) and NS26630 (to M.A.P.-V.). Grant support from the Muscular Dystrophy Association (to M.A.P.-V. and M.C.S.) is gratefully acknowledged.

Appendix

Neural-Network Structure and Theory

Artificial neural networks were first developed in an attempt to learn more about the functioning of the brain and how it trains itself to process information and to learn (e.g., see Hinton 1992). Very simply, in the human brain each neuron collects signals from a set of other neurons, through a series of electrical impulses that either excite or inhibit activity in the neuron. When a neuron receives such signals, it, in turn, processes those signals and sends out a resulting signal to other neurons. Learning occurs when these signals are modified so that the influence of one neuron on others changes. Artificial neural networks are composed of a set of units organized into two or more layers, working in parallel, in which units in each layer are connected, by weighted links, to all units in the subsequent layer. Each unit converts the pattern of all incoming "activity," represented by the weighted links, into a single outgoing value that is passed to all units in the next layer. By repeating these processes from one layer of the network to the next and then adjusting the weights by factors determined by the difference between a calculated and a desired output value, one attempts to train the network to provide the most reliable set of output values for given input information.

By way of a very simple example, consider the task of interpretation of the decision of a panel of three judges. That interpretation will be related to an output value for the network. The rules for determination of the output values are as follows: if two or more judges vote yes (+1), then the output is "yes" (+1); otherwise, the output is "no" (−1).

Using the step function, $\theta(x) = +1$ if $x \geqslant 0$ and $\theta(x) = -1$ if $x < 0$, we define an output function: $S = \theta(j_1 w_1 + j_2 w_2 + j_3 w_3)$, where $j_i = \pm 1$ represents the vote of the ith judge and w_i is a weight factor associated with the ith judge's vote. Thus, an output value is the S value obtained by application of the step function to the weighted sum of the judges' votes.

Consider a set of votes for the three judges: +1, +1, and −1. If we arbitrarily choose a set of weights, $w_1 = 2$, $w_2 = 1$, and $w_3 = 4$, then $(j_1 w_1 + j_2 w_2 + j_3 w_3) = -1$ and $S = \theta(-1) = -1$, leading to an erroneous output value of "no" (−1). If the three weights are $w_1 = -1$, $w_2 = 2$, $w_3 = 3$, then $(j_1 w_1 + j_2 w_2 + j_3 w_3) = -2$ and $S = \theta(-2) = -1$, again leading to an erroneous output value of "no." However, if the three weights are $w_1 = 1$, $w_2 = 1$, and $w_3 = 1$, then $(j_1 w_1 + j_2 w_2 + j_3 w_3) = +1$ and $S = \theta(+1) = +1$, leading to the output value of "yes," which satisfies the "majority" rule cited above. By adjusting the weights, we have trained the system to

arrive at the specified output values for various patterns of votes.

In this simple example we do not explicitly need the assistance of the transfer function and step function to interpret the decision of the judges, but in more complex situations the output value might not be obvious from the input values. In such cases, functions similar to those described can be used, along with rules for testing and adjusting the weights, to obtain an optimum set of weights that transform any set of input values into an appropriate output value.

Figure A1*a* shows, schematically, a diagram of a very simple neural network. It consists of an input layer with three "units" and an output layer with one unit. The two layers are connected in such a way that each input unit contributes to the value of the output unit, where that contribution is determined by both an input value, say x_i, and a corresponding weight value, w_i. Thus, for a given set of n input parameters and weights, the output value y is calculated as $y = \Sigma_{i=1}^{n} w_i x_i$.

The objective is to modify the weights in such a way that the calculated value of y is very close to the desired output value, d. If we have only one set of input/output values, this is a relatively simple task, but, if we have a large set of input vectors with a corresponding set of desired output values, the task becomes more complicated, since we now wish to find one set of weight values that will provide reliable estimates of the output values for *all* of the sets of input values.

Say that, for the kth input set we have a vector $\mathbf{X}_k$, $k = 1,2,...,m$, where $\mathbf{X}_k$ has components $x_{1k},...,x_{nk}$ containing information on factors relevant for training a neural network. We can classify an output parameter, y_k, for each input set as a weighted sum of the input parameters. Then, $y_k = \Sigma_{i=1}^{n} w_i x_{ik}$, where the weight vector, $\mathbf{W}$, with components, w_i (taken independent of k), represents the weights for parameters x_{ik}. Now consider a set of such input vectors, $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_m\}$, each with its own "correct" output value, d_k, $k = 1,2,...,m$. We wish to find a set of weights that enables us to most closely estimate the correct output values. We do this by training the neural network in such a way as to minimize the differences between the estimated output values, y_k, and the correct values, d_k, for all k.

Let $\varepsilon_k = d_k - y_k$ represent the error between the true and the estimated values for each item. The mean squared error, ξ, is then $\xi = \frac{1}{m}\Sigma_{k=1}^{m} \varepsilon_k^2$.

We now want to find the weight vector, $\mathbf{W}$, that will minimize ξ. This is done by means of a minimization method to find the minimum of the error surface. Details of such minimization methods can be found in the work of Hertz et al. (1991), pp. 115–120, for example. This formulation of a simple neural network is mathematically equivalent to linear discriminant analysis (Bishop 1995).

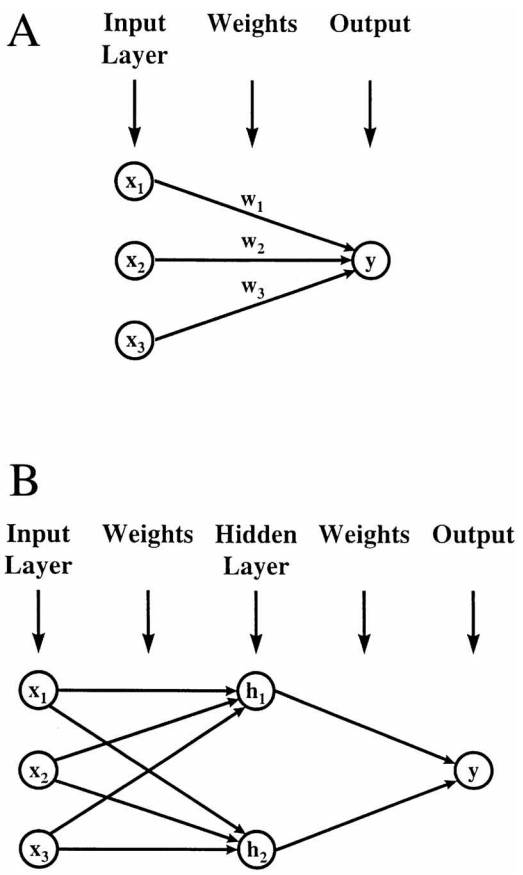

Figure A1 *a*, Illustration of a simple two-layer neural network with three input units, x_i. The input units are connected to the single output value, y, by means of the weights, w_i. *b*, Illustration of a three-layer neural network. Each input layer, x_i, is connected to each hidden unit, h_j, by the six weights, w_{ij}, and each hidden unit is, in turn, connected to the output value, y, by the two weights, w_j. Note that, for simplicity, the weights are not explicitly labeled.

A simple neural network, such as the one described here, does not always provide a suitable set of weight parameters to estimate the output values with an acceptable level of accuracy. Several extensions of this simple procedure have been developed, and they generally improve the ability of a network to be trained. One useful extension is that of using a transformation function between the input and output values. The example represented by fig. A1*a* can be said to have a linear output function, in which the output value is simply a linear combination of the input values and the corresponding weight values. More-useful predictive characteristics are obtained with a nonlinear output function, such as the step function, used above, or the sigmoid function, which we now will introduce and use. Our output value, y_k, would then be estimated from the sigmoid function, as $y_k = 1/(1 + e^{-S_k})$, where $s_k = \Sigma_{i=1}^{n} w_i x_{ik}$. The sigmoid function for large positive s_k

yields $y_k \sim 1$ and, for very negative s_k yields $y_k \sim 0$. Thus, the sigmoid function has the form of a smoothed step function.

A second extension of the simple model is to add an additional layer of units between the input and output layers. This additional layer is called a "hidden layer," since the status of the units in this layer cannot be inspected from the "outside." Figure A1*b* shows a schematic of a neural network that contains the addition of a hidden layer with two units. (Note that subscripts have been simplified in this figure, as they were in figure A1*a*, and that the figures should, therefore, be interpreted as schematic representations of the corresponding networks.). With a three-layer neural network, we have two sets of weights, one connecting the input values, x_i, with the hidden values (here depicted by h_k), and a second set, connecting the hidden values with an output value. As before, each input value contributes to each hidden value, and each hidden value contributes to the output value.

The addition of a hidden layer provides for more-efficient and reliable training of a neural network, but it also creates a new problem when we try to update the weights that go from the hidden-layer units to the output layer. We have no way of knowing the "true" values of the hidden-layer units, h_i, for a given input layer. We therefore have to make use of the fact that the actual output values depend on the hidden-layer weights; that is, these weights form part of the calculation of the hidden-layer outputs, which are then used in the calculation of the output-layer value. We can work "backwards" by determining the "local gradient of the error surface with respect to the hidden-layer weights and use this value to update the weights" (Freeman 1994, p. 64). This creates what has been called a "back propagation network" (also see Hertz et al. 1991). We calculate errors on the output layer first and bring these errors back to the hidden layer, to calculate error-surface gradients at that level. (It should be noted that the use of other error functions for neural-network learning have been studied [Solla et al. 1988].)

We perform this sequence of steps over and over again, using information from a set of input/output values, randomly selecting, at each iteration, a particular input/output combination. We continue this process until the estimated weights provide a set of output values that are *all* acceptably close to the known output values; that is, until the mean squared error, ξ, is less than some predetermined small number. Once the network is trained, the weights are fixed; and the network with those weights can be used to calculate output values for any given set of input parameters.

References

Bishop CM (1995) Neural networks for pattern recognition. Oxford University Press, New York

Brooke MH (1986) A clinician's view of neuromuscular disease. Williams & Wilkins, Baltimore

Falk CT, Gilchrist JM, Pericak-Vance MA, Speer MC (1996) Using neural networks as an aid in determining disease state: comparison of clinical diagnosis with computer predictions in a pedigree with autosomal dominant limb girdle muscular dystrophy. Am J Hum Genet Suppl 59:A178

Freeman JA (1994) Simulating neural networks. Addison-Wesley, Reading, MA

Gilchrist JM, Pericak-Vance M, Silverman L, Roses AD (1988) Clinical and genetic investigation in autosomal dominant limb-girdle muscular dystrophy. Neurology 38:5–9

Goldin LR, Elston RC, Graham JB, Miller CH (1980) Genetic analysis of von Willebrand's disease in two large pedigrees: a multivariate approach. Am J Med Genet 6:279–293

Halkjaer S (1996) Dynamics of learning in neural networks: application to the diagnosis of HIV and Alzheimer patients. Cand Scient thesis, University of Copenhagen, Copenhagen

Hertz J, Krogh A, Palmer R (1991) Introduction to the theory of neural computation. Addison-Wesley, Redwood City, CA

Hinton GE (1992) How neural networks learn from experience. Sci Am 267: 144–151

Kelsoe JR, Ginns, EI, Egeland JA, Gerhard DS, Goldstein AM, Bale SJ, Pauls DL, et al (1989) Re-evaluation of the linkage relationship between chromosome 11p loci and the gene for bipolar affective disorder in the Old Order Amish. Nature 342:238–243

Ripley BD (1996) Pattern recognition and neural networks. Cambridge University Press, Cambridge

Snyder EE, Stormo GD (1993) Identification of coding regions in genomic DNA sequences: an application of dynamic programming and neural networks. Nucleic Acids Res 21: 607–613

Solla SA, Levin E, Fleisher M (1988) Accelerated learning in layered neural networks. Complex Syst 2:625–640

Speer MC (1993) Genetic studies in limb-girdle muscular dystrophy. PhD thesis, Duke University, Durham, NC

Speer MC, Yamaoka LH, Gilchrist JH, Gaskell CP, Stajich JM, Vance JM, Kazantsev A, et al (1992) Confirmation of genetic heterogeneity in limb-girdle muscular dystrophy: linkage of an autosomal dominant form to chromosome 5q. Am J Hum Genet 50:1211–1217

Sweeney WP, Musavi MT, Guidi JN (1994) Classification of chromosomes using a probabilistic neural network. Cytometry 16:17–24

Weinstein JN, Kohn KW, Grever MR, Viswanadhan VN, Rubinstein LV, Monks AP, Scudiero DA, et al (1992) Neural computing in cancer drug development: predicting mechanism of action. Science 258:447–451

Yamaoka LH, Westbrook CA, Speer MC, Gilchrist JH, Jabs EW, Schweins EG, Stajich JM, et al (1994) Development of a microsatellite genetic map spanning 5q31-q33 and subsequent placement of the LGMD1A locus between D5S178 and IL9. Neuromusc Disord 4:471–475