AAPS PharmSciTech **2004; 5 (1) Article 4 (http://www.aapspharmscitech.org).**

Artificial Neural Network as an Alternative to Multiple Regression Analysis in Optimizing Formulation Parameters of Cytarabine Liposomes

Submitted: June 6, 2003; Accepted: November17, 2003

Narayanaswamy Subramanian,[1] Archit Yajnik,[2] and Rayasa S. Ramachandra Murthy[1]

[1] Department of Pharmacy, Faculty of Technology and Engineering, M.S. University of Baroda, Kalabhavan, Vadodara-390001, India

[2] Department of Applied Mathematics, Faculty of Technology and Engineering, M.S. University of Baroda, Kalabhavan, Vadodara-390001, India

ABSTRACT

The objective of the study was to optimize the formulation parameters of cytarabine liposomes by using artificial neural networks (ANN) and multiple regression analysis using 3^3 factorial design (FD). As model formulations, 27 formulations were prepared. The formulation variables, drug (cytarabine)/lipid (phosphatidyl choline [PC] and cholesterol [Chol]) molar ratio (X_1), PC/Chol in percentage ratio of total lipids (X_2), and the volume of hydration medium (X_3) were selected as the independent variables; and the percentage drug entrapment (PDE) was selected as the dependent variable. A set of causal factors was used as tutorial data for ANN and fed into a computer. The optimization was performed by minimizing the generalized distance between the predicted values of each response and the optimized one that was obtained individually. In case of 3^3 factorial design, a second-order full-model polynomial equation and a reduced model were established by subjecting the transformed values of independent variables to multiple regression analysis, and contour plots were drawn using the equation. The optimization methods developed by both ANN and FD were validated by preparing another 5 liposomal formulations. The predetermined PDE and the experimental data were compared with predicted data by paired *t* test, no statistically significant difference was observed. ANN showed less error compared with multiple regression analysis. These findings demonstrate that ANN provides more accurate prediction and is quite useful in the optimization of pharmaceutical formulations when compared with the multiple regression analysis method.

KEYWORDS: artificial neural network, contour plots, cytarabine liposomes, multiple regression, factorial design

Corresponding Author: Rayasa S. Ramachandra Murthy, Department of Pharmacy, Faculty of Technology and Engineering, M.S. University of Baroda, Kalabhavan, Vadodara-390001, India; Tel: 091-265-2794051; Fax: 091-265-2423898; Email: murthyrsr@sify.com

INTRODUCTION

Various formulation and process variables relating to effectiveness, safety, and usefulness should be optimized simultaneously when developing pharmaceutical formulations. The difficulties in optimizing a pharmaceutical formulation are due to the difficulty in understanding the real relationship between casual and individual pharmaceutical responses. A response surface method (RSM) has often been applied to optimize the formulation variables.[1,2] The optimization procedure based on RSM includes statistical experimental designs, multiple regression analysis, and mathematical optimization algorithms for seeking the best formulation under a set of constrained equations. Since theoretical relationships between the response variables and causal factors are not clear, multiple regression analysis can be applied to the prediction of response variables on the basis of a second-order equation. The prediction of pharmaceutical responses based on second-order polynomial equation, however, is often limited to low levels, resulting in the poor estimation of optimal formulations.[3,4] To overcome the limitation of FD, artificial neural network (ANN) was incorporated.[5,6]

ANN is a massively parallel-distributed processor made up of simple processing units that has a natural propensity for storing experimental knowledge and making it available for use.[7] It resembles the brain in the way in which knowledge is acquired by the network from its environment through a learning process, and interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge. ANN could be applied to quantify a nonlinear relationship between causal factors and pharmaceutical responses by means of iterative training of data obtained from a designed experiment.

The ANN has 1 input layer, 1 or more hidden layers, and 1 output layer. Each layer has some units corresponding to neurons. The units in neighboring layers are fully interconnected with links corresponding to synapses. The strengths of the connections between 2 units are called "weights." The number of hidden layers or number of units in hidden layers

is arbitrarily defined. In each hidden layer and output layer, the processing unit sums its input from the previous layer and then applies the sigmoidal/logistic function to compute its output to the next layer according to the following equations[8]:

$$S = \sum_{i=1}^{p} W_{ij} X_i \quad (1)$$

where w_{ij} is the weight from node *i* in the input layer to node *j* in the hidden layer; x_i is the *i*th input element; and *p* is the number of nodes in the input layer. Generally a nonlinear sigmoidal/logistic function is used to regulate the output of a node, shown as follows:

$$F(S) = 1/(1 + e^{-\lambda s}) \quad (2)$$

where $F(S)$ is the output of the *j*th node in the hidden layer; and λ is the parameter relating to the shape of the sigmoidal function. Subsequently, output from the hidden layer is used as input to the output node. Finally, the overall response from the network is obtained via the output node in the output layer.[9] The sum of error squares [Σ(n)] for the nth iteration is defined as

$$\sum(n) = 1/2 \sum e^{-2}(n) \quad (3)$$

where $e^2(n)$ is the error signal at the output neuron and is the difference between desired response and computed response. Based on $\Sigma(n)$ the weights are updated in such a way that the error signal is minimized to the required threshold.

Cytarabine is one of the most effective anticancer agents used for various types of tumors.[10-12] The narrow therapeutic index, high volume of distribution, and poor tissue specificity require cytarabine to be delivered as liposomes. Drugs that are freely soluble in water, such as cytarabine, pose a great challenge to entrap them into the liposomes, as they have very low entrapment efficiency.[13,14] The entrapment may vary significantly from batch to batch as the number of formulation variables increases. Factorial design[15] and ANN[16-18] are useful models for studying the effect of several factors influencing the responses by varying them simultaneously and carrying out a limited number of experiments.

The present investigation aims to optimize the formulation variables of liposomes containing cytarabine prepared by the lipid film hydration method. Cytarabine is a highly water-soluble drug, and it is very difficult to achieve high entrapment efficiency for water-soluble drugs. Hence, percentage drug entrapment (PDE) is taken as the response parameters for the study. The process variables such as drying time, rotation speed, temperature, vacuum applied, and hydration time are kept constant, while the formulation variables, drug/lipid (phosphatidyl choline [PC] and cholesterol [Chol]) molar ratio (X_1), PC/Chol in percentage ratio of total lipids (X_2), and the volume of hydration medium (X_3), which have been predicted to play a significant role in enhancing the PDE are taken as variable parameters. The ratio of drug to lipid is very important for the prediction of PDE because the lower ratio may lead to incomplete entrapment of cytarabine and the higher ratio may lead to the presence of an excess of phospholipids (increase the cost of the formulation). The PC/Chol ratio is also very important because the presence of cholesterol provides rigidity to the liposomal bilayer, which will be useful to retain the drug within the liposomes. The third independent factor, volume of hydration medium, is also a very important factor as the drug is introduced into the liposomes after dissolving in the hydration medium. ANN and 3^3 factorial design are used to study the effects of the formulation variables on the PDE.

MATERIALS AND METHODS

Chemicals

Cytarabine was a gift from Dabur Research Foundation, Ghaziabad, India; egg PC was purchased from Sigma, St Louis, MO; Chol was purchased from S.D. Fine Chemicals, Mumbai, India; and DL-α-tocopherol was purchased from E. Merck India Ltd, Mumbai, India. All other chemicals and solvents were of analytical reagent grade.

Preparation of Liposomes

In the present study, drug/lipid (PC and Chol) in molar ratio, PC/Chol in percentage of total lipids, and the volume of hydration medium were selected as independent variables, whereas PDE within the liposomes was selected as dependent variable. The values of these selected variables along with their transformed values are shown in Table 1.

Twenty-seven batches of cytarabine liposomes were prepared by lipid film hydration method[19] according to the experimental conditions as shown in Table 2. PC, Chol, and α-tocopherol (0.5 mL of 0.1% wt/vol solution in chloroform) were dissolved in 5 mL of chloroform and methanol (2:1 by volume ratio) in a 250-mL round-bottom flask. The flask was rotated in the rotary flash evaporator at 100 rpm for 20 minutes in a thermostatically controlled water bath at 37°C under vacuum (600 mm of mercury). Drug solution (5 mg of drug dissolved in distilled water [hydration medium]) was added to the thin, dry, lipid film formed, and the flask was rotated again at the same speed and temperature as before

Table 1. Coded Values of the Formulation Parameters of Cytarabine Liposomes*

Coded Values	Actual Values		
	X_1	X_2	X_3
–1	1:7	50:50	1 mL
0	1:10	60:40	2 mL
1	1:13	70:30	3 mL

*X_1 represents the drug:lipid (molar ratio); X_2, the PC:Chol (in percentage of total lipids); and X_3, the hydration volume (distilled water).

Table 2. 3^3 Full Factorial Design Layout*

Batch No.	X_1	X_2	X_3	X_1^2	X_2^2	X_3^2	X_1X_2	X_2X_3	X_1X_3	$X_1X_2X_3$	Observed PDE ± SEM	Predicted PDE
1	–1	–1	–1	1	1	1	1	1	1	–1	43.5 ± 1.46	45.86
2	0	–1	–1	0	1	1	0	1	0	0	52.1 ± 2.24	53.60
3	1	–1	–1	1	1	1	–1	1	–1	1	77.8 ± 2.14	76.84
4	–1	0	–1	1	0	1	0	0	1	0	63.1 ± 1.50	60.28
5	0	0	–1	0	0	1	0	0	0	0	59.6 ± 2.10	61.91
6	1	0	–1	1	0	1	0	0	–1	0	78.6 ± 1.04	79.03
7	–1	1	–1	1	1	1	–1	–1	1	1	70.5 ± 2.11	70.05
8	0	1	–1	0	1	1	0	–1	0	0	64.1 ± 1.10	65.57
9	1	1	–1	1	1	1	1	–1	–1	–1	80.4 ± 2.15	76.57
10	–1	–1	0	1	1	0	1	0	0	0	70.4 ± 2.50	66.62
11	0	–1	0	0	1	0	0	0	0	0	66.9 ± 1.10	64.50
12	1	–1	0	1	1	0	–1	0	0	0	77.2 ± 2.34	77.86
13	–1	0	0	1	0	0	0	0	0	0	70.6 ± 2.40	71.52
14	0	0	0	0	0	0	0	0	0	0	69.2 ± 1.70	69.14
15	1	0	0	1	0	0	0	0	0	0	83.5 ± 1.60	82.25
16	–1	1	0	1	1	0	–1	0	0	0	71.3 ± 1.50	71.27
17	0	1	0	0	1	0	0	0	0	0	67.8 ± 1.13	69.14
18	1	1	0	1	1	0	1	0	0	0	77.4 ± 2.11	82.00
19	–1	–1	1	1	1	1	1	–1	–1	1	70.1 ± 1.11	72.21
20	0	–1	1	0	1	1	0	–1	0	0	64.2 ± 1.30	60.21
21	1	–1	1	1	1	1	–1	–1	1	–1	59.2 ± 2.19	63.70
22	–1	0	1	1	0	1	0	0	–1	0	64.5 ± 2.20	67.58
23	0	0	1	0	0	1	0	0	0	0	58.7 ± 1.62	61.20
24	1	0	1	1	0	1	0	0	1	0	75.4 ± 1.13	70.30
25	–1	1	1	1	1	1	–1	1	–1	–1	59.7 ± 1.32	58.32
26	0	1	1	0	1	1	0	1	0	0	60.2 ± 2.54	57.54
27	1	1	1	1	1	1	1	1	1	1	71.3 ± 1.68	72.25

*PDE indicates percentage drug entrapment; SEM, Standard error of mean. (n = 3).

but without vacuum for 30 minutes for lipid film removal and dispersion. The liposomal suspension so formed was then transferred to a suitable glass container and sonicated for 30 minutes using a probe sonicator (model RR-120, Ralsonics, Mumbai, India) in an ice bath for heat dissipation. The sonicated dispersion was then allowed to stand undisturbed for about 2 hours at room temperature for swelling. Each batch was prepared 3 times and stored in refrigerator.

Estimation of Entrapped Drug in Liposomes

Cytarabine entrapped within the liposomes was estimated after removing the unentrapped drug. The unentrapped drug was separated from the liposomes by subjecting the dispersion to centrifugation[19] in a cooling centrifuge (Remi Equipments, Mumbai, India) at 15 000 rpm at a temperature of –4°C for 30 minutes, whereupon the pellets of liposomes and the supernatant containing free drug were obtained. The liposome pellets were washed again with distilled water to remove any unentrapped drug by centrifugation. The com-

Table 3. Model Coefficients Estimated by Multiple Regression Analysis

Factor	Coefficient	Computed *t* Value	*P* Value
Intercept	69.141	40.742	.000000*
X_1	5.366	6.769	.000005*
X_2	2.323	2.930	.009813*
X_3	–0.356	–0.453	.656601
X_1^2	7.745	5.678	.000034*
X_2^2	–2.322	–1.702	.108040
X_3^2	–7.590	–5.509	.000048*
X_1X_2	–0.253	–0.252	.804281
X_2X_3	–3.658	–3.806	.001553*
X_1X_3	–4.008	–4.170	.000722*
$X_1X_2X_3$	5.863	4.980	.000136*

* *P* value is very significant at $P < .01$

bined supernatant was analyzed for the drug content after suitable dilution with methanol by measuring absorbance at 274 nm using a Hitachi U-2000 double-beam spectrophotometer (Hitachi Ltd, Tokyo, Japan). The PDE in the liposomes was calculated from the difference between the initial drug added and the drug detected in the supernatant. The amount of drug exactly present within the liposomes was also analyzed by dissolving the liposomes in methanol to countercheck the PDE and to arrive at a mass balance. The analysis of drug in liposomes was carried out using the empty liposomes dissolved in methanol as blank in order to nullify the interference of the excipients. The mean PDE of all 27 batches is shown in Table 2.

Multiple Regression Analysis

A prior knowledge and understanding of the process and the process variables under investigation are necessary for achieving a more realistic model. Based on the results obtained in preliminary experiments, drug/lipid ratio, PC/Chol ratio, and hydration volume were found to be the major variables in determining the PDE. Hence, these variables were selected to find the optimized condition for higher PDE using 3^3 factorial design and contour plots.

Twenty-seven batches of different combinations were prepared by taking values of selective variables X_1, X_2, and X_3 at different levels as shown in Table 1. The prepared batches were evaluated for PDE, a dependent variable, and the results are recorded in Table 2. A multilinear stepwise regression analysis was performed using Microsoft Excel software. Mathematical modeling was carried out by using Equation 4 to obtain a second-order polynomial equation.[20]

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_1{}^2X_{11} + b_2{}^2X_{22} + b_3{}^2X_{33} + b_{12}X_1X_2 + b_{23}X_2X_3 + b_{13}X_1X_3 + b_{123}X_1X_2X_3 \quad (4)$$

where Y is the dependent variable (PDE), while b_0 is the intercept; b_i (b_1,b_2, and b_3), b_{ij} (b_{12}, b_{23}, and b_{13}), and b_{ijk} (b_{123}) represent the regression coefficient for the second-order polynomial; and X_i represents the levels of independent formulation variables. A full model (Equation 5) was established after putting the values of regression coefficients in Equation 4. The predicted values along with their observed values are shown in Table 3, which gives information about the percentage of error obtained when the predicted value was compared with the observed values, and the predicted values were calculated by using the mathematical model derived from the coefficients of the model as shown in Table 4.

$$Y = 69.141 + 5.366X_1 + 2.323X_2 - 0.356X_3 + 7.745X_1^2 - 2.322X_2^2 - 7.59X_3^2 - 0.253X_1X_2 - .658X_2X_3 - 4.008X_1X_3 + 5.863X_1X_2X_3 \quad (5)$$

Neglecting nonsignificant ($P < .01$) terms from the full model established a reduced model (Equation 6), which facilitates the optimization technique by plotting contour plots and response surface plots keeping 1 independent formulation variable constant and varying the other 2 independent formulation variables, to establish the relationship between independent and dependent variables.

$$Y = 67.556 + 5.394X_1 + 2.29X_2 + 7.71X_1^2 - 7.53X_3^2 - 3.658X_2X_3 - 4.008X_1X_3 + 5.863X_1X_2X_3 \quad (6)$$

Results of analysis of variance (ANOVA) of full model and reduced model were carried out and the F statistic was applied to check whether the nonsignificant terms can be omitted or not from the full model, which is shown in Table 5.

Table 4. Analysis of Variance of Full and Reduced Model*

		DF	SS†	MS‡	F§	R	R^2	Adj R^2
Regression	FM	10	1980.817	198.082	17.866	0.9580	0.9178	0.8664
	RM	7	1944.702	277.814	24.723	0.9493	0.9011	0.8646
Error	FM	16	177.390 (E1)	11.087				
	RM	19	213.504 (E2)	11.237				

*DF indicates Degree of freedom; E1 and E2 indicated Sum of squares of error of full and reduced model respectively; F, Fischer ratio; FM, full model; MS, Mean squares; RM, reduced model; and SS, Sum of squares. Number of parameters omitted = 3.
†SSE2 – SSE1 = 213.504 – 177.390 = 36.114
‡MS of error (full model) = 11.087
§F calculated = (36.114/3)/11.087 = 1.08

Table 5. Validation of the Established Relationships*

No.	X_1 (Drug:Lipid)	X_2 (PC:Chol)	X_3 (Hydration Volume)	Y†	Y-BP‡	Y-FD§
1	1: 12.40	67.6:32.4	1	79.2	78.90	80.00
2	1:11.00	57.2:42.8	2	73.1	74.25	70.00
3	1: 11.60	65.2:34.8	2	77.8	78.29	75.00
4	1: 12.40	57.6:42.4	3	74.6	74.72	75.00
5	1: 10.66	56.0:44.0	3	59.9	59.10	60.00
		$t_{calculated}$			0.7122	0.3337
		$t_{tabulated}$			2.7765	
		Normalized Error			0.02197	0.05681

†Y is the percentage drug entrapment (PDE), obtained from experiments.
‡Y-BP is the PDE predicted by Back Propagation Network (artificial neural network).
§Y-FD is the PDE predicted by multiple regression analysis.

Contour Plots

Two-dimensional contour plots were established using reduced polynomial equation (Equation 6). Values of X_1 and X_2 were computed at prefixed values of PDE. Three contour plots were established between X_1 and X_2 at fixed level if –1, 0, and 1 level of X_3 as shown in Figure 1 (A, B, and C).

Artificial Neural Network

A commercial Microsoft Windows-based ANN software, Matlab Version 6.1 (The MathWorks, Natick, MA) was used throughout the study with a P-4 personal computer. This software allows the user to select the number of hidden layers and hidden layer nodes (neurons), iterations used dur-

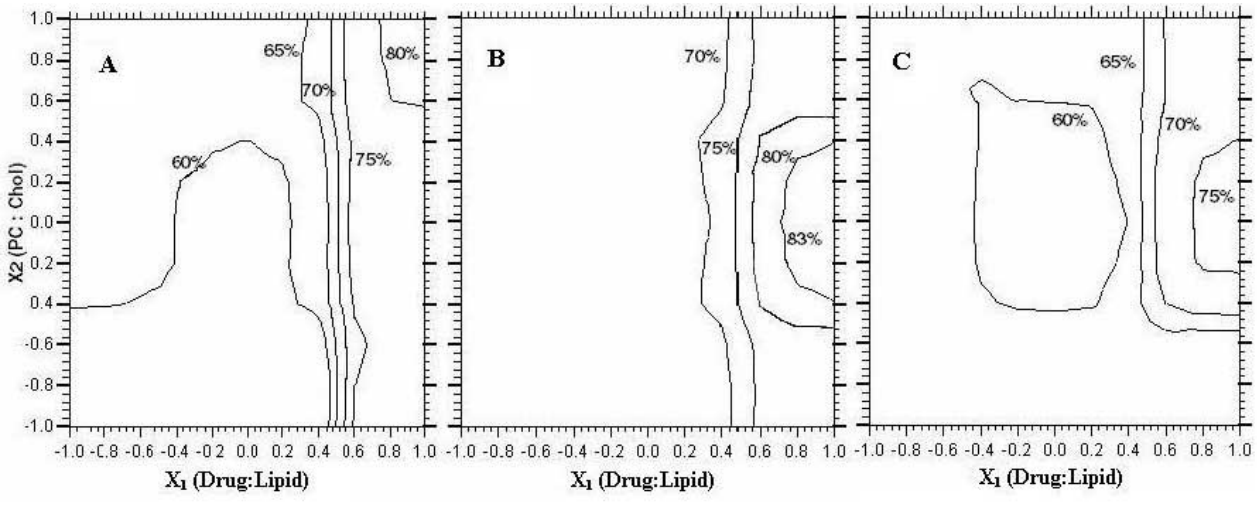

Figure 1. Contour plots (A) at –1 level of variable X_3, (B) at 0 level of variable X_3, (C) at 1 level of variable X_3.

ing the model training, learning algorithm, and transfer functions.

A multilayer feed-forward back-propagation network, which was created by generalizing the Levelberg-Marquardt's learning rule to multiple layer networks and nonlinear differential transfer functions, was used to predict PDE of the liposomal formulations. Our network architecture consisted of an input layer with 3 neurons, an output layer with 1 neuron, and a hidden layer (Figure 2). The number of hidden nodes in a network is critical to network performance. Too few nodes can lead to underfitting. Too many nodes can lead the system toward memorizing the patterns in the data.[21] According to Kolmogorov's theorem, it was understood that twice the number of input nodes plus one is sufficient to compute any arbitrary continuous function.[22] Hence, we started off with Kolmogorov's number of hidden nodes and increased the number until a network with the least mean-squared error was attained.

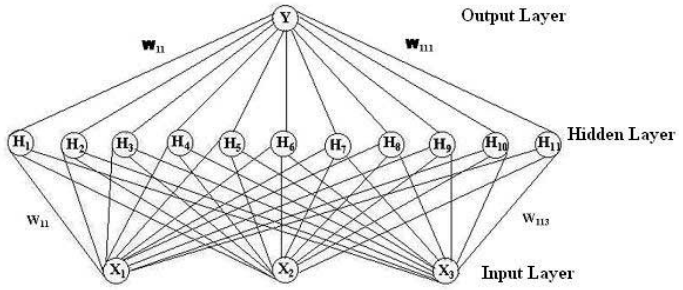

Figure 2. The feed forward back propagation network used in the study. X_1, Drug:Lipid, X_2, PC:Chol, X_3, volume of hydration medium,Y, PDE; H_1 -H_{11}, nodes of the hidden layer; W_{11}, connection from first input node to the first hidden node; W_{11}, connection from the hidden node to the output node; W_{113}, connection from the third input node to the eleventh hidden node; W_{111}, connection from the eleventh hidden node to the output node.

Input vectors and the output vector (response) were used to train the network until it could approximate a function (ie, associate input vectors with specific output vectors). Trained back-propagation networks tend to give reasonable answers when presented with inputs that they have never seen.

In the experiment, based on the 3^3 factorial design, 27 batches were prepared and PDE of these batches was used for training. Learning rate and error goal were selected on a trial and error basis in such a way so as to keep the minimum distance between the actual and predicted value. The second set was used for validation of the trained network. Here the paired *t* test was applied between experimental and predicted value.

Normalized Error Determination

The quantitative relationship established by both techniques (ANN and FD) was confirmed by preparing experimentally 5 liposomal formulations by random selection of causal factors. PDEs predicted from the ANN and FD were compared with those generated from physical experiment using Normalized Error (NE). The equation of NE is expressed as follows:

$$NE = [\Sigma\{(Pr - Er)/Er\}^2]^{1/2} \quad (7)$$

where Pr and Er represent predicted and experimental response, respectively.

RESULTS AND DISCUSSION

Multiple Regression Analysis

By using 3^3 factorial design (Table 1), 27 batches of cytarabine liposomes were prepared by the lipid film hydration method varying 3 independent variables: drug:lipid (molar ratio) (X_1); PC:Chol (in percentage of total lipids) (X_2); and volume of hydration medium (X_3). The PDE, which was taken as a dependent variable, was determined and the results were recorded (Table 2). A substantial high drug entrapment achieved in liposomes prepared by lipid film hydration method was 83.5% at 1 level of X_1 (1:13), 0 level of X_2 (60:40), and 0 level of X_3 (2 mL). The reasons for high entrapment at 0 levels of X_2 and X_3 may be because the increase in Chol content above 40% in the bilayer led to a reduction in PDE, and as the drug was dissolved in the hydration medium, the increase in hydration volume beyond 2 mL led to a reduction in PDE. The liposome formulations were lyophilized using sucrose as cryoprotectant. The lyophilized liposome powder was coated with gold and kept in the sampling unit as a thin film; then a photomicrograph was taken at 11 000 magnification using Jeol Scanning Electron Microscope (Jeol, JSM-840 SEM,Tokyo, Japan), which showed that the liposomes formed are spherical in shape (see Figure 3).

The PDE (dependent variable) obtained at various levels of 3 independent variables (X_1, X_2 , and X_3) were subjected to multiple regression to yield a second-order polynomial equation (full model). The PDE values for the 27 batches showed a wide variation from 43.5% to 83.5% (Table 2). The values X_3, X_2^2, and X_1X_2 in Equation 5 are regarded as least contributing in the preparation of cytarabine liposomes by lipid film hydration method. Hence, these terms are neglected from the full model considering nonsignificance and a reduced polynomial equation (Equation 6) was obtained.

The significance of each coefficient of Equation 5 was determined by Student *t* test and *P* value, which are listed in

Table 3. The larger the magnitude of the *t* value and the smaller the *P* value, the more significant is the corresponding coefficient.[23,24] This implies that the quadratic main effects of drug/lipid ratio and PC/Chol ratio are significant. The second-order main effects of both drug/lipid ratio and volume of hydration are significant, as is evident from their *P* values. The interaction between X_2X_3, X_1X_3, and $X_1X_2X_3$ are found to be very significant from their *P* values (Table 3).

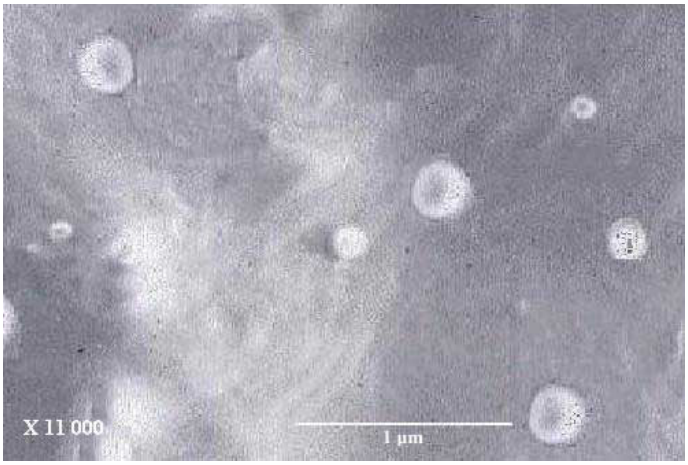

Figure 3. SEM photograph of cytarabine liposomes using original magnification ×11 000.

The results of ANOVA of the second-order polynomial equation are given in Table 4. F statistic of the results of ANOVA of full and reduced model confirmed omission of nonsignificant terms of Equation 5. Since the calculated F value (1.086) is less than the tabulated F value (3.25) (α = 0.05, V_1 = 3 and V_2 = 16), it was concluded that the neglected terms do not significantly contribute in the prediction of PDE. When the coefficients of the 3 independent variables in Equation 5 were compared, the value for the variable X_1 (b_1 = 5.366) was found to be maximum; hence the variable X_1 was considered to be a major contributing variable for PDE of cytarabine liposomes. The Fisher F test with a very low probability value demonstrates a very high significance for the regression model. The goodness of fit of the model was checked by the determination coefficient (R_2). In this case, the values of the determination coefficients (R_2 = 0.9178 for full model and 0.9011 for reduced model) indicated that over 90% of the total variations are explained by the model. The values of adjusted determination coefficients (adj R_2 = 0.8664 for full model and 0.8646 for reduced model) are also very high, which indicates a high significance of the model. The higher values of correlation coefficients (R = 0.958 for full model and 0.9493 for reduced model) signify an excellent correlation between the independent variables.[25-27]

Contour Plots

Values of X_1 and X_2 were computed at prefixed values of PDE and contour plots were established. The variable X_3, being least significant, was kept constant in drawing the contour plots. Three contour plots were established between X_1 and X_2 at fixed level if –1, 0, and 1 level of X_3 as shown in Figure 1 (A, B, and C). The contour plots showed very clearly the relationship between the independent variables and the PDE. The developed contour plots were used to predict the PDE. Five checkpoints were selected from the contour plots, and the predicted PDE was compared with the experimental PDE using paired Student *t* test. The results of the *t* test proved that the difference between the predicted and experimental PDE was not statistically significant (*t* value = 0.3337). This demonstrates the effective use of reduced polynomial equation and contour plots in determining the PDE in the preparation of cytarabine liposomes.

Artificial Neural Network

A multilayer feed-forward back-propagation network using Levelberg-Marquardt's learning rule was used to predict PDE of the liposomal formulations. Three causal factors corresponding to different levels of drug:lipids (X_1), PC:Chol (X_2), and volume of hydration medium (X_3) were used as each unit of input layer. The output layer was composed of one response variable Y, PDE. A set of release parameters and causal factors was used as tutorial data for ANN and fed into a computer. Several iterations were conducted with different numbers of nodes of hidden layer and training times in order to determine the optimal ANN structure.[28] For selecting the number of hidden nodes, we started with 3 hidden nodes and gradually increased the number of nodes until a network of least mean squared error was attained. Increase in the number of nodes led to decrease in least mean squared error. Finally, with 11 hidden nodes, we could achieve the least mean squared error and excellent prediction of the response variable. Further increase in hidden nodes produced high error, when the network was validated with another set of test data (Table 5). The Student *t* test carried out between the predicted results (*t* value = 0.722) from the ANN and the experimental results showed no statistically significant difference between them. The NE between the predicted and experimental response variables was employed as an evaluation standard between ANN and FD. The NE value observed with the optimal ANN structure was 0.021 97, while it was 0.056 81 in case of second-order polynomial equation (FD).

Comparison of ANN and FD

Both ANN and FD visualized similar results, and their predictions regarding the PDE coincided very well. To check the accuracy of these predictions, we prepared experimentally 5 liposomal formulations by random selection of causal factors. Experimental results were comparable with the predicted results (Table 5). Data analyzed using paired Student *t* test revealed that there was no statistically significant difference between the experimental results and the predicted results of ANN and FD. A close look at both ANN and FD reveals the following facts. The normalized error obtained from ANN was less, compared with the multiple regression analysis, and shows the higher accuracy in prediction. ANN can easily handle more input variables and is extremely helpful when the number of experiments is greater, but in the case of factorial design, a large number of input variables leads to a polynomial with many coefficients, which involves tedious computation. Another major advantage with ANN is the flexibility to work with the theoretical data for better prediction, but FD does not accommodate theoretical or historical data.

CONCLUSION

The study demonstrated that 3^3 factorial design (FD) and back-propagation network (ANN) are useful tools to understand the effects of the various formulation parameters in the preparation of cytarabine liposomes by lipid film hydration method and to predict the best composition for a particular response. The optimal formula for the high PDE (83.5%) was found to be drug:lipid (molar ratio), 1:13; PC:Chol (in percentage of total lipids), 60:40; and 2 mL of hydration volume. Thus, desirable goals can be achieved by a systematic formulation approach in the shortest possible time with a reduced number of experiments, thereby reducing the cost of development of the formulations.

ACKNOWLEDGEMENTS

The authors are thankful to the Council of Scientific and Industrial Research, New Delhi, India, for funding this project (Grant No. 9/114(137)/2K2/EMR-I).

REFERENCES

1. Levison KK, Takayama K, Isowa K, Okaba K, Nagai T. Formulation optimization of indomethacin gels containing a combination of three kinds of cyclic monoterpenes as percutaneous penetration enhancers. *J Pharm Sci*. 1994;83:1367-1372.

2. Shirakura O, Yamada M, Hashimoto M, Ishimaru S, Takayama K, Nagai T. Particle size design using computer optimization technique. *Drug Dev Ind Pharm*. 1991;17:471-483.

3. Takayama K, Morva A, Fujikawa M, Hattori Y, Obata Y, Nagai T. Formula optimization of theophylline controlled-release tablet based on artificial neural networks. *J Control Release*. 2000;68:175-186.

4. Takahara J, Takayama K, Nagai T. Multi-objective simultaneous optimization technique based on an artificial neural network in sustained release formulations. *J Control Release*. 1997;49:11-20.

5. Takayama K, Fujikawa M, Nagai T. Artificial neural network as a novel method to optimize pharmaceutical formulations. *Pharm Res*. 1991;16:1-6.

6. Takahara J, Takayama K, Isowa K, Nagai T. Multi-objective simultaneous optimization based on artificial neural network in a ketoprofen hydrogel formula containing O-ethylmenthol as a percutaneous absorption enhancer. *Int J Pharm*. 1997;158:203-210.

7. Simon Haykin. *Neural Networks: A Comprehensive Foundation*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall; 1999:156-254.

8. Achanta AS, Kowalski IG, Rhodes CT. Artificial neural networks: implications for pharmaceutical sciences. *Drug Dev Ind Pharm*. 1995;21:119-155.

9. Baughman DR, Liu YA. *Neural Networks in Bioprocessing and Chemical Engineering*. New York, NY: Academic Press; 1995.

10. Tricot G, De Bock R, Dekker AW. Low dose cytosine arabinoside (Ara C) in myelodysplastic syndromes. *Br J Haematol*. 1984;58:231-240.

11. Roberts JD, Ershlew WB, Tindle BH. Low-dose cytosine arabinoside in the treatment of myelodysplastic syndromes and acute myelogenous leukemia. *Cancer*. 1985;56:1001-1005.

12. Winter JN,Variakojis D, Gaynor ER. Low-dose cytosine arabinoside (Ara-C) therapy in myelodysplastic syndromes and acute leukemia. *J Cancer*. 1985;56:443-449.

13. Allen TM, Mehra T, Hansen C, Chin YC. Stealth liposomes: an improved sustained release system for 1-ß-D-arabinofuranosylcytosine. *Cancer Res*. 1992;52(9):2431-2439.

14. Zou Y, Ling YH, Van NT, Priebe W, Perz Solar R. Antitumor activity of free and liposome-entrapped annamycin, a lipophilic antracycline antibiotic with non-cross-resistance properties. *Cancer Res*. 1994;54(6):1479-1484.

15. Cochran WG, Cox GM. *Experimental Designs*. 2nd ed. New York, NY: John Wiley & Sons; 1992.

16. Murtonemi E, Yliruusi J, Kinnunen P, Merkku P, Leiviska K. The advantages by the use of the neural networks in modeling the fluidized bed granulation process. *Int J Pharm*. 1994;108:155-164.

17. Hussian AS. Application of neural computing in pharmaceutical product development. *Pharm Res*. 1991;8(10):1248-1252.

18. Ebube NK, McCall T, Chen Y, Meyer MC. Relating formulation variables to in vitro dissolution using an artificial neural network. *Pharm Dev Technol*. 1997;2(3):225-232.

19. New RRC. Preparation of liposomes. In: *Liposomes: A Practical Approach*. Oxford, UK: Oxford University Press; 1990:33-104.

20. Armstrong NA, James KC. *Pharmaceutical Experimental Design and Interpretation*. Bristol, PA: Taylor and Francis Publishers; 1996:131-192.

21. Erb RJ. Introduction to backpropagation neural network computation. *Pharm Res*. 1993;10:165-170.

22. Nielsen RH. Kolmogrov's mapping neural network existence theorem. In: Proceedings of the Second IEEE International Conference on Neural Networks; June 21-24, San Diego, CA. 1987: 11-14.

23. Akhnazarova S, Kafarov V. *Experiment Optimization in Chemistry and Chemical Engineering*. Moscow, Russia: Mir Publications; 1982.

24. Adinarayana K, Ellaiah P. Response surface optimization of the critical medium components for the production of alkaline protease by a newly isolated *Bacillus* sp. *J Pharm Pharm Sci*. 2002;5(3):281-287.

25. Box GEP, Wilson KB. On the experimental attainment of optimum conditions. *J Roy Stat Soc*. "Ser C Appl Stat" 1951;13:1-45.

26. Box GEP, Hunter WG, Hunter JS. *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*. New York, NY: John Wiley & Sons; 1978.

27. Yee L, Blanch HW. Defined media optimization for the growth of recombinant *Escherichia coli* x90. *Biotechnol Bioeng*. 1993;41:221-227.

28. Jha BK, Thambe SS, Kulkarni BD. Estimating diffusion coefficients of a micellar system using an artificial neural network. *J Colloid Interface Sci*. 1995;170:392-398.