CHAPTER III

MATERIALS AND METHODS
It has been proposed to analyze fat and water soluble serum vitamin concentrations for the selected subjects using High Performance Liquid Chromatography (HPLC) and executed statistically by artificial neural networks to estimate the optimum vitamin values.

Selection of subjects: The subjects (male and female) with various universally prevalent diseases, conditions and unhealthy habits were selected. In this study the following physical, personal and clinical parameters details of the subjects were recorded.

Physical parameters:
- Gender (Male or Female)
- Age of the subject (yrs)
- Weight of the subject (kgs)

Clinical parameters:
- Blood group with Rh factor
- Hemoglobin % (mg/dL)
- Heamatocrit %(Packed cell volume)

During the present study vitamin concentrations in subjects with different ailments, unhealthy habits and certain specified conditions were determined.

Pathological conditions:
- Human immune deficiency virus (H.I.V)
- Coronary Artery Disease (C.A.D)
- Diabetes Mellitus (D.M)
- Hyper and hypo Tension (H.T)
- Tuberculosis (T.B)
- Alzheimer’s disease.

Unhealthy habits:
- Smoking.
- Consumption of alcohol.
Specified conditions in females:

- Menopause.
- Subjects who underwent hysterectomy.

People of 30 to 40 yrs age having H.I.V and those from 41 to 70 yrs age having various diseases namely coronary artery disease, diabetes mellitus, hyper and hypo tension, tuberculosis and Alzheimer’s diseases were selected and placed into categories. In addition to this, women having ages from 41 to 60 yrs passing through menopause and underwent hysterectomy were considered.

Further, persons having certain unhealthy habits like smoking and consumption of alcohol were placed in another group with age from 61 to 70 yrs. These diverse groups were made three main groups:

**Group-I:** Subjects who were having any of the above mentioned parameters but under supplementation were denoted as “Supplemented group” (S.G). The supplemented group in this work received one multivitamin capsule per day with the following composition of the vitamins:

- Vitamin A – 2500 IU
- Vitamin D₃ – 200 IU
- Vitamin E – 5 mg
- Vitamin C – 50 mg
- Vitamin B₁ – 1 mg
- Vitamin B₂ – 1.5 mg
- Vitamin B₃ – 10 mg
- Vitamin B₆ – 1 mg
- Vitamin B₉ – 0.15 mg
- Vitamin B₁₂ – 1 mcg

**Group-II:** Subjects who were having any of the above listed parameters but not under supplementation were denoted as “Non supplemented group” (N.S.G).

**Group-III:** Subjects who were not having any of the above mentioned parameters were denoted as “Control group” (C.G).
**Collection of blood samples:** Under the medical supervision, the subjects are reported to the hospital and baseline parameters were taken which includes gender, age, body weight, blood group with Rh factor, hemoglobin and hematocrit percentages. Besides, working status, place of living (urban/rural) and economic status were recorded as shown in Annexure I. Then the blood samples (2 mL) were collected and subjected to centrifugation for 15 min in low speed REMI cooling centrifuge and the serum samples were collected in high density polyethylene vials.

**Determination of serum vitamin concentration by HPLC:** The concentration of Vitamin A (Retinol), Vitamin E (δ-Tocopherol), Vitamin D (Cholecalciferol), Vitamin K (Menaquinone), Vitamin C (Ascorbic acid), Vitamin B₁ (Thiamine), Vitamin B₂ (Riboflavin), Vitamin B₃ (Nicotinic acid), Vitamin B₆ (Pyridoxine), Vitamin B₉ (Folic acid) and Vitamin B₁₂ (Cyanocobalamin) were determined using HPLC.

The fat soluble vitamins were analyzed at a single run; next to this B-complex vitamins except B₉ were detected in second run, in last run was used to identify C and B₉. (It is necessary for the biological samples to undergo a pretreatment method, Solid Phase Extraction for removing the interfering components before injecting into HPLC).

**Instrumentation and chemicals:** Shimadzu liquid chromatography 2010 CHT consisting of a Shimadzu system controller, a Shimadzu Photo Diode Array detector, a Shimadzu auto sampler is shown in Figure 24 and system aligned HP Laser Jet printer is used. All necessary evaporation were performed in the evaporator. Organic solvents like methanol, acetonitrile and ammonium acetate were of HPLC grade. Water used throughout the analyses was from a millipore system. All vitamins used during this analysis were of analytical grade. The SPE cartridges used for sample clean-up were Supelclean LC-18 and LC-SAX.

**Preparation of standard solutions:** The aqueous stock solutions of the water soluble vitamins were prepared. The working standards, in the range
of 20-110 mg in 100mL, were prepared by appropriate dilution of the stocks. Folic acid was dissolved in an aqueous solution of 1 M NaHCO₃. The stock methanolic solutions of the fat-soluble vitamins were prepared. The working standards were prepared every day by appropriate dilution of the concentrated stock standard solutions. All solutions were stored in a refrigerator, covered with aluminum foil in order to protect them from light.

**Biological Sample preparation:** Aliquots of 40 µL pooled human blood serum were treated with 200 µL of CH₃CN in order to precipitate proteins. After vortex mixing for 2 min, the sample was centrifuged at 4000 rpm for 15 min. The supernatant was subsequently applied to the solid-phase cartridge after removal of the organic solvent by evaporation.

**Solid phase extraction procedure:** SPE optimization was performed in terms of retention and elution. Several different SPE cartridges were tested. Different washing and elution solvents were assessed. Optimum SPE conditions include Supelclean LC-18 cartridges (500 mg/3 mL), conditioned immediately prior to use with 1 mL of methanol and 1 mL of deionized water. After sample application vitamins were retained on the sorbent and subsequently eluted by 2 mL of an 85:15 v/v mixture of methanol: water. The eluate was collected in eppendorf tubes and evaporated to dryness using evaporator, at a low drying rate. The dry residue of the vitamin fraction was reconstituted in 100 µL of an aqueous solution of the chromatographic internal standard. 20 µL of each solution was injected into the HPLC column.

**Chromatographic conditions:** For the separation of the vitamins, a Phenomenex column, type Luna 3µC18 (2) (150mm × 4.60 mm, 3µm) coupled with a Phenomenex security guard pre-column, at 30 ºC temperature was used.

For the separation of the fat-soluble vitamins (A, D, E and K), the mobile phase of the HPLC system consisted of acetonitrile (solvent A) and methanol (solvent B). A simple linear gradient was used, starting from 50/50 (solvent A/solvent B) and ending after 20 min at 70/30 solvent A/solvent B). A 10 min equilibration time was observed between injections.
The mobile phase flow rate was 1.5 mL/min and the injection volume each time was 20 µL.

For the separation of B-complex vitamins except B₉, the mobile phase of the HPLC system delivered at a flow rate of 1.5 mL/min, consisted of A: 0.05 M CH₃COONH₄/CH₃OH (99/1) and B: H₂O/CH₃OH (50/50). A multi-step gradient was used, starting at an A: B v/v composition of 99:1 and remaining isocratic for 4 min. This composition was changed linearly to reach 100% of solvent B after 18 min and finally elution was performed isocratically for 8 min. A 30 min equilibration time was observed between injections. The injection volume was 20 µL.

For the separation of vitamins C and B₉, the mobile phase of the HPLC system consisted of buffer (solvent A) which is made up of hexane-1-sulfonic acid sodium salt in H₂O, tri ethyl amine with acetic acid and methanol (solvent B). A simple linear isocratic elution of 50/50 (solvent A/solvent B) was used. A 10 min equilibration time was observed between injections. The mobile phase flow rate was 1.5 mL/min and the injection volume each time was 20 µL[90, 91]. A series of experiments were performed in triplicate using HPLC to generate the chromatographic response surfaces of the selected samples using LC solutions version 2.1 software which were represented in Annexure II.

For all the runs performed detection was carried out with a photodiode array detector monitoring the eluent at 280 nm. The chromatographic conditions were chosen in terms of peak shape, column efficiency, chromatographic analysis time, selectivity, and resolution.

The determination of relative peak areas of the individual vitamin was carried out. Using these peak areas of the standard and samples (peak areas of all the vitamins of the samples were given in the Annexure III) and known concentration of the internal standard, concentration of the vitamins present in the sample were calculated individually. The concentration is expressed in terms of ng/mL.

**Calculation:**

\[
\text{Conc.} \text{sample} = \frac{\text{sample peak area \times std. wt. \times dilution \times purity of std.}}{\text{std. peak area \times dilution \times wt. wt. \times 100}} \times 100
\]
The outcome (vitamin concentration) of the above experimentation was statistically executed by artificial neural networks for optimizing and estimating the best possible levels of dietary vitamin supplements for the selected subjects. A nonlinear relationship exists between the chosen given input variables and the level of vitamin concentration as a response in this study. Several subjects had been optimized to acquire potential and desired responses using this technique.

The type of ANN used in this study was supervised neural networks (multi-layered perceptrons). These were specified by the network architecture, topology, neuron characteristics and learning rules they use.

The architecture of a supervised neural network defines the way the network is composed of its interconnected layers of neurons (nodes or processing elements). Here, it was the multi-layered feed-forward back-propagation network.

The number of neurons in each layer defines the network topology. The selection of an optimal neural-network topology is of critical importance for a successful application \cite{116}. In this case, there were input neuron layer, in which input (independent) data were fed forward to the hidden neurons in hidden layer(s), then to output neurons in output layer, which gives the network’s response. Each neuron is connected to every neuron in the preceding layer and a value for each neuron (activation function) is calculated as the weighed sum of all its inputs minus a threshold value (bias) for that neuron. At the output layer, the activation of the output neuron was compared with the observed (dependent) value for each case in turn and an overall error term was calculated for the entire data-set \cite{115}.

The number of neurons in the input and output layers was determined by the number of preselected input and output variables in each case. The input variables selected were the age of the specimen and the parameters (pathological conditions, unhealthy habits and other specified conditions) and the output variable is the vitamin concentration of the specimen. The number of hidden neurons can be considered as a design parameter, but there are no universal methods to determine their optimal number. This is usually done by the trail and error method and they were illustrated in Table 05.
Table 05: Number of neurons used in the hidden layer.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Number of Neurons in the Hidden Layer</th>
<th>Vitamins</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A</td>
</tr>
<tr>
<td>H.I.V</td>
<td>Alone</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>10</td>
</tr>
<tr>
<td>C.A.D</td>
<td>Alone</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>10</td>
</tr>
<tr>
<td>D.M</td>
<td>Alone</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>10</td>
</tr>
<tr>
<td>H.T</td>
<td>Alone</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>10</td>
</tr>
<tr>
<td>T.B</td>
<td>Alone</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>6</td>
</tr>
<tr>
<td>A.D</td>
<td>Alone</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>10</td>
</tr>
<tr>
<td>U. Hab.</td>
<td>Alone</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>10</td>
</tr>
<tr>
<td>M.C</td>
<td>Alone</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>11</td>
</tr>
<tr>
<td>H.C</td>
<td>Alone</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>8</td>
</tr>
</tbody>
</table>

The activation functions of the hidden and output layer were “tansig” and “purelin” respectively, by these the overall error term was calculated.

The network learns by virtue of a training algorithm, whereby the connection weights and biases were adjusted as a function of the root mean squared error at the end of each run through the data-set (epoch). The maximum number of epochs used was 600 in this study. The most commonly used algorithm was back propagation. The data-set is thus presented repeatedly to the network and this training process continues for either the maximum fixed number of epochs or until a pre-defined minimum root mean squared error was reached. Once trained, the connection weights and the neuron biases compromise the model parameters with a convessional equation or model. The training functions used for this study were trainscg, trainlm, trainbr and trainc.

**Trainscg (scaled conjugate gradient back propagation):** This training function updates weight and bias values according to the scaled conjugate gradient method. This can train any network as long as its weights, net input and transfer functions have derivatives function. Back propagation is used to calculate derivatives of performance (perf) with respect to the
weight and bias variables (X). The scaled conjugate gradient algorithm is based on conjugate directions, but search is performed in a line for each iteration.

**Trainlm (Levenberg-Marquardt back propagation):** This training function updates weight and bias values according to Levenberg-Marquardt optimization. It can train any network as long as its weight, net input and transfer functions have derivative functions.

**Trainbr (Bayesian regulation back propagation):** In this the weight and bias values are updated according to Levenberg-Marquardt optimization. It minimizes a combination of squared errors and weights and then determines the correct combination so as to produce a network that generalizes well. This process is called Bayesian regulation. This function can train any network as long as its weight, net input and transfer functions have derivative functions.

**Trainc (cyclical order incremental training):** In this the cyclical order of incremental training was done with learning functions. This trains a network with weight and bias learning rules with incremental updates after each presentation of an input. Inputs are presented in cyclic order for each epoch; each vector (sequence) is presented in order to the network with weight and bias values updated accordingly after each individual presentation.

The network can also be used to make predictions based on independent input data. The consistency of the network was given by the coefficient of determination, $R^2$ \cite{8, 47}. The ‘goodness of fit’ in both learning and testing was given by $R^2$, which describes the variance in the modeled variable that can be explained with the model \cite{116}. Thus, if the value of $R^2$ is nearer to 1, the model predicts every experimental point exactly and if the value of $R^2$ is zero then the fit does don’t exist at all.

Knowledge is acquired by the neural network by iterative adjustment of the synapse weights which is done by back propagation. In this after each presentation the input and output vectors are coupled together until the error between the output and target is minimized. The computation is carried out in parallel by the large number of interconnected neurons.
**Back propagation procedure:**

In training, when a set of inputs is given to a network, there the weights are pre-assigned randomly. Each neuron in the hidden and output layers first calculates the weighted sum of its inputs and passes the result through a transfer function to produce an estimated output that fits to the input data set. The result is compared to the corresponding desired values and the error is back propagated through the network to adjust the connection weights according to the learning rule. This procedure is repeated iteratively, until either the predetermined target, mean square of the output error (MSE) or a given maximum number of iterations is reached.

For this process the neural network computes the error derivative of the weights (W), i.e., to calculate the changes in error even the weight changes (increases or decreases) slightly. The algorithm computes the rate of error change with the activity level of a unit change (R). For output neurons, R is the difference between the actual and desired output. R is computed for hidden layer by multiplying the weights between the hidden and output layer by the R of those outputs and add the products. After calculating the R for the hidden layer, the other layers are also computed in the same fashion, moving layer to layer in the opposite direction through which the network activities propagates \[^{[P\cdot B/R37]}\]. This is why the network is called back propagation. After computing the R for each and every neuron, W is calculated for every incoming connection of the unit. W is the product of the R and the activity through the incoming connection. This process is repeated iteratively till the goal is reached. These entire neural network models described in this study was executed in Matlab/simulink version 7.0 and run under Microsoft windows XP environment.