Journal of Biomedical Informatics 43 (2010) 774–781

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Prediction of influenza vaccination outcome by neural networks and logistic regression

Ljiljana Trtica-Majnaric [a,*], Marijana Zekic-Susac [b], Natasa Sarlija [b], Branko Vitale [c]

[a] Department of Family Medicine, Medical School Osijek, University of Osijek, Osijek, Croatia
[b] Faculty of Economics, University of Osijek, Osijek, Croatia
[c] Institute Rudjer Bošković, Zagreb, Croatia

ARTICLE INFO

Article history:
Received 10 December 2009
Available online 6 May 2010

Keywords:
Influenza vaccination
Outcome
Prediction
Neural networks
Multilayer perceptron
Radial-basis function
Probabilistic network
Logistic regression

ABSTRACT

The major challenge in influenza vaccination is to predict vaccine efficacy. The purpose of this study was to design a model to enable successful prediction of the outcome of influenza vaccination based on real historical medical data. A non-linear neural network approach was used, and its performance compared to logistic regression. The three neural network algorithms were tested: multilayer perceptron, radial basis and probabilistic in conjunction with parameter optimization and regularization techniques in order to create an influenza vaccination model that could be used for prediction purposes in the medical practice of primary health care physicians, where the vaccine is usually dispensed. The selection of input variables was based on a model of the vaccine strain which has frequently been changed and on which a poor influenza vaccine response is expected. The performance of models was measured by the average hit rate of negative and positive vaccine outcome. In order to test the generalization ability of the models, a 10-fold cross-validation procedure revealed that the model obtained by multilayer perceptron produced the highest average hit rate among neural network algorithms, and also outperformed the logistic regression model with regard to sensitivity and specificity. Sensitivity analysis was performed on the best model and the importance of input variables was discussed. Further research should focus on improving the performance of the model by combining neural networks with other intelligent methods in this field.

1. Introduction

Prevention and control of influenza epidemics is a major challenge for public health care services. The current approach is annual vaccination with a trivalent inactivated influenza vaccine (against the A/H1N1, A/H3N2 and B influenza virus strains) [1]. Although this approach is generally safe and effective in preventing influenza, there is a need for influenza vaccines with improved efficacy in the elderly [2]. This need is based on the observation that available vaccines are less effective in the elderly than in the younger population [2,3]. Factors responsible for these differences have been identified and include: older age, previous exposures to influenza viruses, pre-existing antibody titres and chronic aging diseases [4,5].

Several new vaccine preparations and vaccination approaches are currently being pursued in order to improve the efficacy of influenza vaccines in the elderly [6,7]. Decision on the introduction of new vaccines into national vaccination programs is, however, connected with the demand for cost-effectiveness analyses and development of an influenza vaccination action plan [8]. The main challenge is to predict which individual will most likely adequately respond to conventional influenza vaccines and which individual will not.

Experience from clinical medicine suggests the use of a biology systems approach within the context of artificial intelligence, when obtaining models of prediction [9]. This assumption is based on the observation that chronic aging diseases are characterized by multiple factors, interacting with each other in a non-linear manner. This is the reason for the huge variability in disease expression and severity among individuals [10]. The artificial neural networks (ANN) has been shown to be a suitable computer-based method which can incorporate non-linear effects and interactions between multiple variables in a valid probability model [11].

Neural networks (NNs) as one of the artificial intelligence methods has been successfully used for classification, prediction and association in different fields, including general purpose, as well as some specific fields, such as diagnosis of disease [12]. Together with genetic algorithms, clustering algorithms, decision trees and other methods, NNs are widely used in Data Mining methodology for revealing hidden non-linear relationships among data [13].

* Corresponding author. Address: Department of Family Medicine, Medical School Osijek, University of Osijek, Strossmayerova 105, Osijek, Croatia. Fax: +385 31 376353.
E-mail address: ljiljana.majnaric@hi.t-com.hr (L. Trtica-Majnaric).

1532-0464/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2010.04.011

Previous research showed that, as a non-parametric method, NNs has remarkable information processing characteristics, pertinent mainly to non-linearity, high parallelism, fault and noise tolerance, learning and generalization capabilities [14]. Some of the other advantages of NNs are the ease of optimization, cost-effectiveness and flexible non-linear modeling of large datasets, and accuracy for predictive inference showing that NNs could serve as a valuable decision support tool in different areas, including medicine [15]. Three NN algorithms were tested in this paper: multilayer perceptron (MLP), radial-basis function network (RBFN) and probabilistic network (PNN). MLP is a general purpose feedforward network, and one of the most frequently used NN algorithms. In order to optimize the error function it uses the classical backpropagation algorithm based on deterministic gradient descent algorithm originally developed by Paul Werbos in 1974, extended by Rumelhart, Hinton, Williams (in [16]). RBFN is based on a clustering procedure for computing distances among each input vector and the center, represented by the radial unit. The ability of RBFN with one hidden layer to approximate any non-linear function has been demonstrated by Park and Sandberg (in [17]). The PNN algorithm was tested due to its fast learning and efficiency in classification. It is a stochastic-based NN developed by Specht (in [16]). The architecture of the PNN is based on Bayes'classifier, using the Parzen window estimator to estimate the probability distributions of the class samples [18].

Logistic regression (LR) modeling is widely used for analyzing multivariate data involving dichotomous responses dealt with in this paper. It provides a powerful technique analogous to multiple regression and ANOVA for continuous responses. Since the likelihood function of mutually independent variables $Y_1, \ldots, Y_n$ with outcomes measured on a binary scale is a member of the exponential family with $(\log(\frac{\pi_1}{1-\pi_1}), \ldots, \log(\frac{\pi_n}{1-\pi_n}))$ as a canonical parameter (π_j is a probability that Y_j becomes 1), the assumption of the logistic regression model is a linear relationship between a canonical parameter and the vector of explanatory variables $\mathbf{x}_j$ (dummy variables for factor levels and measured values of covariates):

$$\log\left(\frac{\pi_j}{1-\pi_j}\right) = \mathbf{x}_j^{\tau}\beta \tag{1}$$

This linear relationship between the logarithm of odds and the vector of explanatory variables results in a non-linear relationship between the probability of Y_j equals 1 and the vector of explanatory variables:

$$\pi_j = \exp(\mathbf{x}_j^{\tau}\beta)/(1+\exp(\mathbf{x}_j^{\tau}\beta)) \tag{2}$$

Detailed description of the logistic regression can be found in Harrel [19].

Three NN algorithms, as well as logistic regression were used in order to provide the influenza vaccination probability model that could be used for prediction purposes in the practice of primary health care physicians, where the vaccine is usually dispensed.

2. Materials

The study was based on original data collected in primary health care in Croatia. A total number of 90 patients, out of 150 individuals requiring the influenza vaccine during 2003/2004, gave their consent and were enrolled in the study. There were 35 male and 55 female patients, 50–89 years old (median 69), all suffering from multiple chronic medical conditions. Study protocol was approved by the local ethics committee.

The commercially licensed trivalent inactivated split vaccine was used for the vaccination, manufactured by Solway, the Netherlands, containing the following influenza virus strains: A/H1N1/New-Caledonia/20/99-like, A/H_3H_2/Moscow/10/99-like and B/Hong Kong-330/2001-like. Specific antibody production was measured by the Hemagglutination Inhibition (HI) test, a standard microtitre technique. At least a fourfold increase in antibody titre was used for expression of the specific antibody induction. For calculation of geometric mean titres (GMT), a titre of <1:10 was arbitrarily set at 5. The influenza B vaccine strain was also tested on the B/Sicuan 379/99 strain, contained in the vaccine in the recent past, for a heterologous reaction [1].

The target attribute in our study was the vaccination reaction to the influenza vaccine virus strain B/Hong Kong. The output variable used in NN models was binary, expressed in the form of two categories, where 0 value denoted the category of negative vaccine reaction, or less than the fourfold increase in antibody titre, while the value of 1 denoted the positive vaccine reaction, or the fourfold and more increase in antibody titre.

Available input space included a large set of variables, such as: the physician's diagnoses of the main groups of chronic diseases, anthropometric measures, hematological and biochemical laboratory tests. Blood tests were chosen on the basis of two criteria: (a) to determine the main age-related pathogenetic changes and (b) to be available in a real health care system setting. Based on these criteria, we performed blood tests to determine: (1) inflammation, (2) nutritional status, (3) metabolic status, (4) chronic renal impairment, (5) latent infections, (6) humoral immunity and (7) the neuroendocrine status. Blood samples were collected from subjects three times prior to the vaccination and once 4 weeks after the vaccination (for paired serum measuring). Hematological analyses were carried out on fresh blood samples, while sera for biochemical analyses and serological tests were separated by centrifugation and stored at −40 °C until assayed.

Due to the large dimension of initial input space, it was necessary to reduce the number of input variables before NN and LR modeling. Non-linear Data Mining algorithms were used for this purpose, resulting in a final set of 26 input variables [20]. Consequently, the results obtained by NN and LR modeling could be biased by the choice of the preprocessing method. Future research should be focused on the use of other preprocessing methods in modeling, and use of more datasets.

The total 26 input variables used in the NN and LR models can be divided into three groups: (1) data related to previous exposure to influenza viruses (the number of previous vaccinations and pre-existing antibody titres for all four influenza virus vaccine strains, measured in the study), (2) the set of medical data, and (3) age (implicating age-related changes in the immune system). All available input variables and their descriptive statistics (mean and standard deviation for continuous variables, frequencies for categorical variables) are presented in Table 1.

The whole sample consisted of 60 patients with negative vaccine outcome, and 30 patients with positive vaccine outcome.

Many authors, such as Liu and Tourassi, emphasize that model selection should be performed on the basis of generalization error [21,22]. Some of the well-known methods for testing the generalization ability of models are n-fold cross-validation, jackknifing, bootstrapping and round robin technique [13,22]. All of them have the purpose of reducing the small-sample estimation bias and variance contributions [22,23]. Cross-validation was used in this paper because it produces no statistical bias of the result since each tested sample is not a member of the training set. According to Witten and Frank, extensive tests on numerous datasets have shown that 10 is a sufficient value for n in the n-fold cross-validation [13]. Therefore, a 10-fold cross-validation procedure (or leave k cases out, where $k = 1/10$ of the total sample) was performed according to a slightly modified description of Masters including the following steps: (1) the in-sample data were divided into 10 equally-sized independent subsamples, (2) each NN model estimated 10 times, each time using a different

Table 1
Input variables and their descriptive statistics.

Variable No.	Variable code	Variable description	Descriptive statistics
1	VACC	The number of previous vaccinations 0 = vaccinated for the first time previously vaccinated: 1 = once, 2 = two or three times 3 = four or more times	0 = 39.79% 1 = 20.43% 2 = 13.98% 3 = 25.81%
2	H1N1_1	Pre-existing antibody titre on the influenza virus A/H1N1 strain	Mean = 11.08 stdev = 22.38
3	H3N2_1	Pre-existing antibody titre on the influenza virus A/H3N2 strain	Mean = 69.68 stdev = 63.54
4	KONG_1	Pre-existing antibody titre on the influenza virus B/Hong Kong strain	Mean = 43.44 stdev = 99.90
5	SICM_1	Pre-existing antibody titre on the influenza virus B/Sicuan strain	Mean = 30.32 stdev = 44.64
6	GLU	Fasting blood glucose	Mean = 6.52 stdev = 2.10
7	SKINFOLD	Triceps skinfold thickness (indicating malnutrition)	Mean = 33.37 stdev = 7.38
8	AGE	Age	Mean = 67.66 stdev = 7.96
9	PSYCH	Neuropsychiatric diseases (anxiety/depression, Parkinson's disease, cognitive impairments) (0 = no, 1 = yes)	0 = 40.86% 1 = 59.13%
10	HPA	*Helicobacter pylori* specific antibodies type IgA (indicating chronic gastritis)	Mean = 32.61 stdev = 51.39
11	HPG	*Helicobacter pylori* specific antibodies type IgG (indicating chronic gastritis)	Mean = 66.53 stdev = 63.42
12	EO	Eosinophils % in White Blood Cell differential (indicating humoral immunity)	Mean = 3.85 stdev = 2.65
13	MO	Monocytes % in White Blood Cell differential (indicating immune cell activation)	Mean = 8.10 stdev = 2.26
14	LY	Lymphocytes % in White Blood Cell differential (indicating lymphopenia)	Mean = 35.45 stdev = 8.99
15	MCV	Mean Cell Volume (indicating vitamin B12 deficiency)	Mean = 91.03 stdev = 5.03
16	ALB	Serum albumin (indicating inflammation/malnutrition)	Mean = 46.11 stdev = 3.13
17	CRCLEA	Creatinine clearance(indicating chronic renal impairment)	Mean = 1.69 stdev = 0.45
18	HOMCYS	Amino acid homocysteine (indicating nutritional status/chronic renal impairment)	Mean = 12.35 stdev = 3.81
19	BETA	Beta-globulins in serum proteins electrophoresis (indicating low-grade chronic inflammation)	Mean = 8.44 stdev = 0.94
20	GAMA	Gamma-globulins in serum proteins electrophoresis (indicating low-grade chronic inflammation/chronic humoral immune reaction)	Mean = 12.47 stdev = 2.29
21	VITB12	Vitamin B12 (indicating vitamin B12 deficiency/the nutritional status)	Mean = 284.33 stdev = 158.79
22	PRL	Hormone prolactin (indicating hyperprolactinemia)	Mean = 124.57 stdev = 120.39
23	TSH	TSH (thyroid-stimulating hormone) (indicating thyroid gland hormone hypofunction)	mean = 2.04 stdev = 2.61
24	FT3	Free triiodothyronine (thyroid gland hormone)(indicating thyroid gland hormone hypofunction)	Mean = 5.46 stdev = 0.53
25	FT4	Free thyroxine (thyroid gland hormone) (indicating thyroid gland hormone hypofunction)	Mean = 14.01 stdev = 2.21
26	IGE	Immunoglobulin type E (indicating impaired/decreased humoral immunity)	Mean = 135.91 stdev = 245.59
27	Output	Vaccine response (0 = negative, less then fourfold increase in antibody titre) (1 = positive, fourfold and more increase in antibody titre)	0 = 32.26% 1 = 67.74%

set of 9 subsamples for training, and tested on 1 sample that was left out of training, (3) 10 different results were obtained for each model, (4) an average of 10 obtained results, i.e. average error was computed [16]. The generalization ability in our study was measured by the average error, and the model with the lowest average error was selected as the best model.

The 10-fold CV procedure was performed on each of the three NN models: MLP, RBFN and PNN, as well as on the LR model. The size of the subsamples is presented in Table 2.

The purpose of the study was to design a computer-based neural network (NN) model that will enable successful prediction of the outcome of influenza vaccine efficacy based on data related to influenza viruses and influenza vaccination, in combination with historical medical data. The creation of the NN and LR models in this study was based on the immune response to the influenza virus strain B whose content in the vaccine was recently changed (new influenza virus vaccine strain) and on which a poor antibody response is expected [23].

3. Methods

3.1. Neural network methodology

Three NN algorithms were tested: MLP, radial basis and probabilistic. The output layer of all three NN models consisted of one neuron (valued as 1 for the positive response, and 0 for the negative response). One hidden layer was used in all NN models in our experiments. With regard to the number of hidden units, the method of pruning was used which eliminates weights which are lower than the threshold (0.05 in our experiments) input and

Table 2
Sampling in the 10-fold cross-validation procedure.

Sample	Total	
	Number of patients	Proportion (%)
Train	80	90
Test	10	10
Total	90	100

hidden units at the end of the training process in order to produce smaller and faster networks with equivalent performance. The initial number of hidden units was set to 15 in MLP networks, and to one-half of the size of the training sample in RBFN and to the size of the training sample in the PNN. Overtraining was avoided by a split-sample process which alternatively trains and tests the network (using a separate test sample) until the performance of the network on the test sample does not improve for n number of iterations. The maximum number of training epochs was set to 600. The generalization ability of all three NN models was determined by a 10-fold cross-validation procedure described in Section 2.

The level of output sensitivity to each input variable in NN models was computed by sensitivity analysis. Sensitivity analysis studies how the variation in the output of a mathematical model can be apportioned to different sources of variation in the input of a model. It is a technique that systematically changes model parameters to determine the effects of such changes [25]. The basic principle is that experimenting with a wide range of values gives insight into the behavior of a system in extreme situations, and can lead to identification of parameters whose specific value can significantly influence the behavior of the model.

There are various approaches to sensitivity analysis. We used a common OAT (one-factor-at-a-time) approach which changes the values of one input variable to see what effect it produces on the output, while all other variables are fixed to their central or baseline value [26]. The level of importance of each input variable is computed by a sensitivity index, which represents the relative sensitivity of the output to the changes of an input. Among the number of developed sensitivity indices, we used the importance index which can be interpreted to show the relative importance of an input variable to the output. A higher value of importance index means higher sensitivity of the output to changes of that particular input.

3.2. Logistic regression methodology

The aim of LR modeling in this study was to estimate the risk of reaction to influenza vaccine and to extract variables which are found to be important in risk prediction. The LR model used the same initial set of input variables as the NN model (see Table 1) with the forward selection procedure (selection criteria was $p < 0.05$). At the output, a binary variable was used with one category representing a patient with a negative influenza vaccine outcome (0) and the other representing a patient with a positive influenza vaccine outcome (1). The SAS software was used to conduct the procedure, with standard overall fit measures such as likelihood ratio and score, as well as c statistics which measure discriminative power of logistic equation.

3.3. Evaluating model performance

The performance of the NN and LR models was measured by the hit rate of negative vaccine outcome (i.e. the "negative hit rate" – hit_0), hit rate of positive vaccine outcome (i.e. the "positive hit rate" – hit_1), and the average hit rate ($ave\ hit$) computed by:

$$hit_0 = \frac{c_0}{t_0},\ hit_1 = \frac{c_1}{t_1},\ ave\ hit = \frac{hit_0 + hit_1}{2} \tag{3}$$

where c_0 is the number of patients accurately predicted to have negative vaccine outcome (i.e. the number of true negatives), t_0 is the number of patients with actual (target) negative vaccine outcome, c_1 is the number of patients accurately predicted to have positive vaccine outcome (i.e. the number of true positives), and t_1 is the number of patients with actual positive vaccine outcome. The above performance measures were computed on test samples for all 10 NN and LR models. In order to test the generalization ability of the models, the average hit rate of all ten samples was also computed and used as the measure of model performance, as well as the model selection criterion. The positive hit rate is equivalent to the term of model sensitivity, while the negative rate is equivalent to the model specificity, which is important when investigating the ability of the model to accurately recognize positive and negative outcomes. The computation of sensitivity and specificity also explains the proportion of false negatives or false positives that the models produce. For this purpose, their sensitivity and specificity ratios were computed in each of the 10 test samples used in the 10-fold cross-validation procedure, and type I and type II errors were calculated. The sensitivity is computed according to:

$$\text{sensitivity} = \frac{c_1}{(c_1 + d_0)} \tag{4}$$

where c_1 is the number of true positives and d_0 is the number of false negatives [26]. The specificity is computed according to:

$$\text{specificity} = \frac{c_0}{(c_0 + d_1)} \tag{5}$$

where c_0 is the number of true negatives and d_1 is the number of false positives. The sensitivity is equivalent to the positive hit rate hit_0, while the specificity is equivalent to the negative hit rate hit_1 (see Eq. (3)). The false positive rate (i.e. type I error) is computed as $\alpha = 1 -$ specificity, while the false negative rate (i.e. type II error) is computed as $\beta = 1 - sensitivity$. Comparison of false positive and false negative rates explains the tendency of a model to misclassify positive patients into the group of negatives, or negative patients into the group of positive patients [26]. The model with a high sensitivity could be used to screen for disease, since it has a tendency to misclassify more negative patients into the group of positive patients. The model with a high specificity could be used to confirm the test results, since it is more specific in recognizing the actual positive patients. The ideal situation would be that both the sensitivity and specificity of a model have high values [26].

4. Results

Among the three tested NN algorithms (MLP, RBFN and PNN), the best performance with regard to the average hit rate was obtained by the MLP algorithm. The results of the best NN model and LR models of the 10-fold cross-validation procedure are presented in Table 3.

It can be seen from Table 3 that the MLP hit rates across 10 samples in the 10-fold CV procedure varied, with the average hit rate of 72.52%. Following the modeling strategy described in Section 3, the best model is the model with the highest average hit rate out of the 10 samples. Therefore, the procedure showed that the NN model was more successful than the LR model. In order to analyze the statistical significance of the difference in the average hit results of the two models, non-parametric Mann–Whitney test was conducted since it is more appropriate in this case due to the small-sample size [27]. The results of the test showed that the difference between the average hit rates of the NN model and the LR model was significant at the level 10% ($p = 0.0587$).

Due to the fact that the difference in performance (i.e. the average hit rate) was statistically significant, it can be concluded that the NN model was more accurate in predicting vaccine outcome

Table 3
Neural network and logistic regression results of the 10-fold cross-validation procedure.

Test sample in CV procedure	Results of NN model			Results of LR model			
	Ave. hit rate (%)	Positive (1) hit rate (%)	Negative (0) hit rate (%)	Ave. hit rate (%)	Positive (1) hit rate (%)	Negative (0) hit rate (%)	LR model fitting measures
1	50.00	33.33	66.67	83.33	100.00	66.67	Likelihood ratio: 22.72, $p = 0.0009$, Score: 18.56, $p = 0.005$, $c = 0.808$
2	50.00	100.00	0.00	52.50	80.00	25.00	Likelihood ratio: 18.33, $p = 0.0004$, Score: 15.67, $p = 0.0013$, $c = 0.782$
3	28.57	57.14	0.00	42.50	60.00	25.00	Likelihood ratio: 36.72, $p < 0.0001$, Score: 28.93, $p = 0.0001$, c = 0.876
4	83.33	66.67	100.00	35.71	71.43	0.00	Likelihood ratio: 21.20, $p = 0.0003$, Score: 17.40, $p = 0.0016$, $c = 0.797$
5	83.33	100.00	66.67	65.00	80.00	50.00	Likelihood ratio: 36.83, $p < .0001$, Score: 28.94, $p = 0.0001$, $c = 0.872$
6	83.33	66.67	100.00	53.57	57.14	50.00	Likelihood ratio: 21.44, $p < .0001$, Score: 18.76, $p = 0.0003$, $c = 0.791$
7	91.67	100.00	83.33	60.71	71.43	50.00	Likelihood ratio: 19.34, $p = 0.0002$, Score: 16.56, $p = 0.0009$,$c = 0.795$
8	75.00	83.33	66.67	66.67	66.67	66.67	Likelihood ratio: 23.27, $p = 0.0003$, Score: 20.11, $p = 0.0012$, $c = 0.799$
9	90.00	100.00	80.00	58.33	50.00	66.67	Likelihood ratio: 28.31, $p = 0.0002$, Score: 22.42, $p = 0.0021$, $c = 0.843$
10	90.00	80.00	100.00	33.33	66.67	0.00	Likelihood ratio: 27.01, p = 0.0001,Score: 21.80, p = 0.0013,0.842
Ave. hit rate across 10 samples	72.52	78.71	66.33	55.17	70.33	40.00	

than the LR model. When the hit rate of each output category in the NN model is observed separately, it can be seen that the average hit rate was higher for the positive group of patients (78.71%) than for the negative group of patients (66.33%) indicating that this algorithm recognizes positive patients more easily.

Table 4 presents the average sensitivity and specificity, as well as type I and type II errors of the NN model, while these ratios obtained by the LR model are presented in Table 5. It can be seen from Table 4 that the specificity of the NN model is 0.79, while its sensitivity is 0.66. The false positive rate (i.e. type I error) of the NN model is 0.21, while the false negative rate (i.e. type II error) is 0.24. It reveals that the NN model is more sensitive than specific, tending to misclassify more patients who actually had negative vaccine outcome into the category of positive vaccine outcome. However, the difference between type I and type II errors is small, implying that the NN model is able to balance between sensitivity and specificity.

Although the LR model is also more sensitive than specific (see Table 5), both its sensitivity and specificity ratios are smaller than the same ratios of the NN model. The LR model's false positive rate (0.30) is twice as small as its false negative rate (0.60).

In order to further investigate the importance of each input variable to the output, sensitivity analysis was performed on each NN model in the 10-fold cross-validation procedure, and the average sensitivity ratio of the influenza vaccine outcome to input variables is presented in Fig. 1.

The analysis shows that the VACC is the variable with the highest influence on the vaccine outcome, closely followed by the HPG. The average sensitivity ratio is also high for variables PSYCH, BETA, HPA, EO, VITB12 and CRCLEA (sensitivity ratio higher than 1.1), while the ratio of other input variables is around 1.0. The variables with the lowest ratios (less than 1.0) are: AGE, H1N1_1, LY and HOMCYS.

In the process of LR modeling, a forward selection procedure was used to select input variables important for the model. By using such selection criteria, the LR extracted a total set of nine input variables: EO, VITB12, GLU, AGE, VACC, HPG, BETA, LY and IGE. Although the procedure of selecting input variables differs from the one used in NN modeling, some of the highly ranked variables identified by the NNs were also extracted by the LR models. If the rank of input variables is compared across NN and LR models, it can be seen that almost all variables extracted by the LR also had very high importance for the NNs (such as VACC, HPG, EO, VITB12, BETA, GLU), while some of the variables (such as AGE, IGE, and LY) were found to be significant for the LR model, although not very important for the NN model.

Table 4
The average sensitivity and specificity of the NN model.

Vaccine outcome predicted by the NN models	Actual vaccine outcome	
	1 (positive)	0 (negative)
1 (positive)	0.79	0.24
0 (negative)	0.21	0.66

Table 5
The average sensitivity and specificity of the LR model.

Vaccine outcome predicted by the NN models	Actual vaccine outcome	
	1 (positive)	0 (negative)
1 (positive)	0.70	0.60
0 (negative)	0.30	0.40

5. Discussion

The results clearly indicate that both types of data, those related to previous influenza viruses exposure and those describing the health status of examinees, influence outcome values of performed predictive models and could be used as efficient predictors of the influenza vaccine efficacy (Fig. 1). The reason for the low influence of the variable AGE in the NN models could be found in the sample structure, since most of the patients in the observed sample were elderly (mean age 67.67 years) (Table 1, Fig. 1). Furthermore, the variable AGE can hardly be separated from the factors related to previous antigenic exposure, and from factors related to chronic health disorders, both known to be age-dependent.

With regard to the influenza vaccine component, the content of which was recently changed (new influenza virus strain), such as

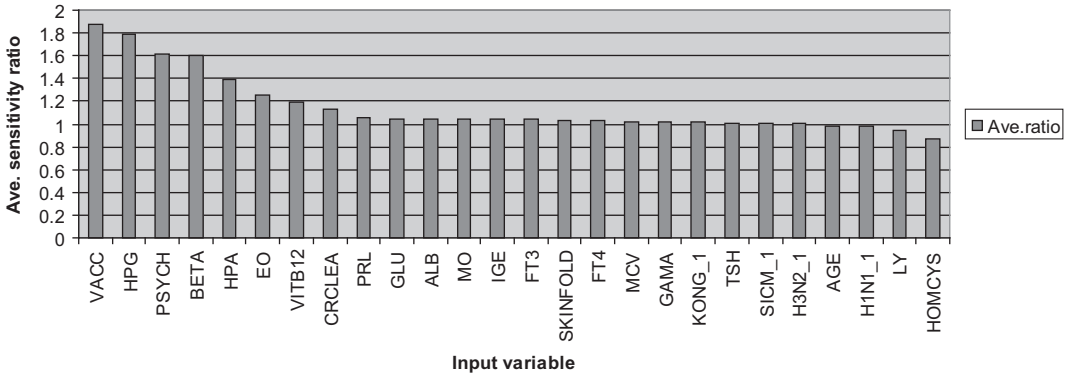

Fig. 1. Sensitivity ratios of the output variable towards the input variables.

the component B/Hong Kong, the most important information could be the number of previous vaccinations (variable VACC) (Fig. 1). In this connection, we analyzed the postvaccination antibody titres in relation to the number of previous vaccinations for all three influenza virus strains contained in the vaccine preparation used in the study (A/H1N1, A/H3N2 and B/Hong Kong). Indeed, a statistically significant decrease was only found for the component B/Hong Kong and only for individuals frequently vaccinated (previously vaccinated four or more times) (The results are not presented in the study). This may be a consequence of the negative impact of heterologous immune reaction (to influenza B virus strains from vaccine preparations used for vaccination in the recent past, before the last vaccine change) [23].

The results also show that for building a model of prediction the health status of examinees must be described with regard to many aspects (Fig. 1). Hematological and biochemical laboratory tests are preferable, in comparison to clinical diagnoses of diseases (only one diagnosis, that of neuropsychiatric disease, was selected as being relevant for prediction of low antibody response to influenza vaccination) (Table 3 and Fig. 1).

Some of the relevant variables, those with the highest influence on the vaccine outcome, are likely to indicate chronic clinical conditions which, in elderly people, could affect the antibody immune response. These variables are: (1) PSYCH, indicating neuropsychiatric diseases, (2) HPA (and HPG), indicating *Helicobacter pylori* positive chronic gastritis, and (3) CRLEA, indicating chronic renal impairment. Other variables, shown in this study to be necessary for the modeling, may represent common intermediate mechanisms, associated with these main clinical conditions (Fig. 1).

In accordance with these explanations, low influenza vaccination antibody response was found in the elderly with dementia/depression [28]. Evidence suggests that, in elderly people, both disorders, anxiety/depression and neurodegenerative/cognitive impairments, are closely related and frequently occur together [29]. Mechanisms such as neuroendocrine disorders, including also hyperprolactinemia and thyroid gland hormone hypofunction (in our results indicated by the variables PRL, FT3, FT4 and TSH), may link these disorders with impaired immune response to influenza vaccination [30–32]. Another proposed mechanism may be chronic activation of immunocomponent cells, during the course of neurodegenerative processes [33]. This can lead to a deficiency of vitamin B12 and, in turn, to impaired immune reaction, because this vitamin is necessary for proliferation of immunocompetent cells [34,35]. These assumptions, in our results, are likely to be represented by the variables: VITB12, MCV, MO and LY (Table 1, Fig. 1). The above explanation, based on existing knowledge on the issue, may be the reason for differences in variable selection, between the two used methods of prediction: the LR and the NN. Namely, because of the mutual interdependence between variables VITB12 and LY, they could gain different statistical power in the two models, based on the different performance due to linearity and dependence among the data.

Chronic gastritis is a frequent disorder in older people. In the majority of cases, it is caused by *H. pylori* infection. This disorder can be described as chronic inflammatory reaction in the gastric submucosa. Chronic inflammation and mechanisms such as increased oxidative stress and cytokine production may contribute to enhanced lymphocyte apoptosis and lymphopenia and, consequently, to impaired immune reaction [36]. In addition, because of intensive activation of B-lymphocytes in gastric submucosa, the disorder may lead to impairment of the specific humoral immune response [37]. In our results all these mechanisms, linking chronic gastritis with insufficient antibody production to influenza vaccination, are likely to be implicated by variables such as: BETA, EO, ALB, MO, IGE, GAMA and LY (Table 1 and Fig. 1). Close functional association between these variables may be the reason for complementariness among some of them, such as, e.g., the paired variables EO and IGE, indicating the same pathogenetic disorder – impaired humoral immunity. This may also provide a reasonable explanation for differences in selection of these variables in two different models of prediction: the LR and the NN.

Some other mechanisms can be proposed, linking chronic gastritis with low antibody response to influenza vaccination. Such mechanisms include vitamin B12 deficiency, probably due to malabsorption, and thyroid gland hormone hypofunction, due to evidence showing that chronic gastritis and chronic autoimmune thyroiditis, both common aging diseases, usually appear in comorbidity [38,39].

In general, B-vitamin deficiency and mild hyperhomocysteinemia, closely related metabolic disorders, may be markers of impaired cell cycling [40]. This may have the greatest negative impact on the function of cells with a high cell turn-over, such as immunocomponent cells [35,41].

Chronic renal impairment, a clinical condition in our results negatively influencing antibody response to influenza vaccination, has frequently been found to be associated with numerous disorders. Many of them are intermediate mechanisms, already mentioned above, including: inflammation, lymphopenia, increased mononuclear leukocyte activity, B-vitamin deficiency, hyperhomocysteinemia, as well as multiple neuroendocrine disorders [42–46]. Low-grade inflammation, common in patients with decreased renal function, is also associated with strong protein malnutrition, which in our results is most likely indicated by variables ALB and SKINFOLD (Table 1 and Fig. 1). Inflammation and protein malnutrition are mechanisms which both may affect the immune system function [47]. Moreover, metabolic disorders, including hyperhomocysteinemia and impaired glucose metabolism, are usually associated with inflammation/malnutrition, forming a unique syndrome

[43]. This fact, in our results, is likely to be supported by the variable GLU, indicating fasting blood glucose concentration (Table 1 and Fig. 1).

Hence, for influenza vaccine strains which are frequently changed (such as the B/Hong Kong vaccine strain), the poorest antibody response after vaccination can be expected in individuals previously vaccinated several times and who, in addition, are burdened by health disorders, with great impact on the immune reaction.

6. Conclusion

The aim of the study was to design an intelligent computer-based neural network model that will enable successful prediction of the outcome of influenza vaccine efficacy. The model is based on the results of vaccination by the influenza vaccine strain, the content of which was recently changed (in this case the B/Hong Kong vaccine strain) and on which, therefore, a poor antibody response was expected. The results were compared with a standard logistic regression approach. The input space consisted of 26 input variables, comprising both variables related to previous influenza virus exposure and previous vaccinations and variables describing many aspects of the health status of a group of high-risk older patients vaccinated against influenza.

Multilayer perceptron, radial-basis function, probabilistic neural network and logistic regression were used to predict the outcome of influenza vaccination, based on a set of previously selected input variables. Due to the small-sample size, it was necessary to perform a 10-fold cross-validation procedure in order to estimate the generalization ability of the model. The procedure showed that the multilayer perceptron algorithm had the highest average performance obtained on 10 samples, and therefore can be proposed as the model that generalizes better. The sensitivity and specificity ratios of the NN model were also higher than those ratios of the LR model. The neural network model was also able to balance between the false positive and false negative hit rates, showing that it was able to determine important features necessary to correctly classify patients with negative vaccine outcome and those with positive vaccine outcome.

The sensitivity analysis showed that the following predictor variables have the greatest importance for the output: VACC (the number of previous vaccinations), HPG (*H. pylori* specific antibodies type IgG), PSYCH (neuropsychiatric disease), BETA (beta-globulins in serum proteins electrophoresis), HPA (*H. pylori* specific antibodies type IgA), EO (eosinophils % in White Blood Cell differential), VITB12 (vitamin B12), and CRCLEA (creatinine clearance). Although the logistic regression approach extracted a more narrow set of predictors, it produced a significantly lower average hit rate compared to the neural network approach, implying that the multilayer structure of neural networks is more capable of understanding the interconnections among input variables and the output in order to successfully predict the vaccine output. Results obtained by both methods clearly indicate that the health status of the examined patients, for the purpose of prediction, must be determined by many aspects, and by a sufficiently large set of variables. This is due to the nature of chronic health disorders, characterized by many factors, each having only small predictive power.

Since this is a preliminary study, improvement in model performance can be expected by increasing the number of patients included in the dataset, and by using other intelligent methods, such as genetic algorithms, support vector machines, and other classifiers which could combine with neural networks in order to build a more successful model. The potential of the methodology in this field is evident, and benefit from this research in primary health care can be anticipated, as well as in more global strategic planning of influenza vaccination.

References

[1] Keitel WA, Cate TR, Couch RB. Efficacy of sequential annual vaccination with inactivated influenza virus vaccine. Am J Epidemiol 1988;127:353–64.
[2] Gross PA, Hermogenes AW, Sacks HS, Lau J, Levandowski RA. The efficacy of influenza vaccine in elderly persons. A meta-analysis and review of the literature. Ann Intern Med 1995;123:518–27.
[3] Goodwin K, Viboud C, Simonsen L. Antibody response to influenza vaccination in the elderly: a quantitative review. Vaccine 2006;24(8):1159–69.
[4] Remarque EJ. Influenza vaccination in elderly people. Exp Gerontol 1999;34(3):445–52.
[5] Beyer WE, Palache AM, Sprenger MJ, Hendriksen E, Tukker JJ, Darioli R, et al. Effects of repeated annual influenza vaccination on vaccine sero-response in young and elderly adults. Vaccine 1996;14(14):1331–9.
[6] Tosh PK, Poland GA. Emerging vaccines for influenza. Expert Opin Emerg Drugs 2008;13(1):21–40.
[7] Tripp RA, Tompkins SM. Recombinant vaccines for influenza virus. Curr Opin Investig Drugs 2008;9(8):836–45.
[8] Kim SY, Goldie SJ. Cost-effectiveness analyses of vaccination programmes: a focused review of modelling approaches. Pharmacoeconomics 2008;26(3):191–215.
[9] Cruz JA, Wishart DS. Application of Machine Learning in cancer prediction and prognosis. Review. Cancer Informatics 2006;2:59–78.
[10] Buchanan AV, Weiss KM, Fullerton SM. Dissecting complex disease: the quest for the philosopher's stone? Int J Epidemiol 2006;35:562–71.
[11] Linder R, Konig IR, Weimar C, Diener HC, Poppl SJ, Ziegler A. Two models for outcome prediction – a comparison of logistic regression and neural networks. Methods Inf Med 2006;45(5):536–40.
[12] Soucek B, the IRIS Group. Neural and intelligent systems integration. New York: John Wiley & Sons; 1991.
[13] Witten H, Frank E. Data mining, practical machine learning tools and techniques with java implementations. San Francisco: Morgan Kaufmann Publishers; 2000.
[14] Basheer IA, Hajmeer M. Artificial neural networks: fundamentals, computing, design, and application. J Microbiol Methods 2000;43:3–31.
[15] Lisboa PJ, Taktak AFG. The use of artificial neural networks in decision support in cancer: a systematic review. Neural Netw 2006;19:408–15.
[16] Masters T. Advanced algorithms for neural networks, A C++ sourcebook. New York: John Wiley & Sons; 1995.
[17] Karayiannis NB, Weigun GM. Growing radial basis neural networks: merging supervised and unsupervised learning with network growth techniques. IEEE Trans Neural Netw 1997;8:1492–505.
[18] Patterson DW. Artificial neural networks. New York: Prentice Hall; 1995.
[19] Harrel Jr FE. Regression modeling strategies with applications to linear models, logistic regression and survival analysis. Berlin: Springer; 2001.
[20] Majnaric TLJ, Vitale B, Martinis M. Non-linear approach in establishing influenza vaccine priority groups. Elsevier editorial system for vaccine™, manuscript draft, available from http://ees.elsevier.com/jvac/default.asp.
[21] Liu Y. Unbiased estimate of generalization error and model selection in neural network. Neural Netw 1995;8:215–9.
[22] Tourassi GD, Floyd CE. The effect of data sampling on the performance evaluation of artificial neural networks in medical diagnosis. Med Decis Making 1997;17:186–92.
[23] Hilden J. Neural networks and the roles of cross validation. Med Decis Making 1998;18:122–3.
[25] Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, et al. Global sensitivity analysis. The Primer, Chicester: John Wiley & Sons; 2008.
[26] Simon D, Boring III JR. Sensitivity, specificity, and predictive value. In: Walker HK, Hall WD, Hurst JW, editors. Clinical methods: the history, physical, and laboratory examinations. 3rd ed. Stoneham, MA: Butterworths; 1990. p. 49–54.
[27] Sheskin DJ. Handbook of parametric and nonparametric statistical procedures. Washington, DC: CRC Press; 1997.
[28] Vedhara K, Cox NK, Wilcock GK, Perks P, Hunt M, Anderson S. Chronic stress in elderly carers of dementia patients and antibody response to influenza vaccination. J Psychosom Res 2003;55(6):569–72.
[29] Magri F, Cravello L, Barili L, Sarra S, Cinchetti N, Salmoiraghi F, et al. Stress and dementia: the role of the hypothalamopituitaryadrenal axis. Aging Clin Exp Res 2006;18(2):167–70.
[30] Carta MG, Loviselli A, Hardoy MC, Massa S, Cadeddu M, Sardu C, et al. The link between thyroid autoimmunity (antithyroid peroxidase autoantibodies) with anxiety and mood disorders in the community: a field of interest for public health in the future. BMC Psychiatry 2004;18:4–25.
[31] Sonino N, Navarrini C, Ruini C, Fallo F, Boscaro M, Fava GA. Life events in the pathogenesis of hyperprolactinemia. Eur J Endocrinol 2004;151(1):61–5.
[32] Dorshkind K, Horseman ND. The roles of prolactin, growth hormone, insulin-like growth factor-I, and thyroid hormones in lymphocyte development and function: insights from genetic models of hormone and hormone receptor deficiency. Endocr Rev 2000;21(3):292–312.
[33] Widner B, Leblhuber F, Frick B, Laich A, Artner-Dworzak E, Fuchs D. Moderate hyperhomocysteinemia and immune activation in Parkinson's disease. J Neural Transm 2002;109(12):1445–52.
[34] Scott TM, Tucker KL, Bhadelia A, Benjamin B, Patz S, Bhadelia R, et al. Homocysteine and B vitamins relate to brain volume and white-matter changes in geriatric patients with psychiatric disorders. Am J Geriatr Psychiatry 2004;12(6):631–8.

[35] Schroecksnadel K, Frick B, Wirleitner B, Winkler C, Schennach H, Fuchs D. Moderate hyperhomocysteinemia and immune activation. Curr Pharm Biotechnol 2004;5:107–18.
[36] Peek RM, Miller GG, Tham Kyi T, Perez-Perez GI, Zhao X, Atherton JC. Heightened inflammatory response and cytokine expression in vivo to cagA$^+$ *Helicobacter pylori* strains. Lab Invest 1995;71(12):760–70.
[37] Hussell T, Isaacson PG, Crabtree JE, Dogan A, Spencer J. Immunoglobulin specificity of low grade B cell gastrointestinal lymphoma of mucosa-associated lymphoid tissue (MALT) type. Am J Pathol 1993;142(1):285–92.
[38] Sipponen P, Laxen F, Huotari K, Harkonen M. Prevalence of low vitamin B_{12} and high homocysteine in serum in an elderly male population: association with atrophic gastritis and *Helicobacter pylori* infection. Scand J Gastroenterol 2003;12:1209–16.
[39] Centanni M, Marignani M, Gargano L, Corleto VD, Casini A, Delle Fave G, et al. Atrophic body gastritis in patients with autoimmune thyroid disease: an underdiagnosed association. Arch Intern Med 1999;159(15):1726–30.
[40] Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet Suppl 2003;33: 245–54.
[41] Zhang C, Cai Y, Adachi MT, Oshiro S, Aso T, Kaufman RJ, et al. Homocysteine induces programmed cell death in human vascular endothelial cells through activation of the unfolded protein response. J Biol Chem 2001;276: 35867–74.
[42] van Guldener C, Robinson K. Homocysteine and renal disease. Semin Thromb Hemost 2000;26(3):313–24.
[43] Pecoits-Filho R, Lindholm B, Stenvinkel P. The malnutrition, inflammation and atherosclerosis (MIA) syndrome – the heart of the matter. Nephrol Dial Transplant 2002;17(Suppl. 11):28–31.
[44] Luger A, Lang I, Kovarik J, Stummvoll HK, Templ H. Abnormalities in the hypothalamic–pituitary–adrenocortical axis in patients with chronic renal failure. Am J Kidney Dis 1987;9(1):51–4.
[45] Kutlay S, Atli T, Koseogullari O, Nergizoglu G, Duman N, Gullu S. Thyroid disorders in hemodialysis patients in an iodine-deficient community. Artif Organs 2005;29(4):329–32.
[46] Fernandez-Fresnedo G, Ramos MA, Gonzalez-Pardo MC, de Francisco AL, Lopez-Hoyos M, Arias M. B lymphopenia in uremia is related to an accelerated in vitro apoptosis and dysregulation of Bcl-2. Nephrol Dial Transplant 2000;15(4):502–10.
[47] Hara M, Tanaka K, Hirota Y. Immune response to influenza vaccine in healthy adults and the elderly: association with nutritional status. Vaccine 2005;23(12):1457–63.