JP Journal of Biostatistics
Volume 7, Number 2, 2012, Pages 61-76
Published Online: November 2012
Available online at http://pphmj.com/journals/jpb.htm
Published by Pushpa Publishing House, Allahabad, INDIA

USING A REDUCED TOPOLOGY OF AN ARTIFICIAL NEURAL NETWORK TO IDENTIFY BIOLOGICAL INTERACTION IN GENETIC EPIDEMIOLOGY FOR TWO-LOCUS DISEASE MODELS

Frauke Günther[1,*], **Karin Bammann**[1,2] **and Iris Pigeot**[1]

[1]BIPS - Institute for Epidemiology and Prevention Research GmbH
Achterstraße 30, 28359 Bremen, Germany
e-mail: guenther@bips.uni-bremen.de
bammann@uni-bremen.de
pigeot@bips.uni-bremen.de

[2]Institute of Public Health and Nursing Science (IPP)
University of Bremen
Grazer Straße 4, 28359 Bremen, Germany

Abstract

Many diseases with genetic background are not just caused by single genes, but by a complex interplay between several genes and environmental factors. Thus, the investigation of their interactions gains more and more importance. However, statistical modeling of gene-gene interactions is a challenge. For example, regression models do not fully capture biological interaction. We may look for statistical approaches that offer more flexibility. Artificial neural networks do

2010 Mathematics Subject Classification: 62J12, 62P10, 92B20, 92D30.

Keywords and phrases: artificial neural networks, biological independence, gene-gene interaction, two-locus disease model.

[*]Corresponding author

Received June 13, 2012

not depend on pre-specified model structures and may thus be more appropriate to identify biological independence or interaction. Our approach is based on the idea that neural networks with reduced topology, i.e., a topology where the independent loci are not connected, should be able to reflect a biological independence model. Thus, we compare the model fits of two neural networks one with a reduced and one with a fully connected topology which should allow to decide on the presence of biological interactions. We perform a simulation study to investigate whether our approach leads to satisfactory results assuming different biological models of independence and interactions. It can be concluded that obviously similar problems as with standard regression models occur such that our approach in its present form is not able to identify biological independence.

1. Introduction

Complex diseases are caused by an interplay between various genes and environmental factors. Their investigation gains increasing attention in genetic epidemiology and requires that, besides main effects, interactions are considered for a better understanding of the effect of the covariates. Here, we face the problem that statistical interactions are not able to fully capture the complexity of biological interactions [9, 20]. Statistical interaction is usually defined as deviation from an additive effect of the covariates on the untransformed or transformed outcome. For instance in logistic regression models, an additive effect on the logit-transformed outcome means a multiplicative effect on the untransformed outcome and statistical interaction can be interpreted as deviation from a multiplicative effect. In contrast, biological interaction is presented if one gene changes the effect of another gene [8]. Thus, several statistical approaches have been proposed to model and to identify biological gene-gene interactions such as for instance support-vector machines [6], random forests [2, 4], multi-factor dimensionality reduction (MDR [24]), combinatorial partitioning methods [19], focused interaction testing framework [16], classification and regression trees (CART [7]), logic regression [26], and lasso regression [28]. Nevertheless, none of

these methods can be considered as standard which is in part due to the fact that none of the approaches seems to be adequate for all two-locus disease models (see, e.g., [5, 13, 14, 17]). Despite this obvious problem caused by non-flexible modeling approaches, appropriate statistical methods are needed and still looked for to identify biological interaction for a meaningful interpretation of genetic studies [18, 29].

Artificial neural networks offer a great flexibility to model any functional relationship between response variables and covariates. In addition, it is not necessary to prespecify the model structure in advance. It has been shown that neural networks are able to reflect various types of biological models that are typically used to capture biological gene-gene interactions [11]. Thus, the advantages of neural networks may be further exploited to identify gene-gene interactions by especially accounting for the topology of neural networks. For this purpose, we start from the biological two-locus disease model representing biological independence, the so-called additive model by Risch [23]. We assume that the additive model requires less parameters to be represented by a neural network. To be more specific, it might be expected that if the two loci are indeed independent there are no synapses connecting these two loci in the fitted neural network. After having fitted a network with a complete topology and one with a reduced topology, i.e., without the connecting synapses, a comparison of both model fits may enable the researcher to decide whether the reduced topology is sufficient to capture the association structure of the underlying genetic data and thus to decide on the presence of biological interactions. This means that the trained neural network would then allow to distinguish between biological independence and biological interaction.

The paper is organized as follows: Section 2 gives a brief overview of the methods used in this paper: Subsection 2.1 summarizes biological models for two loci representing biological independence and different types of biological interaction. Subsection 2.2 introduces artificial neural networks without dwelling into technical details. In addition, our approach to compare two different fitted neural networks in order to identify biological gene-gene

interactions is described in somewhat more detail in Subsection 2.3. To investigate whether our approach is successful we perform a simulation study where the design and its results are presented in Section 3. Finally, Section 4 gives a critical discussion of our approach.

2. Methods

2.1. Two-locus disease models

In genetic epidemiology, two-locus models that are characterized by penetrance tables have been mainly used to model biological independence and interaction [8, 27]. The so-called additive model introduced by Risch [23] represents biological independence of two biallelic loci. Let Y denote the case-control status, where $Y = 1$ indicates a case, i.e., a person having the disease of interest, and $Y = 0$ indicates a control, i.e., a person without the disease of interest, and let G_A as well as G_B denote the genotype of the two biallelic candidate loci A and B. G_A and G_B take 0, 1, 2 as possible values, i.e., the genotypes count the number of alleles at risk for both loci. Each two-locus disease model is defined by a special structure of the so-called penetrance matrix $f = (f_{ij})_{i,j}$, where the penetrances are defined as conditional probability of being a case $(Y = 1)$ given the joint genotype $G_A = i$ and $G_B = j$, $i, j \in \{0, 1, 2\}$. For the additive model, it is assumed that the penetrances can be represented as sum of the so-called penetrance summands a_i and b_j:

$$f = (f_{ij})_{i,j} = (P(Y = 1 | G_A = i, G_B = j))_{i,j}$$
$$= (a_i + b_j)_{i,j}, \quad i, j \in \{0, 1, 2\}. \tag{1}$$

To ensure that $0 \le f_{ij} \le 1$, the penetrance summands have to fulfill $0 \le a_i, b_j \le 1$ and $0 \le a_i + b_j \le 1$.

To distinguish the additive independence model from interaction models, we investigate two disease models that represent biological gene-gene

interaction. For the multiplicative disease model, the penetrances are given as product of so-called penetrance factors a_i and b_j:

$$f = (f_{ij})_{i,j} = (P(Y = 1 | G_A = i, G_B = j))_{i,j}$$
$$= (a_i \cdot b_j)_{i,j}, \quad i, j \in \{0, 1, 2\}. \tag{2}$$

The second model is an epistatic model. Epistatic models provide a very flexible framework for two-locus disease models [15]. We consider the recessive epistatic model given by

$$f = (f_{ij})_{i,j} = (P(Y = 1 | G_A = i, G_B = j))_{i,j}$$
$$= \begin{pmatrix} c & c & c \\ c & c & c \\ c & c & rc \end{pmatrix}, \quad i, j \in \{0, 1, 2\}, \tag{3}$$

where c denotes a baseline risk and r is a risk increase or decrease. There is only a risk change if both loci carry two alleles of risk.

2.2. Artificial neural networks

We choose a feed-forward multilayer perceptron (MLP [3]) as artificial neural network. The underlying structure of an MLP is a directed and weighted graph and consists of vertices (neurons) that are organized in layers and edges (synapses). A synaptic weight is attached to each synapse indicating the effect of the related neuron. The input layer consists of all covariates and each covariate is represented by a separate neuron. The response variables are located in the output layer. The organization of neurons, synapses and weights is called *network topology*. In the special case of only one input layer and one output layer with one response variable, an MLP calculates the following function:

$$\mu(\mathbf{x}) = \sigma\left(w_0 + \sum_i w_i \cdot x_i \right) = \sigma(w_0 + \mathbf{w}^T \cdot \mathbf{x}), \tag{4}$$

where w_0 denotes the intercept, $\mathbf{w} = (w_1, ..., w_n)$ the vector consisting of all

synaptic weights without intercept, and $\mathbf{x} = (x_1, ..., x_n)$ the vector of all covariates. The activation function σ regulates the output of the neural network. If the response variable is binary, σ should be chosen as the logistic function, so that the output is converted to the interval [0, 1]. In this case, the neural network is equivalent to the logistic regression model and all trained weights are identical to the estimates of the regression parameters. In particular, the output $\mu(\mathbf{x})$ of the neural network can also be interpreted as conditional probability of being a case given the covariates **x**.

So-called hidden layers of neurons can be included to increase the flexibility of the modeled function. Subsequent layers are fully connected, i.e., each neuron is connected by a synapse to each neuron of the following layer. However, one hidden layer is sufficient to model any piecewise continuous function [12]. The related MLP calculates the following function:

$$\begin{aligned}\mu(\mathbf{x}) &= \sigma\left(w_0 + \sum_j w_j \cdot \sigma\left(w_{0j} + \sum_i w_{ij} \cdot x_i \right) \right) \\ &= \sigma\left(w_0 + \sum_j w_j \cdot \sigma(w_{0j} + \mathbf{w_j}^T \cdot \mathbf{x}) \right),\end{aligned} \tag{5}$$

where w_0 denotes the intercept of the output neuron and w_{0j} the intercept of the *j*th hidden neuron. In addition, w_j denotes the synaptic weight corresponding to the synapse starting at the *j*th hidden neuron and leading to the output neuron, $\mathbf{w_j} = (w_{1j}, ..., w_{nj})$ the vector of all synaptic weights corresponding to the synapses leading to the *j*th hidden neuron, and $\mathbf{x} = (x_1, ..., x_n)$ the vector of all covariates. This MLP calculates a weighted and transformed sum of all incoming signals twice: first, the covariates are weighted, added up and transformed by σ at each hidden neuron; second, this is done again for the resulting sums at the output neuron. Therefore, artificial neural networks can be regarded as direct generalizations of generalized linear models (GLM). Although the weights cannot be interpreted as regression parameters of a GLM, the output $\mu(\mathbf{x})$ can still be interpreted as

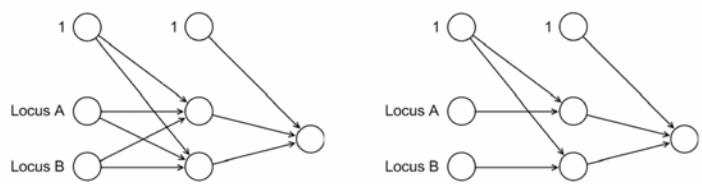
conditional probability of being a case given **x**, if σ is chosen as logistic function.

Artificial neural networks are fitted to the underlying data by a learning algorithm. By adjusting all synaptic weights, this algorithm minimizes an error function that for each observation depends on the given response variable and the related output of the neural network. In this paper, the resilient backpropagation algorithm is used that is based on an iterative gradient descent [22]. In the case of binary output, cross-entropy should be used as error function to ensure equivalence to maximum likelihood estimation. In addition, this allows that criteria based on the likelihood can easily be calculated as transformations of the error function.

2.3. Approach for identifying biological interaction

Starting from the additive model by Risch, we pursue the aim to identify biological independence from a fitted neural network with a reduced topology. This implies that two neural networks have to be fitted to a data set and their model fits have to be compared where we assume that a reduced topology is sufficient to represent the additive independence model (see Figure 1). Both network topologies consist of three input neurons - two loci and one constant variable belonging to the intercepts of the hidden layer –, one output neuron and one hidden layer with two hidden neurons and one constant variable belonging to the intercept of the output neuron. On the one hand, neural networks with complete topology are fitted to the data (see Figure 1(a)), where subsequent layers are fully connected. On the other hand, a reduced topology is assumed, where two synapses are deleted, namely the two synapses that allow interconnections between the two involved loci (see Figure 1(b)). The idea behind is that the flexibility to model statistical associations should be in general higher for a neural network with complete topology than for one with reduced topology due to the higher number of parameters. Thus, it is expected that the model fit is in general better for a neural network with complete topology. Given that the additive model is a disease model of biological independence, it is of interest to learn whether the reduced topology without interconnections is as powerful as the fully

connected network to describe this disease model. If this were the case, complete and reduced topology would lead to identical model fits and the additive model could be characterized based on missing interconnections, i.e., based on an identical model fit to the underlying data by a neural network with reduced topology.



(a) Complete topology (b) Reduced topology

Figure 1. Comparison of both network topologies.

Following the notation introduced in equations (4) and (5), the neural network with complete topology calculates

$$\mu(\mathbf{x}) = \sigma\left(w_0 + \sum_{j=1}^{2} w_j \cdot \sigma\left(w_{0j} + \sum_{i=1}^{2} w_{ij} \cdot x_i \right) \right), \tag{6}$$

where the one with reduced topology calculates

$$\mu(\mathbf{x}) = \sigma\left(w_0 + \sum_{j=1}^{2} w_j \cdot \sigma(w_{0j} + w_{jj} \cdot x_j) \right). \tag{7}$$

3. Simulation Study

3.1. Design

The proposed idea to identify biological independence by an nearly identical model fit of the reduced topology is investigated in a simulation study where first the design of the study is chosen in a very idealized way to check whether the new approach is able to detect the biological independence in this optimal situation. If the approach is successful, then more realistic designs will be considered.

Neural networks with both topologies are fitted to simulated data that represent the two-locus disease models introduced in Subsection 2.1. All samples are generated via a two-step procedure. In a first step, a population is generated that consists of 1,000,000 observations. This population contains the genetic information of two marginally independent and biallelic loci, which are coded by the genotype, and a related case-control status. Both loci have a minor allele frequency (MAF) of 30% to ensure a sufficient frequency of all genotype combinations. The case-control status is randomly allocated with the probabilities of the given penetrance matrix *f*. The penetrance matrix is obtained from the three models given in equations (1)-(3). The following two risk scenarios concerning the genotype relative risks are considered for the additive and the multiplicative model

$$\begin{aligned} &\text{Scenario 1: } a_1 = 2 \cdot a_0 && \text{Scenario } a_1 = 5 \cdot a_0 \\ &\qquad\quad a_2 = 4 \cdot a_0 && \qquad\quad a_2 = 10 \cdot a_0 \\ &\qquad\quad b_1 = 5 \cdot b_0 && \qquad\quad b_1 = 5 \cdot b_0 \\ &\qquad\quad b_2 = 10 \cdot b_0 && \qquad\quad b_2 = 10 \cdot b_0. \end{aligned}$$

This yields, for instance, the following penetrance matrix for the additive model as introduced in equation (1) in the first risk scenario

$$\begin{aligned} f = (f_{ij})_{i,j} &= \begin{pmatrix} a_0 + b_0 & a_0 + b_1 & a_0 + b_2 \\ a_1 + b_0 & a_1 + b_1 & a_1 + b_2 \\ a_2 + b_0 & a_2 + b_1 & a_2 + b_2 \end{pmatrix} \\ &= \begin{pmatrix} a_0 + b_0 & a_0 + 5 \cdot b_0 & a_0 + 10 \cdot b_0 \\ 2 \cdot a_0 + b_0 & 2 \cdot a_0 + 5 \cdot b_0 & 2 \cdot a_0 + 10 \cdot b_0 \\ 4 \cdot a_0 + b_0 & 4 \cdot a_0 + 5 \cdot b_0 & 4 \cdot a_0 + 10 \cdot b_0 \end{pmatrix}. \end{aligned} \tag{8}$$

The epistatic model is considered in two situations with $r = 5$ (scenario 1) and $r = 10$ (scenario 2). The penetrance summands a_0 and b_0, the penetrance factors a_0 and b_0 as well as the baseline risk *c* in the epistatic model are determined such that the prevalence in the population is equal to

1%. As already mentioned, we first consider such risk scenarios with high genotype relative risks to investigate whether our approach works in extreme and well distinguishable situations.

In a second step, case-control samples with 1,000 cases and 1,000 controls are randomly drawn from the simulated population. Neural networks with both network topologies are fitted to the data with starting weights randomly drawn from a standard normal distribution. Ten repetitions are calculated for both network topologies to enhance the chance to find a global minimum of the error function. For both network topologies, the best model is selected based on Akaike's information criterion (AIC [1]). In addition, we fit neural networks with identical starting weights for both topologies to ensure that the results are not affected by the choice of the starting weights, i.e., instead of twenty sets of starting weights — ten for the complete and ten for the reduced topology — only ten are drawn and used for both network topologies. The two best models — one for each network topology — are again selected based on the AIC. Resilient backpropagation is appropriate to train neural networks with both network topologies since it is mainly based on the chain rule for differentiating composite functions [25] and missing edges can be completely ignored. Thus, the learning algorithm does not affect our results. For each situation, one hundred repetitions are calculated and the mean output is determined.

The population and case-control samples differ in all characteristic probabilities like penetrance matrix and allele frequencies due to the different prevalences. Therefore, it is necessary to determine a theoretical penetrance matrix of the case-control sample as [11]. The theoretical penetrance matrix corresponds to the penetrance matrix of a perfect sample and can be used to assess the quality of the model fits. For this purpose, the mean output of the trained neural networks as an estimate of the penetrance matrix of the case-control sample is compared to the theoretical penetrance matrix. The deviation (**D**) between the mean estimated and the theoretical penetrances provides a measure for the quality of the average model fit:

$$\mathbf{D} = (D_{ij})_{i,j} = (|\, f_{ij}^{s} - \hat{f}_{ij}^{s} \,|)_{i,j}$$

$$= (|\, f_{ij}^{s} - \mu(x_A = i,\, x_B = j) \,|)_{i,j}, \tag{9}$$

$i, j \in \{0, 1, 2\}$, where f_{ij}^{s}, $i, j \in \{0, 1, 2\}$, denote the theoretical penetrances for $x_A = i$ and $x_B = j$ and $\hat{f}_{ij}^{s} = \mu(x_A = i, x_B = j)$ denote the related mean output of the neural network, i.e., the mean estimation of the conditional probability of being a case given **x**. Obviously, the smaller the values of D_{ij} are, the better the model fit is and a value of zero indicates a perfect model fit on average.

All simulations are performed using the software R [21]. We use the package `neuralnet` that was implemented by our group and is published on CRAN as package for training neural networks [10].

3.2. Results

Table 1 summarizes the simulation results that we obtained for the additive model. It compares the theoretical and estimated penetrance matrix in four cases, namely risk scenarios 1 and 2 as well as random and equal starting weights as described in Subsection 3.1. The sum of all absolute, elementwise differences between the estimated and the theoretical penetrance matrix, i.e., $\sum D_{ij}$ as defined in equation (9), is provided. It is obvious that for the additive model, a neural network with reduced topology has on average a worse model fit than a neural network with complete topology, since the sum is much smaller for the complete topology in each case. Using for example random starting weights, the sums obtained from the reduced topology are about six times as large as those obtained from the complete topology (0.223 vs. 0.035 in scenario 1 and 0.373 vs. 0.060 in scenario 2).

Table 1. Additive model: comparison of complete and reduced network topology. Sum of absolute elementwise differences $\left(\sum D_{ij}\right)$

	Random starting weights		Equal starting weights	
	Scenario 1	Scenario 2	Scenario 1	Scenario 2
Complete topology	0.035	0.060	0.047	0.068
Reduced topology	0.223	0.373	0.226	0.375

The results for the multiplicative models are shown in Table 2. Reduced topology and complete topology have similar model fits with a tendency of the neural network with reduced topology to show somewhat better results, i.e., to show smaller values of the differences between the estimated and the theoretical penetrance matrix. Both topologies are well able to capture the multiplicative model since both network topologies obtain small sums of all absolute, elementwise differences between the estimated and the theoretical penetrance matrix $\sum D_{ij}$.

Table 3 summarizes the results for the recessive epistatic model. These are similar to those of the additive model. The reduced topology has a worse model fit and is therefore not sufficient to model the penetrance matrix of the epistatic model.

Table 2. Multiplicative model: comparison of complete and reduced network topology. Sum of absolute elementwise differences $\left(\sum D_{ij}\right)$

	Random starting weights		Equal starting weights	
	Scenario 1	Scenario 2	Scenario 1	Scenario 2
Complete topology	0.025	0.052	0.028	0.042
Reduced topology	0.014	0.024	0.019	0.025

Table 3. Recessive epistatic model: comparison of complete and reduced network topology. Sum of absolute elementwise differences $\left(\sum D_{ij}\right)$

	Random starting weights		Equal starting weights	
	Scenario 1	Scenario 2	Scenario 1	Scenario 2
Complete topology	0.089	0.046	0.079	0.058
Reduced topology	0.491	0.697	0.492	0.709

4. Discussion

The aim of the simulation study was to investigate the ability of our new approach to identify biological gene-gene interactions based on trained neural networks. For this purpose, the model fit of two neural networks differing in their network topology was compared. The idea was to characterize the two-locus disease model of independence by missing interconnections, i.e., by an identical model fit to the underlying data by a neural network with reduced topology. The simulations showed that this approach failed since the flexibility of the reduced topology does not seem to be sufficient to adequately describe the additive model. Although the additive model represents biological independence, neural networks need interconnections between both considered loci to capture this disease model. Additional analyses showed for the risk scenario 1 systematic deviations using the reduced topology, i.e., the reduced topology led to over- or underestimated penetrances, such that the structure of the additive model was ignored. The results were supported by the trained weights of the complete topology as there were no two weights that connected both loci in the hidden layer and that were simultaneously estimated as zero (data not shown).

Even more surprisingly, the results for the multiplicative models indicated that the reduced topology was sufficient to capture the assumed biological interactions. Although the complete topology in general describes a more complex relationship than the reduced one, the reduced topology

yields slightly better results, i.e., the estimated penetrance of the reduced topology is closer to the theoretical penetrance matrix than that of the complete topology. This might be due to the fact that the activation function in this case is comparable to the one of the logistic regressions. This means that interaction effects on the additive scale might appear as non-interaction effects on the multiplicative scale and can therefore not be detected by our approach.

The recessive epistatic model behaved like the additive model. Neural networks with reduced topology do not seem to be able to adequately capture this model. Since the epistatic model represents biological gene-gene interaction and the additive model represents biological independence, we can conclude that our approach in its present form is not able to distinguish biological gene-gene interaction from biological independence. Due to these unexpected results, no further simulations with alternative scenarios, like scenarios with a lower minor allele frequency or modeling other two-locus disease models were carried out.

Nevertheless, neural networks should be further investigated regarding their capability to capture disease models of biological independence or interaction. Our current research is on the estimated synaptic weights instead of the whole topology of a fitted neural network where we try to derive statistical tests based on appropriate confidence intervals that may allow to decide on the presence of biological interactions.

Acknowledgement

The authors gratefully acknowledge the financial support of this research by the grant PI 345/3-1 from the German Research Foundation (DFG).

References

[1] H. Akaike, Information theory and an extension of the maximum likelihood principle, B. N. Petrov and B. F. Csaki, eds., Second International Symposium on Information Theory, Academiai Kiado, Budapest, 1973, pp. 267-281.

[2] Y. Amit and D. Geman, Shape quantization and recognition with randomized trees, Neural Comput. 9 (1997), 1545-1588.

[3] C. M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, New York, 1995.

[4] L. Breiman, Random forests, Mach. Learn. 45 (2001), 5-32.

[5] L. Briollais, Y. Wang, I. Rajendram, V. Onay, E. Shi, J. Knight and H. Ozcelik, Methodological issues in detecting gene-gene interaction in breast cancer susceptibility: a population-based study in Ontario, BMC Med. 5 (2007), 22.

[6] S. H. Chen, J. Sun, L. Dimitrov, A. R. Turner, T. S. Adams, D. A. Meyers, B. L. Chang, S. L. Zheng, H. Grönberg, J. Xu and F. C. Hsu, A support vector machine approach for detecting gene-gene interaction, Genet. Epidemiol. 32 (2008), 152-167.

[7] N. R. Cook, R. Y. L. Zee and P. M. Ridker, Tree and spline based association analysis of gene-gene interaction models for ischemic stroke, Stat. Med. 23 (2004), 1439-1453.

[8] H. J. Cordell, Epistasis: what it means, what it does not mean, and statistical methods to detect it in humans, Hum. Mol. Genet. 11 (2002), 2463-2468.

[9] R. Foraita, K. Bammann and I. Pigeot, Modeling gene-gene-interactions using graphical chain models, Hum. Hered. 65 (2008), 47-56.

[10] F. Günther and S. Fritsch, Neuralnet: training of neural networks, R J. 2 (2011), 30-38.

[11] F. Günther, N. Wawro and K. Bammann, Neural networks for modeling gene-gene interactions in association studies, BMC Genet. 10 (2009), 87.

[12] R. Hecht-Nielsen, Neurocomputing, Addison-Wesley, Reading, 1990.

[13] A. G. Heidema, J. M. A. Boer, N. Nagelkerke, E. C. M. Mariman, D. L. van der A and E. J. M. Feskens, The challenge for genetic epidemiologists: how to analyze large numbers of SNPs in relation to complex diseases, BMC Genet. 7 (2006), 23.

[14] M. B. Lanktree and R. A. Hegele, Gene-gene and gene-environment interactions: new insights into the prevention, detection and management of coronary artery disease, Genome Med. 1 (2009), 28.

[15] W. Li and J. Reich, A complete enumeration and classification of two locus disease models, Hum. Hered. 50 (2000), 334-349.

[16] J. Millstein, D. V. Conti, F. D. Gilliland and W. J. Gauderman, A testing framework for identifying susceptibility genes in the presence of epistasis, Am. J. Hum. Genet. 78 (2006), 15-27.

[17] R. L. Milne, R. Fagerholm, H. Nevanlinna and J. Benítez, The importance of replication in gene-gene interaction studies: multifactor dimensionality reduction applied to a two-stage breast cancer case-control study, Carcinogenesis 29 (2008), 1215-1218.

[18] J. H. Moore and S. M. Williams, Traversing the conceptual divide between biological and statistical epistasis: systems biology and a more modern synthesis, Bioessays 27 (2005), 637-646.

[19] M. R. Nelson, S. L. R. Kardia, R. E. Ferrell and C. F. Sing, A combinatorial partitioning method to identify multilocus genotypic partitions that predict quantitative trait variation, Genome Res. 11 (2001), 458-470.

[20] B. North, D. Curtis and P. C. Sham, Application of logistic regression to case-control association studies involving two causative loci, Hum. Hered. 56 (2005), 79-87.

[21] R. Development Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2008.

[22] M. Riedmiller, Advanced supervised learning in multi-layer perceptrons-from back propagation to adaptive learning algorithms, Comput. Stand. Inter. 16 (1994), 265-278.

[23] N. Risch, Linkage strategies for genetically complex traits. I. Multilocus models, Am. J. Hum. Genet. 46 (1990), 222-228.

[24] M. D. Ritchie, L. W. Hahn, N. Roodi, L. R. Bailey, W. D. Dupont, F. F. Parl and J. H. Moore, Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer, Am. J. Hum. Genet. 69 (2001), 138-147.

[25] R. Rojas, Neural Networks, Springer, Berlin, 1996.

[26] I. Ruczinski, C. Kooperberg and M. LeBlanc, Logic regression, J. Comput. Graph. Stat. 12 (2003), 475-511.

[27] K. Strauch, R. Fimmers, M. P. Baur and T. F. Wienker, How to model a complex trait. 2. Analysis with two disease loci, Hum. Hered. 56 (2003), 200-211.

[28] R. Tibshirani, Regression shrinkage and selection via the lasso, J. Roy. Stat. Soc. B. 58 (1996), 267-288.

[29] M. J. Wade, R. G. Winther, A. F. Agrawal and C. J. Goodnight, Alternative definitions of epistasis: dependence and interaction, Trends Ecol. Evol. 16 (2001), 498-504.