

**IN VITRO-IN VIVO CORRELATIONS OF SELF EMULSYFING DRUG
DELIVERY SYSTEMS COMBINING THE DYNAMIC LIPOLYSIS MODEL
AND NEURO-FUZZY NETWORKS**

Dimitrios G. Fatouros^{*, 1, †}, Flemming Seier Nielsen^{1, †}, Dionysios Douroumis²,
Leontios J. Hadjileontiadis^{*}, Anette Mullertz¹

¹Department of Pharmaceutics and Analytical Chemistry, The Faculty of of
Pharmaceutical Sciences, University of Copenhagen, Universitetsparken 2, DK-2100
Copenhagen, Denmark

²Greenwich University, Medway Campus, Central Avenue, Chatham Maritime, Kent
ME4 4TB, UK

³Department of Electrical and Computer Engineering, Aristotle University of
Thessaloniki, Thessaloniki, Greece

*Corresponding authors: Dr Leontios J. Hadjileontiadis, Dr Dimitrios G. Fatouros
e-mail: leontios@auth.gr, df@dfuni.dk

Contact author: Dr Dimitrios G. Fatouros

Tel: + 45 35 30 62 88

Fax: + 45 35 30 60 30

[†]These authors contributed equally to the study

ABSTRACT

The aim of the current study was to evaluate the potential of the dynamic lipolysis model to simulate the absorption of a poorly soluble model drug compound, probucol, from three lipid based formulations and to predict the *in vitro-in vivo* correlation (IVIVC) using neuro-fuzzy networks. An oil solution and two self micro and nano-emulsifying drug delivery systems were tested in the lipolysis model. The release of probucol to the aqueous (micellar) phase was monitored during the progress of lipolysis. These release profiles compared with plasma profiles obtained in a previous bioavailability study conducted in mini-pigs at the same conditions. The release rate and extent of release from the oil formulation were found to be significantly lower than from SMEDDS and SNEDDS. The rank order of probucol released (SMEDDS~SNEDDS>oil formulation) was similar to the rank order of bioavailability from the *in vivo* study. The employed neuro-fuzzy model (AFM-IVIVC) achieved significantly high prediction ability for different data formations (correlation greater than 0.91 and prediction error close to zero), without employing complex configurations. These preliminary results suggest that the dynamic lipolysis model combined with the AFM-IVIVC can be a useful tool in the prediction of the *in vivo* behavior of lipid based formulations.

Keywords: *In vitro-in vivo* correlations (IVIVC); *in vitro* dynamic lipolysis model; self emulsifying drug delivery systems, mathematical modeling, neuron-fuzzy networks

1. Introduction

In recent years, much attention has turned to lipid based formulations with the aim of improving the oral bioavailability of poorly water soluble drugs. Lipid based formulations encompass a diverse group of formulations, very different in physical appearance, ranging from a simple triglyceride vehicle to more sophisticated formulations such as self-emulsifying drug delivery systems [1] and [2].

Most lipid based formulations are designed to deliver the entire dose in solution thereby bypassing the dissolution process in the gastro-intestinal (GI) tract, which has been recognized as one of the main prerequisite for the efficiency of these formulations [3]. However, the GI tract is a complex organ where the environment (e.g. liquid volume, pH, ionic strength and concentration of bile salt) varies depending on dietary state. In addition, lipid based formulations containing hydrolysable lipids like triglycerides are prone to lipolysis by gastric and pancreatic lipase in the GI tract.

The use of an *in vitro* lipolysis model has been proposed as an approach to probe solubilisation in the aqueous phase during the progress of enzymatic degradation of lipid based formulations [4], [5], [6], [7], [8] and [9]. However, none of these studies directly compared lipid formulations with the same composition, relating the absorption of a drug from a lipid based formulation to the solubilization of drug in the micellar phase during lipolysis.

In the present study we have attempted to relate the *in vivo* performance of three different lipid based formulations using the *in vitro* solubilization data obtained from a lipid digestion model.

Additionally a mathematical model was used to correlate the *in vitro* and *in vivo* data. Modelling of *in vitro-in vivo correlation* (IVIVC) has been reported in the literature [10], [11], [12] and [13] as an effort to successfully predict *in vivo* drug concentration-time profiles using the *in vitro* dissolution data. The degree of the assumed correlation between *in vivo* absorption and *in vitro* dissolution properties of drug product reflects the existence or not of a point-to-point relationship between *in vitro* dissolution and *in vivo* input rate [14] and [15]. Efficient performance of an IVIVC model can contribute to the optimization of the drug formulation during clinical phases I and II, achieving the target concentrations of drug in plasma. Furthermore, it can support changes made to

production procedures (clinical phase III), such as variations or scale-up, and during the post approval period [16].

Many of the IVIVC models adopt a linear approach, i.e., they relate a parameter of a time point descriptive of the dissolution to a parameter or a time point descriptive of the pharmacokinetic absorption [10], [11] and [12]. However, these IVIVC models are unable to account for any nonlinear relationship between the *in vitro* dissolution data and the *in vivo* pharmacokinetics and the intrinsic variability of the parameters involved in the IVIVC modelling procedure [17], [18] and [19]. Newer IVIVC modelling tools involve artificial intelligence, i.e., models using artificial neural networks (ANNs). The latter have the ability to incorporate a large number of possible variables and relationships without a predefined model structure [20], [21] and [22]. This characteristic has fostered the use of ANNs in pharmacokinetics and pharmacodynamics [23], [24] [25] and [26] and in product development [27]. A thorough exploration of the efficiency of the ANN-IVIVC modelling has been presented by Dowell *et al.* [28], who developed a methodological approach to ANN-IVIVC modelling and examined its feasibility based on different ANNs configurations and data formats. Furthermore, stochastic IVIVC modelling based on Bayesian approach [29] has been recently proposed [30]. In that IVIVC model, compartment models and related prior information are used, whereas the results are interpreted as probability distributions of the parameters; hence, predictions of plasma concentration profiles with probability distributions based on *in vitro* dissolution data are feasible.

In this work, an extension to the ANN-based optimization technique is introduced based on Neuro-Fuzzy Modelling (NFM). The concept of NFM has emerged in recent years as researchers have tried to combine the transparent, linguistic representation of a nonlinear system with the learning ability of ANNs. The adopted neuro-fuzzy model consists of an Adaptive neuro-Fuzzy Modeller (AFM) [31] and [32] which combines ANNs with fuzzy logic (FL) [33] to model nonlinear complex problems, such as IVIVC modelling. Fuzzy logic is a powerful tool, which has been successfully used in many signal processing fields, like system modelling and control, pattern recognition, detection, de-noising, prediction [34] and [35]. Unlike the Boolean logic, FL allows the input and output values to a fuzzy inference model to belong to multiple sets with different degree of membership

in each set defined by a particular membership function [33]. This facilitates the idea that a nonlinear system can be approximated by softly merging locally linear systems avoiding discontinuities if the system state moves from one local model to another [36]. This fuzzy transition is achieved using the membership functions to calculate the validity of the different local models for a certain state [37]. The resulting structure of the fuzzy system has the appearance of a network; hence, the learning methods of an ANN can be easily applied to form a neuro-fuzzy model with favourable characteristics.

To our knowledge, the AFM approach has not been previously used in the context of IVIVC modelling; hence, the aim of the proposed study is to examine the predictive potential of the AFM-IVIVC approach, based on different data formats and using a relative small set of IVIVC data for training and prediction.

2. Materials and Methods

2.1 Materials

Pancreatin (porcine), bile extract (porcine) and sesame oil were purchased from Sigma–Aldrich, USA. 4-bromobenzenboronic acid (BBBA) was purchased from Lancaster, Germany. Cremophor RH 40 was purchased from BASF, Germany and Maisine 35-1 from Gattefossé, France respectively. Phosphatidylcholine (Lipoid E PC, purity >98%) was kindly donated from Lipoid GMBH, Germany. The water used was obtained from a Milli-Q-water purification system, Millipore, MA, USA. All other chemicals were of analytical grade.

2.2 Preparation of lipid formulations

The lipid based formulations tested were: an oil solution and two self-emulsifying drug delivery systems which were prepared as described previously [38] according to the following composition : SNEDDS [Sesame oil : Maisine 35-1 : Cremophor RH 40 : Ethanol, 30 : 30 : 30 : 10] , SMEDDS [Sesame oil : Maisine 35-1 : Cremophor RH 40 : Ethanol, 26.7 : 26.7 : 26.7 : 20] and Oil [Sesame oil : Maisine 35-1 : Ethanol, 45 : 45 : 10] expressed in w/w. The particle size for the SNEDDS formulation was 45.0 ± 3.4 nm and for the SMEDDS formulation was 4.58 ± 0.84 μ m respectively.

2.3 Lipolysis medium

In the current study a concentration of 20 mM of bile salts and 4 mM of Phosphatidylcholine (Lipoid E PC, purity >98%) was used. This simulates the fed state conditions in the GI tract [39]. The composition of the digestion buffer was 150 mM NaCl, 2mM Trizma maleate pH 6.5 and the final volume 300 mL. The Trizma-maleate concentration was chosen to be low (2 mM) to secure that ionized fatty acids are able to change pH in order to trigger the addition of NaOH in the pH-Stat [4]. We chose pH 6.5 as a compromise between the optimum for the pancreas lipase, which is between 6 and 10 [40] and the measured duodenal pH in the fed state, which is around 5.0-5.5 [41].

2.4 Preparation of lipase suspension

The lipase suspension was prepared in accordance with the method described previously [5] to give an activity of 800 USP units/ml. Briefly 16.6 g of pancreatin was weighted accurately, suspended in 110 mL of Millipore water at 37°C and mixed thoroughly. The suspension was centrifuged for 7 minutes at 4,000 rpm at 37°C and the pH of the supernatant was adjusted to 6.5 using 1.00 M NaOH. 100 mL of the pH-adjusted supernatant was used for the study. In order to minimize denaturation, the time spent on preparing the suspension did not exceed 15 minutes. The lipase activity of pancreatin was determined in accordance with United States Pharmacopeia 26, 2003 [42].

2.5 In vitro digestion study

The experimental set-up of the *in vitro* lipolysis model has been described previously [5]. Briefly it consists of a thermostated double wall reaction vessel, the pH-stat with the auto burette for the addition of NaOH, a peristaltic pump for the addition of CaCl₂ and the computer with the software for the titration experiments. The temperature is monitored during the experiment with a thermocouple. The experiment is performed under continuous agitation at 37°C. A pH-stat titrate control the volume of NaOH to maintain the initial pH. The number of OH⁻ ions present in the volume of the titrant can be equated with the fatty acid liberation caused by lipolysis.

The bile salt, phosphatidylcholine and buffer of the lipolysis medium were mixed in a thermostatically controlled vessel (37 °C). 3 g of the each formulation was added to the obtained 300 mL of the bile salt medium, giving a final concentration of 1% w/v. The pH of the obtained medium was adjusted to 6.5 with 1.0 M NaOH. The lipolysis process was initiated by adding the 100 ml lipase suspension. In the dynamic lipolysis model the continuous addition of calcium chloride solution serves to control accumulation of FA on the surface of the emulsions particles by forming insoluble calcium FA soaps, which precipitate thus removing FA from the system [43]. Therefore a continuous addition of a 0.5 M Ca²⁺ solution was started at time zero with a dispensing rate of 0.045 mM/min. Throughout the study pH was kept constant at 6.5 by means of a pH-stat (Titrino 718 with burette from Metrohm, Switzerland). The software used was Tinet, version 2.3 also from Metrohm.

At 0, 15, 30, 45 and 60 minutes, 20-ml samples were withdrawn and the lipase was inhibited immediately with a 4-bromobenzenboronic acid (BBBA) Lancaster, Germany solution as described previously [5]. The time zero sample was taken just after adjustment of the pH in the medium. The samples were centrifuged at 40,000 rpm for 135 min [5] and probucol concentration in the aqueous phase was determined by HPLC [44].

2.6 In vivo studies

The bioavailability study was conducted as crossover study in male Göttingen mini-pigs fed a high fat meal 30 minutes prior to treatment [44]. The meal consisted of standard mini-pig diet enriched with 20 w/w % olive oil to obtain a high fat meal (50% energy from fat).

2.7 Statistics

The solubilization of probucol from SNEDDS, SMEDDS and oil formulations were compared with One Way Analysis of Variance for each time point and in cases of significance ($p < 0.05$) differences between formulations were allocated by multiple comparison with an un-paired t-test (Student-Newman-Keuls Method) using SigmaStat version 3.1 (Systat Software Inc. USA).

2.8 Functional Principles of Neuro-Fuzzy modelling (NFM)

The AFM belongs to the family of NFM. In general, neuro-fuzzy models belong to the category of empirical data-based models (EDMs). These models rely on the fact that the intrinsic features of the observed interactions of a complex system and their mutual interrelations can be learned from the data using a great number of simultaneously co-operating simple processing units or operations. This approach allows the extraction of information (knowledge) from these low-level data into other forms that might be more abstract [45]. EDMs that make use of fuzzy inference system (FIS) combined with adaptive networks provide a neuro-fuzzy network that consists of nodes and directional links through which the nodes are connected. Part or all of the nodes are adaptive; hence, each output of these nodes depends on the parameters pertaining to this node. The learning rule specifies how these parameters should be changed to minimize a prescribed error measure [46]. In a neuro-fuzzy network, the synergism of ANNs and FL manages to model the structure of complex systems by extracting the necessary knowledge from pairs of crisp input-output data. In fact, an ANN can approximate a function, but it is impossible to interpret the result in terms of natural language. The fusion of ANNs and FL in NFM provide learning as well as readability. On the basis of the FL technology, the model can be linguistically described by means of input-output parameterized variables and well-defined set of IF/THEN fuzzy rules (namely rule-base). A general structure for a fuzzy rule is: “IF <*antecedent*> THEN <*consequent*>”, for example: “IF *speed is high* AND *distance is small* THEN *brake force is high*”. In this example, *speed*, *distance* and *brake force* are fuzzy variables and *high* and *small* are fuzzy sets. Fuzzy sets are linguistic terms that are expressed in an exact mathematical way by using the concept of *membership function*, which represents the degree of truth of an assertion (e.g. pressure is high) in a y-axis range [0, 1]. Membership functions can have several shapes, the most common being Gaussian, trapezoidal, triangular and sigmoidal, and their x-axis range forms the so-called *universe of discourse* [34].

The human-perceived information of the set of fuzzy rules is all encoded, at the mathematical level, by means of ‘fuzzy’ representations, which do not pursue precision. On the other hand, the ANNs technology, by means of the precise input-output values,

identifies the parameters involved in the model, by training a generic model to adaptively approximate the relationship between the input-output data [47]. Thus, by means of its structure, a neuro-fuzzy network manages not only to deal with the uncertainties of complex systems, but it can provide their model with transparent and interpretable structure. Furthermore, based upon this modelling, a neuro-fuzzy network manages to generalize; hence, it produces predictive outputs when presented with new ‘proper’ input. The theoretical details of the neuro-fuzzy modelling can be found in [46], [47]. Moreover, a simplified introduction regarding the general issues of FL modelling, fuzzy sets, membership functions and fuzzy clustering is provided in [34]. However, relevant features and context that refer to the adopted means of neuro-fuzzy modelling, i.e., AFM [32], are described in the subsection that follows and in the Appendix.

2.9 Structure of AFM

AFM is a tool that easily allows obtaining a model of a system based on FL data structure, starting from the sampling of a process/function expressed in terms of input-output values pairs (patterns). Its primary capability is the automatic generation of a database containing the inferencing rules and the parameters describing the membership functions. The generated FL knowledge base represents an optimised approximation of the process/function provided as input. AFM is expected to reveal the correlation between *in vitro* inputs (% dissolved) and *in vivo* outputs (pharmacokinetic observations) and generalize, i.e., provide efficient predictions of the output when presented with new input values. To infer this response values, AFM is trained to evaluate the relation between *in vitro* inputs and *in vivo* outputs. However, this initially unknown relation is hidden within the empirical data that are obtained from experiments. Therefore, AFM training is an equivalent procedure to learning from empirical data. This is achieved with the help of the learning capabilities of ANNs to extract a fuzzy rule base automatically from input-output data [48]. In fact, the NFMs can automatically identify the fuzzy rules and tune the membership functions by modifying the connection weights of the networks using the back-propagation algorithm [49]. The AFM used here adopts a six-layer feed-forward network structure [31] and belongs to the first of the three types of NFMs introduced by Horikawa *et al.* [50]; it is characterized by consequents of a constant type

and its schematic structure is depicted in Figure 1. During training, at each level, the parameterized nodes perform specific functions of the incoming signal [50] (see Appendix for some mathematical details). Briefly, this type of NFM has specially designed structures to make the connection weights of the networks correspond to the parameters of the fuzzy inference. Through the learning with the back-propagation algorithm, NFM can identify the fuzzy rules and tune the membership functions automatically [50]. One problem that needs to be addressed when using neuro-fuzzy networks for automatic rule extraction is the exponential dependence on the number of inputs. For example, when a network is structured for 8 inputs and 3 fuzzy sets per input, a base of $3^8=6561$ rules is generated. This shortcoming can be overcome by using appropriate clustering algorithms, like ‘the winner takes all’ [51] that efficiently reduce the number of rules in a knowledge base.

2.9.1 Implementation of the AFM-IVIVC scheme

The selection of different formatting of the input-output association defines the structural characteristics of the AFM-IVIVC scheme, as it varies the number of its inputs and outputs, accordingly. Moreover, different types of pattern files are constructed from the available experimental data to accommodate for each association characteristics.

The dataset reported in this study included solubilization values from three extended-release formulations (OIL, SMEEDS, SNEDDS) with five solubility time points, i.e., 0, 0.25, 0.5, 0.75 and 1 hour, each, at which three batches were tested per formulation. Each formulation was administered to six pigs in a crossover trial. The drug plasma concentrations were sampled at 12 time points, i.e., 0, 0.75, 1.5, 2.25, 3, 4, 5, 6, 8, 12, 24, and 48 hours, following oral formulation administration. The dataset from the first two formulations, i.e., OIL and SNEDDS, was used in the training procedure of the AFM-IVIVC, whereas the third one (SMEDDS) was kept as a validation set; hence, success of the AFM-IVIVC was based on the prediction of the validation profile.

Similarly to the procedure introduced by Dowell *et al.* [28], different types of pattern files constructed from the same data were selected for evaluation. These pattern files, namely ASSOCIATION-1 through ASSOCIATION-4, included different formatting of the input-output association [28], and a diagram of their structure is presented in Figure 2. In the

latter, the structure of the input-output relationship is depicted, as well as the data format across associations to create a pattern file is shown via subscripts. Moreover, a summary of the constructed pattern files is presented in Table 1.

ASSOCIATION-1 refers to a functional relationship of the data employing an input-output association that uses all of the pharmacokinetic concentration values from a pig as an output set associated with an input set that consists of the solubilization profile from an individual formulation. The resulting pattern file contains each pharmacokinetic observation set associated with each of the three solubilization profiles. In our case, this association involves 5 inputs and 12 outputs, hence, it contradicts the output configuration of the AFM, i.e., the number of outputs should not be greater than four [32]. To overcome this contradiction we split it to three subnets with 5 inputs and 4 outputs each, creating 3 training datasets for each output set (5-1:4, 5-5:8, 5-9:12), so the estimation of the pharmacokinetic observations at various time points should employ the relevant subnet (Figure 2(a)).

ASSOCIATION-2 is similar to ASSOCIATION-1, including the complete kinetic set of solubilization values for each tested formulation, but in this case each is associated with a single respective pharmacokinetic output, with the input-output association lines of the pattern file forming a pharmacokinetic time sequence. Moreover, the pharmacokinetic time point is also included as an input (Figure 2(b)).

ASSOCIATION-3 refers to an input-output association consisted of each *in vitro* value as an input associated with each *in vivo* output, excluding pharmacokinetic observations with no direct association with solubilization observations. The time of observation is also added as an input (Figure 2(c)).

ASSOCIATION-4 is a sequential time series and includes previous solubilization values as inputs. The output consists of the pharmacokinetic concentration value, whereas the inputs are the pharmacokinetic time point and all the dissolution values that precede that point in time. Dissolution values occurring after that pharmacokinetic time point are set to zero in the pattern file and are interpreted as null inputs by the model (Figure 2(d)).

The input/output pairs of the aforementioned associations were all presented to the system during AFM-IVIVC training. Learning is implemented in epochs in order to define the values of the premise and consequent parameters by minimizing the *Mean-*

Quadratic-Error (MQE) [32]. Each epoch foresees two passes: a forward pass of the signal, where the premise parameters are kept fixed and the consequent parameters are calculated by the least squares method, and a backward pass, where the consequent parameters are kept fixed and the premise parameters are updated by the gradient descent method [32].

In all realizations of the AFM-IVIVC according to the four associations three membership functions of a Gaussian-bell shape [40], uniformly distributed in their universe of discourse, and three linguistic variables, i.e., ‘low’, ‘medium’, and ‘high’, were defined as the initial conditions of AFM-IVIVC, i.e., before the training procedure. The target *MQE* for the training procedure was set to 10^{-4} . The data configuration according to the input-output relationship of the four associations was performed using Matlab (Version 7.1, The Mathworks, Inc., Natick, MA, 2005), whereas the AFM-IVIVC was implemented using the AFM 2.0 software (SGS-Thomson Microelectronics, 1998) [32].

2.9.2 Performance Indices

Similarly to the work of Dowell *et al.* [28], the performance of the ANF-IVIVC was evaluated using the performance indices defined below:

$$\text{Correlation coefficient:} \quad R^2 = \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2} \quad (10)$$

$$\text{Mean Prediction Error:} \quad MPE = \frac{1}{N} \sum (\hat{y} - y) \quad (11)$$

$$\text{Mean Absolute Error:} \quad MAE = \frac{1}{N} \sum |y - \hat{y}|, \quad (12)$$

where y denotes the actual observation, \hat{y} is the prediction of the AFM-IVIVC, \bar{y} corresponds to the average observation and N is the number of observations. In addition, the statistical significance of the estimation of the correlation coefficient R^2 is measured with the probability of false hypothesis p , with values of $p < 0.05$ denoting statistical significance of the estimated R^2 value. Moreover, the ratio of R^2 between the predictions and training pattern files is also used as an indicator of possible network memorization.

3. Results and discussion

3.1 Lipolysis rate

The consumption of NaOH during the study, reflecting the progress of lipolysis, is depicted in Figure 3. The values presented have been corrected by subtracting the amount of NaOH consumed when carrying out experiment without any formulation present. Initial lipolysis rates (up to 5 minutes) are almost identical for all formulations. After 60 minutes significantly more NaOH had been consumed during hydrolysis of the oil formulation compared to the SMEDDS and SNEDDS formulations. In absolute terms $1.85 \pm 0.15\text{g}$, $0.94 \pm 0.14\text{g}$ and $1.04 \pm 0.17\text{g}$ of the lipid phase was hydrolyzed after 60 minutes. However, when calculated as the percentage of the total amount of hydrolysable ester bonds, the oil, SMEDDS and SNEDDS formulations were hydrolyzed to the same extent of $68.7 \pm 5.6\%$, $58.5 \pm 0.9\%$ and $57.8 \pm 9.2\%$, respectively (not significant different).

3.2 In vitro digestion of lipid formulations

The percentages of probucol released from the oil solution, SNEDDS and SMEDDS into the aqueous phase versus time are depicted in Figure 4. The initial release values (mean \pm SD), before the onset of lipolysis were $27.2 \pm 0.3\%$ for the SMEDDS, $16.2 \pm 0.4\%$ for the SNEDDS and $8.2 \pm 2.1\%$ for the formulation. The initial release values reflect a partition of probucol between formulation and the mixed micelles and other intermediate products present in the aqueous phase.

From 0 to 15 min a steep release of the drug was noticed for the SMEDDS and SNEDDS formulations. The release from the SMEDDS and SNEDDS formulations reaching $95.8 \pm 0.8\%$ and $71.5 \pm 1.5\%$ of the total amount at 15 minutes, followed by a levelling off for the SNEDDS formulation up to 60 minutes (significantly different values at 15 and 60 minutes) and a plateau for the SMEDDS formulation (no significant difference between 15 and 60 minutes). There was no significant difference between 45 and 60 minutes with regard to the release of probucol from SMEDDS and SNEDDS formulations ($p > 0.05$).

The release profile for the oil formulation increases from $8.2 \pm 2.1\%$ at time point 0 to a plateau of $16.7 \pm 9.0\%$ at 45 minutes from where it decreased, but not significantly, to $15.9 \pm 7.7\%$ at 60 minutes ($p > 0.05$).

Despite the same relative level of hydrolysis, lipolysis of the oil formulation releases significantly lower amounts of probucol to the aqueous phase compared to the SNEDDS and SMEDDS ($p < 0.05$). This could be attributed to an up-concentration of probucol in the remaining non-hydrolysed lipid phase [9]. At all sampling times, more non-hydrolysed lipid is present for the oil formulation. In addition, the Cremophor in the SEDDS formulations might partition into the aqueous phase during lipolysis and aid an increased solubilisation of probucol during lipolysis of the SEDDS formulations.

The release profiles of probucol in the lipolysis model were compared with data obtained from a bioavailability study conducted in mini-pigs fed a high fat meal (standard mini-pig diet enriched with 20 w/w % olive oil to obtain 50% energy from fat) 30 minutes prior to treatment [44].

3.3 In vivo studies

The mean plasma concentration-time profiles following oral administration to fed mini-pigs are shown in Figure 5. No significant differences were found between the determined T_{max} values from the orally dosed formulations in the fed minipigs. The SNEDDS and SMEDDS formulations had a median T_{max} of 8 hours, whereas the oil formulation had a slightly longer median T_{max} of 12 hours. The SNEDDS and SMEDDS formulations exhibited the highest C_{max} . The C_{max} for the oil solution was significantly lower than for SMEDDS.

No significant differences were observed between the determined AUC_{0-48h} or the relative bioavailability in the minipigs fed a high fat meal. The SMEDDS formulation exhibited the highest relative bioavailability, which was fixed at 100% for use as a reference. The SNEDDS formulations exhibited a relative bioavailability equivalent to that of SMEDDS (97%), whereas the relative bioavailability of the oil solution was 61%.

The in vivo findings show that the SNEDDS and SMEDDS formulations exhibit the same T_{max} , but C_{max} and relative bioavailability are slightly higher for the SMEDDS. These findings are in accordance with the profiles obtained from the in vitro lipolysis model.

The in vitro release profile for the oil formulation reach a plateau later (after 30 minutes), and at a lower level than the release profiles for the SNEDDS and SMEDDS

formulations. This is in accordance with the in vivo findings showing that the oil formulation exhibit slightly longer T_{\max} , lower C_{\max} and relative bioavailability compared to the SNEDDS and SMEDDS.

Using the area under the curve for the release profiles in ranking the lipid based formulations results in the following order: SMEDDS ~ SNEDDS > oil in vitro. The same trend was observed in vivo as well.

Although there was no significant difference between the different formulations tested in vivo, there was a trend towards higher absorption values for SNEDDS and SMEDDS. This is also the case for the in vitro studies emphasizing that important information may be gained from the in vitro dynamic lipolysis model.

3.4 Mathematical modelling

Each configuration of the AFM-IVIVC based on the four associations was trained as described in the methodology and inputs from the training and validation pattern files were applied to the trained networks resulting in predicted output values; these were then compared to the actual observations. Shown in Table 2 are the results of the performance indices (see (10)-(12)), for both the training and validation pattern files for each configuration of the AFM-IVIVC scheme. These results are measured by the precision and bias of the predicted outputs from the training and validation data. Clearly, Table 2 reflects a success of each AFM-IVIVC configuration when applied to this particular set of IVIVC data. In fact, the AFM-IVIVC attempted to account for the determination of a mean concentration curve based on the information contained in the solubilisation kinetics and for the variability in the pharmacokinetics due to the variability in the solubilisation kinetics. Nevertheless, as it is noted in Table 2, ASSOCIATION-3 did not provide statistically reliable correlation results, as the structure of this configuration combined with the experimental set up resulted in a very limited number of valid cases (two only). In all other cases, the estimated correlation was statistically significant ($p < 0.001$) exhibiting R_T^2 and R_V^2 values greater than 0.94 and 0.91 for the training and validation set, respectively, resulting in a R_V^2 / R_T^2 ratio greater than 0.96. Moreover,

in all cases, the *MPE* and *MAE* were found close to zero for both training and validation sets.

Figures 6-9 depict the comparison between the actual pharmacokinetic observations and the predicted ones by the AFM-IVIVC for each configuration based on ASSOCIATION-1 through ASSOCIATION-4, respectively. In all these figures, the (a) part corresponds to the case of the analysis of the training data set, whereas the (b) part corresponds to the case of the analysis of the validation data set.

Looking at Figure 6, it is clear that both in Figure 6 (a) and (b), the predicted output of the AFM-IVIVC follows the mean concentration curve of PK (shown as a solid line), exhibiting high correlation ($R_{T,V}^2 > 0.99$). These results show that in the case of ASSOCIATION-1, where the data are formatted as a functional relationship, the AFM-IVIVC successfully determines the shape of the mean PK, despite the noticeable variability in PK observations. Moreover, as it is deduced from Figure 6 (a) and (b), the subnets involved in this configuration perform equal well, providing reliable predicted output for each region of pharmacokinetic time points (denoted as triangles with three different orientations), as justified by the corresponding small values of *MPE* and *MAE* (see Table 2).

In Figure 7, which corresponds to the case of ASSOCIATION-2, a performance similar to Figure 6 is noticed, as, in general, the predicted output is distributed around the curve of the mean of the observed PK values (denoted as solid line). However, there is a difference in the concentration of the predicted values around the mean PK between the training (Figure 7(a)) and validation (Figure 7 (b)) data sets. Clearly in the case of the training data set (Figure 7(a)), the predicted values are less dispersed, following the shape of the mean PK curve, whereas in the case of the validation data set (Figure 8 (b)), the predicted values are more dispersed, yet successfully defining the shape of the mean PK curve. This is also reflected in the difference seen in the corresponding correlation values (see Table 1), i.e., $R_T^2 = 0.973$ and $R_V^2 = 0.941$, and in the small increase of *MPE* and *MAE* of the validation data set compared to the ones of the training data set (see Table 2). Nevertheless, the general performance of the AFM-IVIVC under the configuration of ASSOCIATION-2 is still high, accounting both for the determination of the mean concentration curve and the variability of PK values.

As it is already mentioned, the configuration of ASSOCIATION-3 combined with the experimental set up resulted in limited valid cases for the AFM-IVIVC. Figure 8 illustrates the predicted output for this configuration corresponding to two pharmacokinetic time points, i.e., 0 and 0.75 h. Both in the case of training (Figure 8(a)) and validation (Figure 8 (b)) data sets, the predicted values account for the variability of the observed PK values, resulting in small values of *MPE* and *MAE* (see Table 2).

The performance of AFM-IVIVC for the configuration of ASSOCIATION-4 is illustrated in Figure 9. In this case, data are formatted as a memorative pattern file and this pattern file provides the advantage of being a generalized format with a single output, which also allows the AFM-IVIVC to incorporate relationships from the previous inputs. However, this pattern file seems unable to use information from dissolution values with a time of dissolution greater than the corresponding time of the PK observation [ref. ANNs]. Similar to the ASSOCIATION-1 and ASSOCIATION-2, the predicted values of the AFM-IVIVC of ASSOCIATION-4 shape quite well the curve of the mean PK, exhibiting high correlation values for the training (Fig. 5(a)) and the validation (Figure 9(b)) data set, i.e., $R_T^2 = 0.945$ and $R_V^2 = 0.914$, respectively, with better performance for the case of training than the validation data set (see also the corresponding values of *MPE* and *MAE* of Table 2). In general, a higher dispersion is seen in the distribution of the predicted data when comparing ASSOCIATION-4 with ASSOCIATION-1 and ASSOCIATION-2, especially for the predicted values at pharmacokinetic time points lying within the time span of 1.5-8 hours for the case of the training set and for those at the pharmacokinetic time points of 24 and 48 hours for the case of the validation set.

The setting of the AFMs incorporated in the AFM-IVIVC scheme was selected as the standard one (three membership functions of a Gaussian-bell shape corresponding to the linguistic variables of 'low', 'medium', and 'high'). An example of the estimated membership functions corresponding to each input of the AFM-IVIVC of ASSOCIATION-4 is depicted in Figure 10. Other setting, like trapezoid type and greater number of membership functions, was tested resulting in non significant performance enhancement. The number of the estimated fuzzy rules of the AFM-IVICV was 729 in all associations but ASSOCIATION-3 where it was nine; in ASSOCIATION-1, at each subnet 243 fuzzy rules were corresponded. Moreover, 1000 epochs as a mean value

across the four associations were performed for achieving the target MQE during the training procedure. Apparently, increase in the number of inputs and outputs would lead to a more complicated network structure, which in turn would increase the level of the structural complexity, thus decreasing the likelihood of obtaining good solution [28]. The structure of the AFMs incorporated in the proposed AFM-IVIVC scheme has led to a moderate complexity, securing, at the same time, an efficient prediction performance.

The results presented so far denote that the AFM-IVIVC exhibits a high performance in the prediction of the pharmacokinetic values at different time points based on the information contained in the dissolution kinetics. Despite the difference between the data set used in the current study and in the one by Dowell *et al.* [28], a rough comparison between the performance of the AFM-IVIVC and the ANN-IVIVC [28] could be conducted, as the same data configuration and same performance indices were considered in both works. In particular:

1. From the comparison of the performance indices presented in Table 2 (with the exception of $R_{T,V}^2$ for the case of ASSOCIATION-3) with those presented in [28-Table 3] it is clear that the proposed AFM-IVIVC scheme exhibits a higher performance than the ANN-IVIVC [28], as it results in R_T^2 and R_V^2 values (worse case ASSOCIATION-4) greater than 0.94 and 0.91, respectively, when the ANN-IVIVC results in R_T^2 and R_V^2 values (for the best network configurations) greater than 0.85 and 0.77, respectively. Moreover, the estimated errors, i.e., MPE and MAE are significant lower for the case of AFM-IVIVC ($\text{mean}|MPE|_{\text{training}} = 0.009$; $\text{mean}|MAE|_{\text{training}} = 0.176$; $\text{mean}|MPE|_{\text{validation}} = 0.062$; $\text{mean}|MAE|_{\text{validation}} = 0.176$) than those for the case of ANN-IVIVC ($\text{mean}|MPE|_{\text{training}} = 1.455$; $\text{mean}|MAE|_{\text{training}} = 4.716$; $\text{mean}|MPE|_{\text{validation}} = 2.117$; $\text{mean}|MAE|_{\text{validation}} = 5.353$). Furthermore, the $\frac{R_V^2}{R_T^2}$ ratio shows a higher consistency in its convergence to one for the case of AFM-IVIVC (values from 0.967 to 0.996) than for the case of ANN-IVIVC (values from 0.812 to 1.910).
2. As it have been shown by the results, the AFM-IVIVC incorporates quite well for different data formations despite its AFM structure, showing a successful

performance in the identification of the IVIVC, whereas ANN-IVIVC performs well for ASSOCIATION-1 and ASSOCIATION-4, only, showing a variability in its performance across its eight different ANN architectures employed in the structure of the ANN-IVIVC.

3. The AFM-IVIVC expands the structure of ANNs using fuzzy logic, thus, providing more flexibility in the identification of nonlinear relationships like the IVIVC; hence, resulting in more efficient performance than the ANN-IVIVC, despite the relative smaller set of IVIVC data used for training and prediction in the case of AFM-IVIVC.

The promising results presented here pave the way for applying AFM-IVIVC under different types of data and experimental design in order to further justify the potential of the proposed approach to successfully reveal the correlation between *in vitro-in vivo* data; such approach is an ongoing process within our group.

CONCLUSIONS

Our results demonstrate that the *in vitro* dynamic lipolysis model is a potential tool in evaluating new oral lipid based formulations and predicting their *in vivo* behavior and lays the framework for future work. Moreover a novel approach in the prediction of a mean *in vivo* plasma concentration profile using dissolution kinetics using neuro-fuzzy modelling has been presented in this work. As it has been demonstrated, the introduced AFM-IVIVC scheme exhibited efficient predictive performance for different data formations of this data set, without employing complex configurations. The proposed AFM-IVIVC model has the potential to establish complex relationships and may also possess the ability to interpolate pharmacokinetic parameters and profiles given the formulation specification. Further refinement is possible with the application of the AFM-IVIVC to other data types and larger scale experiments; however, the flexibility and efficient performance of the proposed AFM-IVIVC as it has been shown in this study looks promising and enables the use of *in vitro* dissolution for formulation optimization and as a surrogate for *in vivo* bioequivalent formulations.

Acknowledgements

This work was financially supported from Drug Research Academy (DRA), The Danish University of Pharmaceutical Sciences.

Appendix

In this Appendix, the functionality of the AFM structure, depicted in Figure 1, is explained, from a mathematical viewpoint.

Suppose, for simplicity, that AFM has two inputs, $x_j, j=1,2$, one output variable, y , and three membership functions $m_{1j}(x_j), m_{2j}(x_j), m_{3j}(x_j)$, for each input variable, respectively (see Figure 1). For constant consequence, the rule-base could be formed as

$$R^i: \quad \text{IF } x_1 \text{ is } m_{i1}(x_1) \text{ AND } x_2 \text{ is } m_{i2}(x_2) \text{ THEN } y = f_i, \quad i=1,2,\dots,n, \quad (1)$$

where R^i is the i th fuzzy rule and n is the number of fuzzy rules. In this example, each rule has two variables in the antecedent part (input variables) and one variable in the consequent part (output variable). It should be observed that in the reported example the output variable f_i is a constant (i.e., a *crisp* value) while in the general case it is a fuzzy variable. The circles and squares in Figure 1 represent the units of the network, whereas denotations $w_c, w_g, w_f, 1$ and -1 between the units are the connection weights; bias units with outputs of unity are denoted with the symbol 1. The input-output relationships of the units with symbols of f, Σ , and $\hat{\Pi}$ are defined as

$$I_j^{(n)} = \sum_k w_{jk}^{(n,n-1)} O_k^{(n-1)} \quad (2)$$

$$f: \quad O_j^{(n)} = \frac{1}{1 + \exp(-I_j^{(n)})} \quad (3)$$

$$\Sigma: \quad O_j^{(n)} = I_j^{(n)} \quad (4)$$

$$I_j^{(n)} = \prod_k w_{jk}^{(n,n-1)} O_k^{(n-1)} \quad (5)$$

$$\hat{\Pi}: \quad O_j^{(n)} = \frac{I_j^{(n)}}{\sum_k I_k^{(n)}}, \quad (6)$$

where $I_j^{(n)}$ and $O_j^{(n)}$ are the input and the output of the j th unit in the n th layer, respectively, $w_{jk}^{(n,n-1)}$ is the connection weight between the k th unit and the $(n-1)$ th layer and the j th unit in the n th layer. Units without any symbol just deliver their inputs to succeeding layers. The AFM is divided into the premise parts (layer I through layer V) and the consequence parts (layers V and VI) (see Figure 1). The degrees of truth of the membership functions in the premises are computed in layers I-IV; the connection weights w_c and w_g determine the central position and the gradient of the sigmoid function in the units in layer III, respectively. Appropriate initialization of the weights of the AFM connections allocates the membership functions on the universe of discourse. The truth values of the fuzzy rules are obtained by the product of grades of the membership functions in the units in layer V as

$$\text{Inputs: } \mu_i = \prod_j m_{ij}(x_j) \quad (7)$$

$$\text{Outputs: } \hat{\mu}_i = \frac{\mu_i}{\sum_k \mu_k}, \quad (8)$$

where μ_i is the truth value of the i th fuzzy rule and $\hat{\mu}_i$ is the normalized value of μ_i . In the consequence part (layers V-VI) the connection weights w_{f_i} represent f_i in (1). The inferred defuzzified output of the AFM is given as the output of the unit in layer VI, which is the sum of the products between $\hat{\mu}_i$ and w_{f_i} , i.e.,

$$\tilde{y} = \sum_{i=1}^n \hat{\mu}_i w_{f_i}. \quad (9)$$

References

1. S.A. Charman, W.N. Charman, M.C. Rogge, T.D. Wilson, F.J. Dutko and C.W. Pouton, Self-emulsifying drug delivery systems: formulation and biopharmaceutic evaluation of an investigational lipophilic compound, *Pharm. Res.* 9 (1992) 87-93
2. C.W. Pouton, Lipid formulations for oral administration of drugs: non-emulsifying, self-emulsifying and 'self-microemulsifying' drug delivery systems, *Eur. J. Pharm. Sci.* 11 (2000) S93-S98
3. W.N. Charman, Lipids, lipophilic drugs, and oral drug delivery - some emerging concepts, *J. Pharm. Sci.* 89 (2000) 967-978
4. J.P. Reymond, H. Sucker, In vitro model for cyclosporin intestinal absorption in lipid vehicles, *Pharm. Res.* 5 (1988) 673-676
5. N.H. Zangenberg, A. Mullertz, H.G. Kristensen, L. Hovgaard, A dynamic in vitro lipolysis model. I. Controlling the rate of lipolysis by continuous addition of calcium, *Eur. J. Pharm. Sci.* 14 (2001) 115-122
6. N.H. Zangenberg, A. Mullertz, H.G. Kristensen, L. Hovgaard, A dynamic in vitro lipolysis model. II: Evaluation of the model, *Eur. J. Pharm. Sci.* 14 (2001) 237-244
7. C.J. Porter, A.M. Kaukonen, A. Taillardat-Bertschinger, B.J. Boyd, J.M. O'Connor, G.A. Edwards, W.N. Charman, Use of in vitro lipid digestion data to explain the in vivo performance of triglyceride-based oral lipid formulations of poorly water-soluble drugs: Studies with halofantrine, *J. Pharm. Sci.* 93 (2004) 1110-1121
8. C.J. Porter, A.M. Kaukonen, B.J. Boyd, G.A. Edwards, W.N. Charman, Susceptibility to lipase-mediated digestion reduces the oral bioavailability of danazol after administration as a medium-chain lipid-based microemulsion formulation, *Pharm. Res.* 21 (2004) 1405-1412
9. J.O. Christensen, K. Schultz, B. Mollgaard, H.G. Kristensen, A. Mullertz, Solubilisation of poorly water-soluble drugs during in vitro lipolysis of medium- and long-chain triacylglycerols, *Eur. J. Pharm. Sci.* 23 (2004) 287-296

10. T.J. Sullivan, E. Sakmar, J.G. Wagner, Comparative bioavailability: a new type of *in vitro-in vivo* correlation exemplified by prednisone, *J. Pharmacokinet. Biopharm.* 4 (1976) 173-181.
11. C. Graffner, M. Nicklasson, J.-E. Lindgren, Correlations between *in vitro* dissolution rate and bioavailability of alaproclate tablets, *J. Pharmacokinet. Biopharm.* 12 (1984) 367-380.
12. C. Caramella, F. Ferrari, M.C. Bonferoni, M.E. Sangalli, M. De Bernardi Di Valserra, F. Feletti, M.R. Galmozzi, *In vitro/in vivo* correlation of prolonged release dosage forms containing diltiazem HCl. *Biopharm. Drug Dispos.* 14 (1993) 143-160.
13. V.H. Sunesen, B.L. Pedersen, H.G. Kristensen, A. Mullertz. *In vivo in vitro* correlations for a poorly soluble drug, danazol, using the flow-through dissolution method with biorelevant dissolution media. *Eur. J. Pharm. Sci.* 24 (2005) : 305-13
14. Guidance for Industry, Extended Release Oral Dosage Forms. 1997. Development, evaluation, and application of *in vitro/in vivo* correlations. Ed. Rockville, MD: Food and Drug Administration.
15. Note for Guidance on Quality of Modified-release Products. 1999. A: oral dosage forms, B: transdermal dosage forms section I Quality. Ed. London, UK: Committee for Proprietary Medicinal Products.
16. Guidance for Industry, SUPAC-MR. 1997. Modified-release solid oral dosage forms. Rockville, MD: Food and Drug Administration.
17. R.W. Wood, L. Martis, A.W. Gillum, T.J. Roseman, L. Lin, P. Bernardo, *In vitro* dissolution and *in vivo* bioavailability of commercial levothyroxine sodium tablets in the hypothyroid dog model. *J. Pharm. Sci.* 79 (1990) 124-127.
18. W.H. Barr, E.M. Zola, E.L. Candler, S-M. Hwang, A.V. Tendolkar, R. Shamburek, B. Parker, M.D. Hilty, Differential absorption of amoxicillin from the human small and large intestine. *Clin. Pharmacol. Ther.* 56 (1994) 279-285.
19. G. Levy, L.E. Hollister, Inter- and intra subject variations in drug absorption kinetics. *J. Pharm. Sci.* 53 (1964) 1446-1452.
20. S.I. Gallant, *Neural Network Learning and Expert Systems*, The MIT Press, Cambridge, MA, 1993.

21. M.T. Hagan, H.B. Demuth, M. Beale, Neural Network Design, PWS Publishing, Boston, MA, 1996.
22. D.F. Specht, A general regression neural network. IEEE Trans. Neural Networks 2 (1991) 568-576.
23. R. Erb, The back-propagation neural network - A Bayesian classifier. Introduction and applicability to pharmacokinetics. Clin. Pharmacokinet. 29 (1995) 69-79.
24. A.S. Hussain, R.D. Johnson, N.N. Vachharajani, W.A. Ritschel, Feasibility of developing a neural network for prediction of human pharmacokinetic parameters from animal data. Pharm. Res. 10 (1993) 466-469.
25. M.E. Brier, J.M. Zurada, G.R. Aronoff, Neural network predicted peak and trough gentamicin concentrations. Pharm. Res. 12 (1995) 406-412.
26. P.Veng-Pedersen, N.B. Modi, Application of neural networks to pharmacodynamics. J. Pharm. Sci. 82 (1993) 918-926.
27. A.S. Hussain, X.Q. Yu, R.D. Johnson, Application of neural computing in pharmaceutical product development. Pharm. Res. 8 (1991) 1248-1252.
28. J.A. Dowell, A. Hussain, J. Devane, D. Young, Artificial neural networks applied to the in vitro-in vivo correlation of an extended-release formulation: initial trials and experience, J. Pharm. Sci. 88 (1999) 154-160.
29. A. Gelman, F. Bois, J. Jiang, Physiological pharmacokinetic analysis using population modeling and informative prior distributions, J. Am. Stat. Assoc. 91 (1996) 1400-1412.
30. H. Kortejärvi, J. Malkki, M. Marvola, A. Urtti, M. Yliperttula, P. Pajunen, Level A in vitro-in vivo correlation (IVIVC) model with Bayesian approach to formulation series, J. Pharm. Sci. 95 (2006) 1595-1605.
31. J.-S.R. Jang, ANFIS: Adaptive-network-based fuzzy inference systems, IEEE Trans. on Sys. Man and Cyber. 23 (1993) 665-685.
32. SGS-Thomson Microelectronics, www.st.com
33. L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338-353.
34. A.K. Pannier, R.M. Brand, D.D. Jones, Fuzzy modeling of skin permeability coefficients, Pharm. Res. 20 (2003) 143-148.

35. T.J. Ross, Fuzzy logic with engineering applications, McGraw-Hill, Inc., New York, 1995.
36. T. Bernd, M. Kleutges, A. Kroll. Nonlinear black box modelling-fuzzy networks versus neural networks, *Neural Comput. & Applic.* 8 (1999) 151-162.
37. A. Kroll, Identification of functional fuzzy models using multidimensional reference fuzzy sets, *Fuzzy Sets and Sys.* 80 (1996) 149-158.
38. F.S. Nielsen, E. Gibault, H. Ljusberg-Wahren, L. Arleth, J.S. Pedersen, A. Mullertz, Formulation design and characterization of prototype self-microemulsifying formulations of lipophilic compounds, *J. Pharm. Sci.* 96 (2007) 876-892
39. B.L. Pedersen, H. Brondsted, H. Lennernas, F.N. Christensen, A. Müllertz, H.G. Kristensen HG, Dissolution of hydrocortisone in human and simulated intestinal fluids, *Pharm. Res.* 17 (2000) 183-189
40. Y. Gargouri, H. Moreau, R. Verger, Gastric lipases: biochemical and physiological studies, *Biochim. Biophys. Acta* 1006 (1989) 255-271
41. J.B. Dressman, R.R. Berardi, C.L. Dermentzoglou, T.L. Russell, S.P. Schmaltz, J.L. Barnett, K.M. Jarvenpaa, Upper gastrointestinal (GI) pH in young, healthy men and women, *Pharm. Res.* 7 (1990) 756-761
42. The United States Pharmacopoeia / The National Formulary, (USP 26/NF21), United States Pharmacopeia Convection, Inc., Rockville USP 26, 2003
43. K.J. MacGregor, J.K. Embleton, J.E. Lacy, A.E. Perry, L.J. Solomon, H. Seager, C.W. Pouton, Influence of lipolysis on drug absorption from the gastro-intestinal tract, *Adv. Drug Deliv. Rev.* 25 (1997) 33-46
44. F.S. Nielsen, K.B. Petersen, A. Mullertz, Bioavailability of Probuocol from Lipid and Surfactant Based Formulations in Minipigs: Influence of Particle Size and Dietary State, (Submitted for publication)
45. U.M. Fayyad, G. Piatetsky-Sapiro, P. Smyth, From data mining to knowledge discovery: an overview. In U. M. Fayyad, G. Piatetsky-Sapiro, P. Smyth, and R. Uthurusamy (eds.), *Advances in Knowledge Discovery and Data Mining*, AAAI Press/MIT Press, Menlo Park, CA, (1996) 37-54.

46. J.-S.R. Jang, Neuro-fuzzy modeling: architecture, analyses and applications, Ph.D. Thesis, University of California, Berkeley, CA, 1992
47. L.H. Tsoukalas. Fuzzy and neural approaches in engineering, Wiley & Sons, Inc., New York, 1997.
48. T. Yamaguchi, N. Imasaki, K. Haruchi, Fuzzy rule realization on associative memory system. Proc IJCNN '90 2 (1990) 720–723.
49. D.E. Rumelhart, J.L. McClelland, PDP Research Group, Parallel Distributed Processing, MIT Press, Cambridge, MA, 1986.
50. S. Horikawa, T. Furuhashi, Y. Uchikawa, On fuzzy modeling using fuzzy neural networks with the back-propagation algorithm. IEEE Trans Neural Netw 3(5) (1992) 801-806.
51. J.S.R. Jang, C.T. Sun, E. Mizutani, Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence, Prentice Hall Eaglewood Cliffs, NJ, 1997

TABLE 2: Performance of the AFM-IVIVC scheme for each realization based on the corresponding association (see Figure 2) *

Association #	Training Set			Validation Set			R_V^2 / R_T^2
	$R_T^2(p)$	MPE	MAE	$R_V^2(p)$	MPE	MAE	
1	0.998 ($\ll 0.001$)	-0.001	0.201	0.994 ($\ll 0.001$)	0.020	0.251	0.996
2	0.973 ($\ll 0.001$)	-0.005	0.224	0.941 ($\ll 0.001$)	0.007	0.305	0.967
3	NS	0.021	0.042	NS	0.047	0.056	NS
4	0.945 ($\ll 0.001$)	0.010	0.240	0.914 ($\ll 0.001$)	0.177	0.329	0.967

*NS: Not statistically significant estimation due to very limited number of valid cases; R_T^2 : correlation coefficient (see (10)) for the training set; R_V^2 : correlation coefficient (see (10)) for the validation set; MPE : mean prediction error (see (11)); MAE : mean absolute error (see (12)); p : probability of false hypothesis.

CAPTIONS TO FIGURES

FIGURE 1: Schematic representation of the configuration of the AFM for two inputs and one output. The premise and the consequent parts refer to input variables along with their associations and the formation of the output variable, respectively. For explanations of the symbols and unit-types the reader should refer to the Appendix.

FIGURE 2: Diagrams of the input-output associations used in pattern files ASSOCIATION-1 through ASSOCIATION-4. *PK*: Pharmacokinetic observations (in vivo); *DISS*: % dissolved (in vitro); *i*: pig number; *j*: tablet number; t_{PK} : pharmacokinetic time point; t_{DISS} : dissolution time point.

FIGURE 3: The consumption of NaOH as a function of time for the oil formulation (■), the SNEDDS formulation (●), the SMEDDS formulation (▼), (mean ± S.E., n=3)

FIGURE 4: The % of drug released into the aqueous phase as a function of time using the in vitro dynamic lipolysis model for the oil formulation (■), the SNEDDS formulation (●), the SMEDDS formulation (▼), (mean ± S.E., n=3)

FIGURE 5: Plasma concentrations of probucol after administration of different lipid formulations in fed state for the oil formulation (■), the SNEDDS formulation (●), the SMEDDS formulation (▼), (mean ± S.E., n=6). Adopted from Nielsen et al. [44]

FIGURE 6: Prediction performance of the AFM-IVIVC for the case of pattern files of ASSOCIATION-1. (a) Actual pharmacokinetic observations (white squares) from the training data set compared with AFM-IVIVC pharmacokinetic predictions (black triangles) using *in vitro* inputs from the training data set. (b) Actual pharmacokinetic observations (squares) from the validation data set compared with AFM-IVIVC pharmacokinetic predictions (black triangles) using *in vitro* inputs from the validation data set. Downwards, upwards and rightwards triangles correspond to the output from subnets 1 through 3 of the AFM-IVIVC ASSOCIATION-1, respectively (see Figure 3).

FIGURE 7: Prediction performance of the AFM-IVIVC for the case of pattern files of ASSOCIATION-2. (a) Actual pharmacokinetic observations (white squares) from the training data set compared with AFM-IVIVC pharmacokinetic predictions (black triangles) using *in vitro* inputs from the training data set. (b) Actual pharmacokinetic observations (squares) from the validation data set compared with AFM-IVIVC pharmacokinetic predictions (black triangles) using *in vitro* inputs from the validation data set.

FIGURE 8: Prediction performance of the AFM-IVIVC for the case of pattern files of ASSOCIATION-3. (a) Actual pharmacokinetic observations (white squares) from the training data set compared with AFM-IVIVC pharmacokinetic predictions (black triangles) using *in vitro* inputs from the training data set. (b) Actual pharmacokinetic observations (squares) from the validation data set compared with AFM-IVIVC pharmacokinetic predictions (black triangles) using *in vitro* inputs from the validation data set.

FIGURE 9: Prediction performance of the AFM-IVIVC for the case of pattern files of ASSOCIATION-4. (a) Actual pharmacokinetic observations (white squares) from the training data set compared with AFM-IVIVC pharmacokinetic predictions (black triangles) using *in vitro* inputs from the training data set. (b) Actual pharmacokinetic observations (squares) from the validation data set compared with AFM-IVIVC pharmacokinetic predictions (black triangles) using *in vitro* inputs from the validation data set.

FIGURE 10: The estimated membership functions after their automated tuning by the AFM for the input parameters of the AFM-IVIVC ASSOCIATION-4, i.e., (a) t_{PK} , (b) $DISS_j(t_{DISS1})$, (c) $DISS_j(t_{DISS2})$, (d) $DISS_j(t_{DISS3})$, (e) $DISS_j(t_{DISS4})$ and (f) $DISS_j(t_{DISS5})$, corresponding to the three linguistic variables ‘low’, ‘medium’ and ‘high’.

FIGURE 1

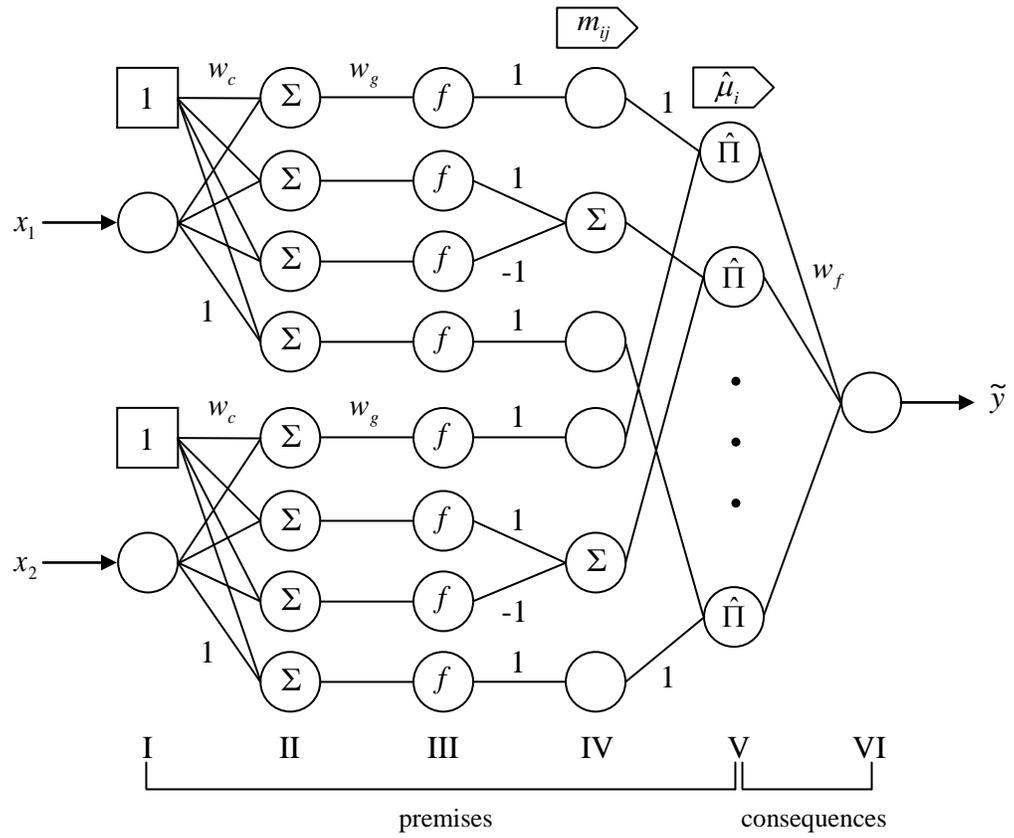


FIGURE 2

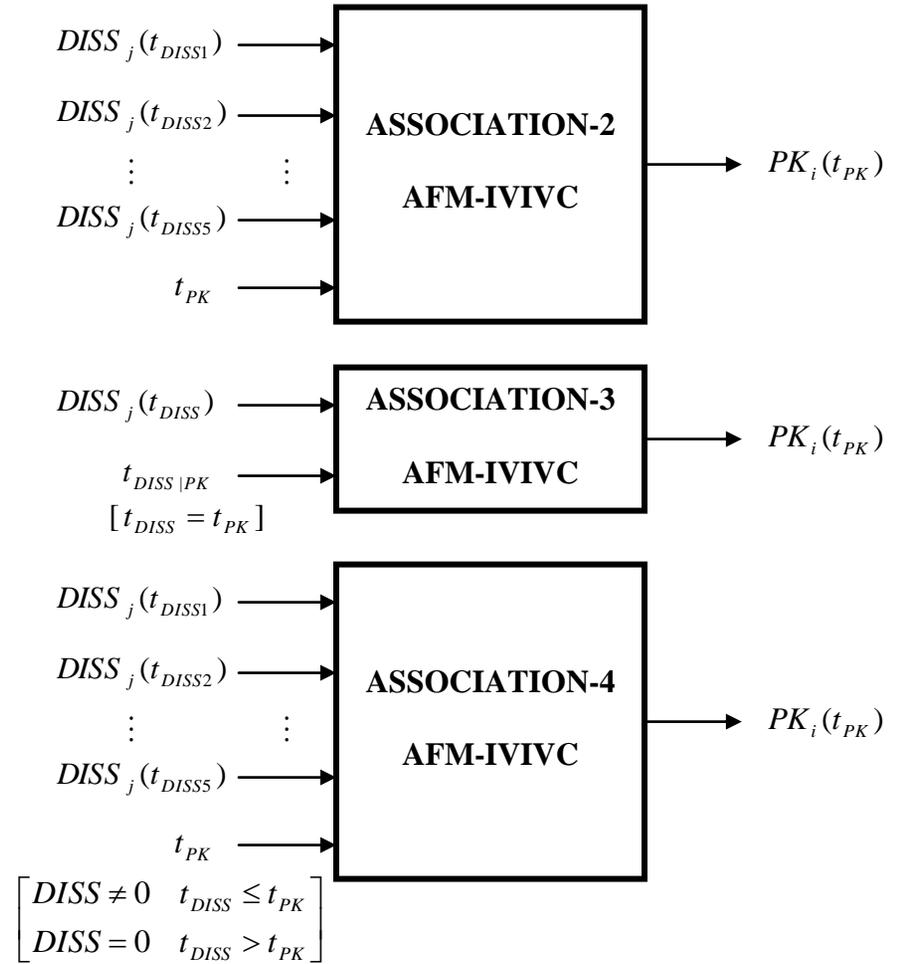
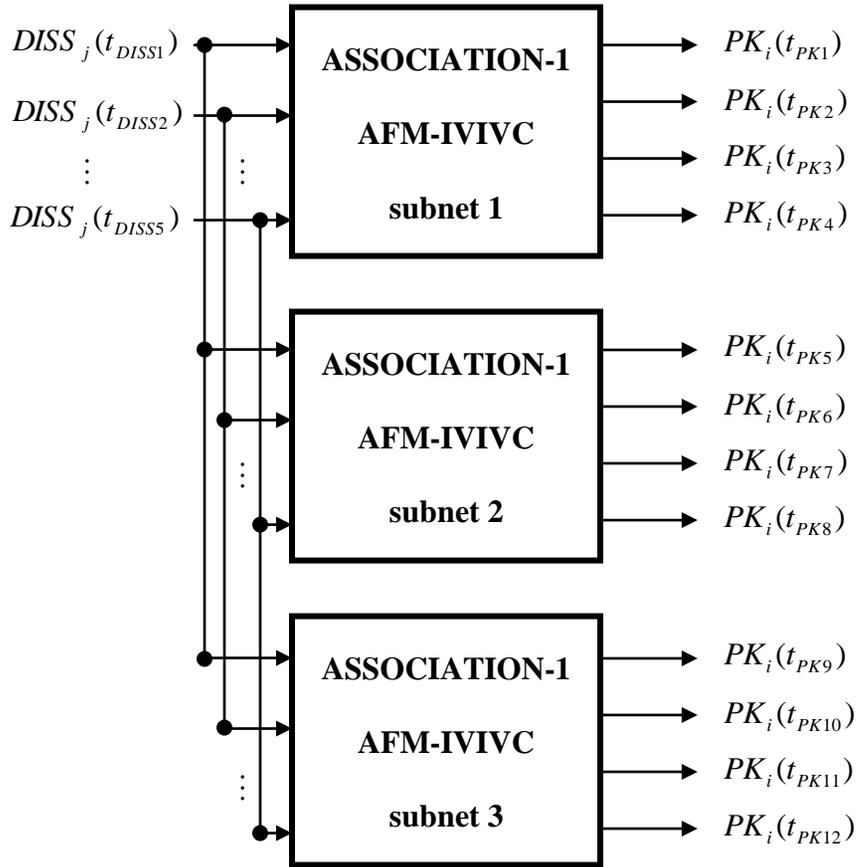


FIGURE 3

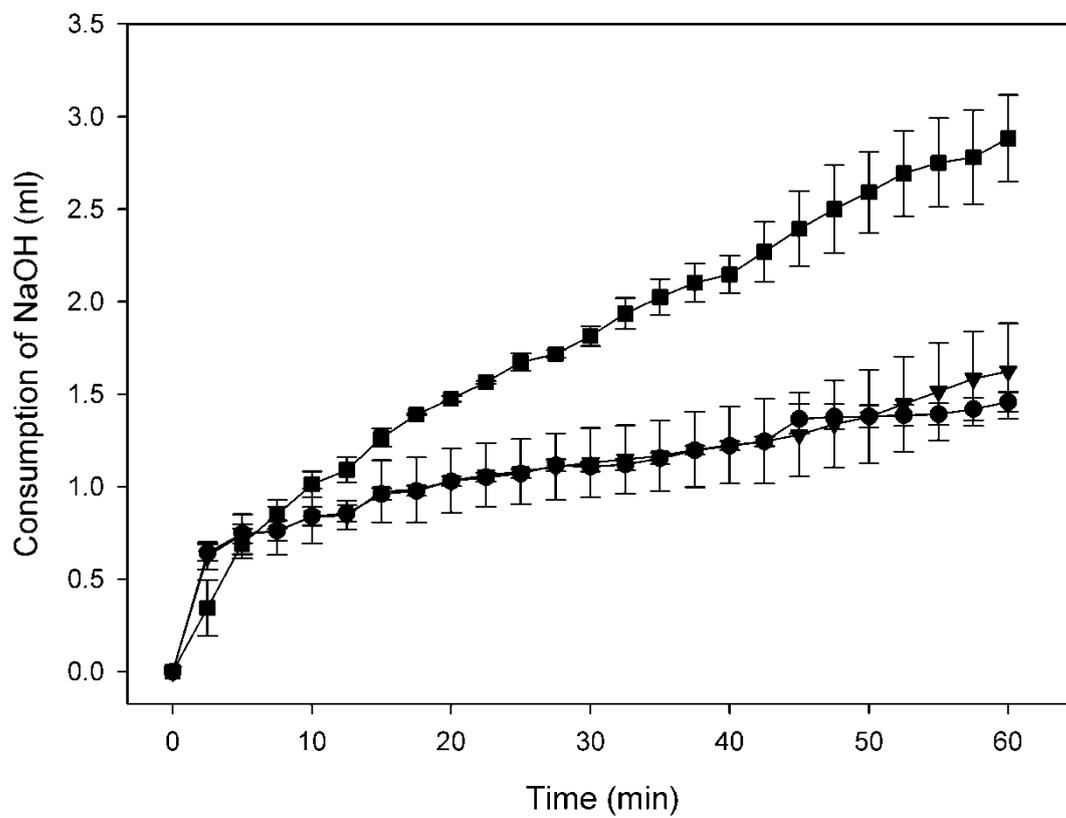


FIGURE 4

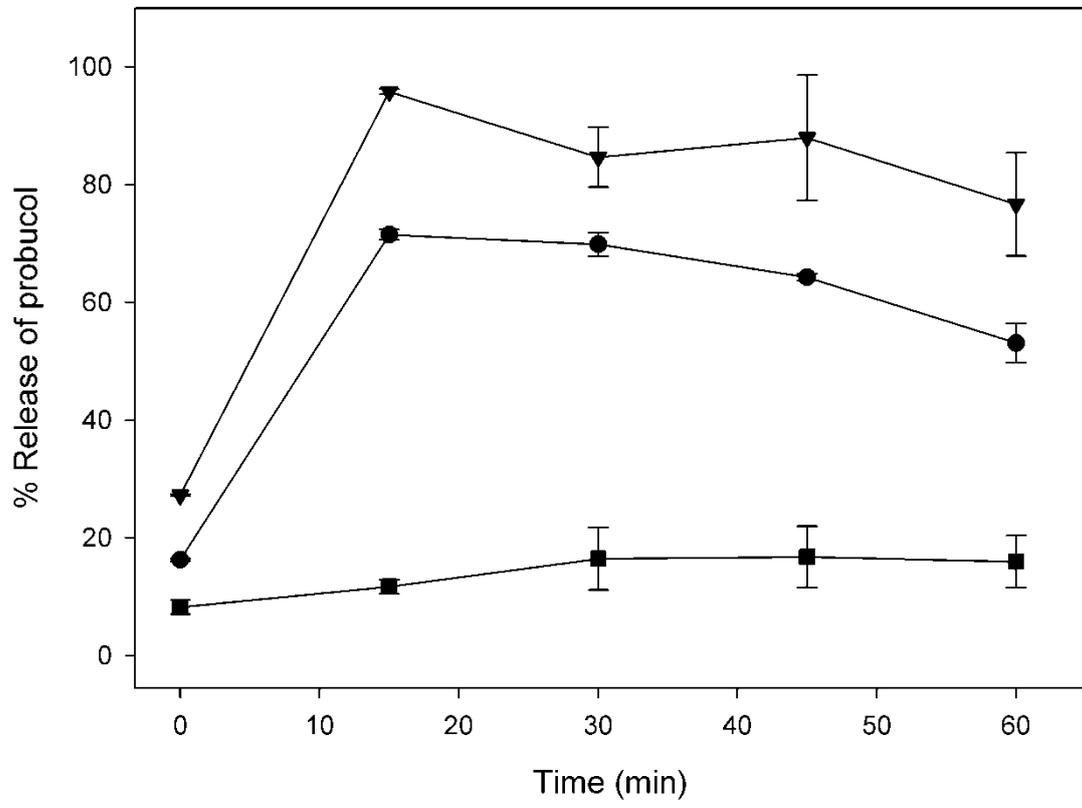


FIGURE 5

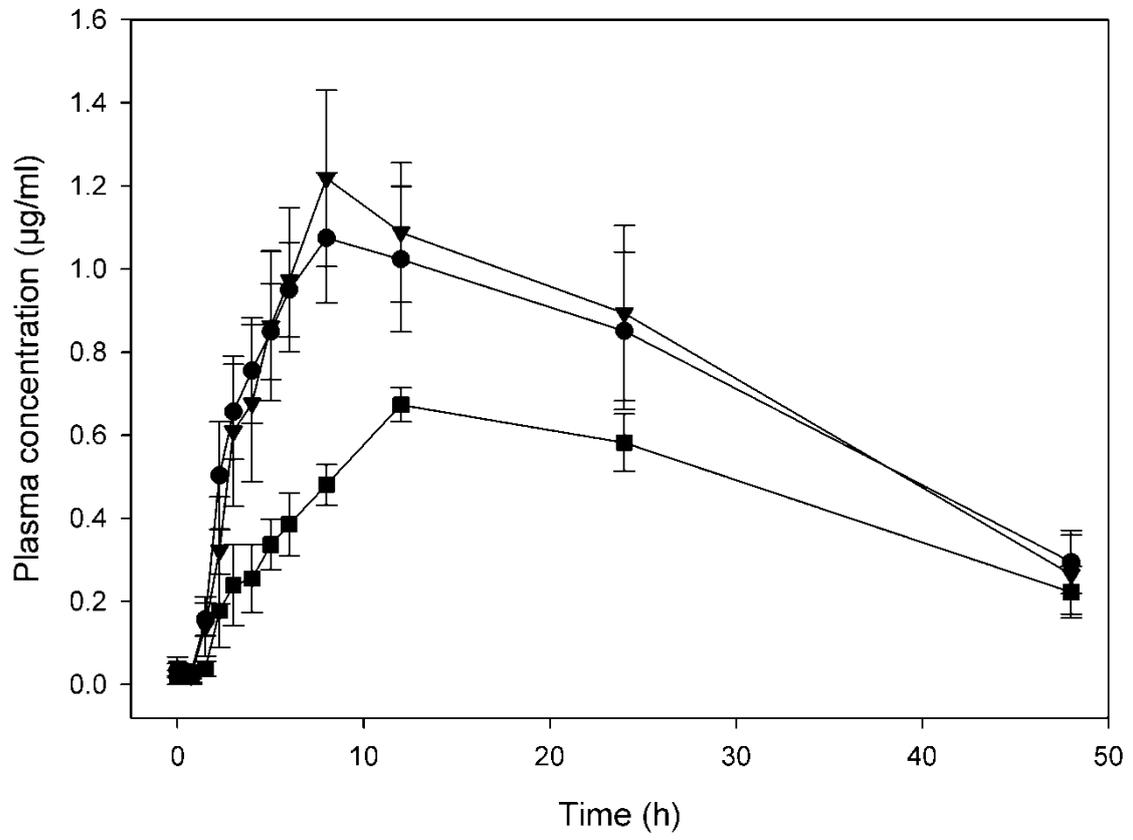
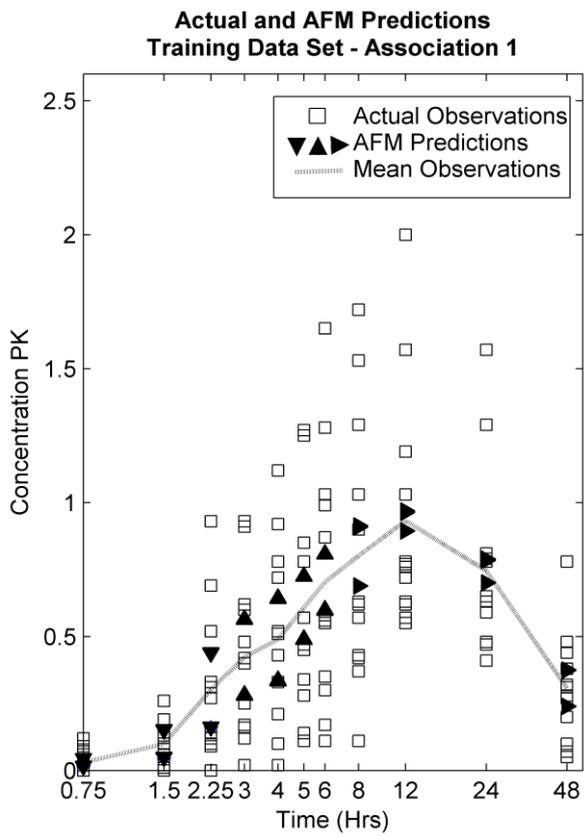
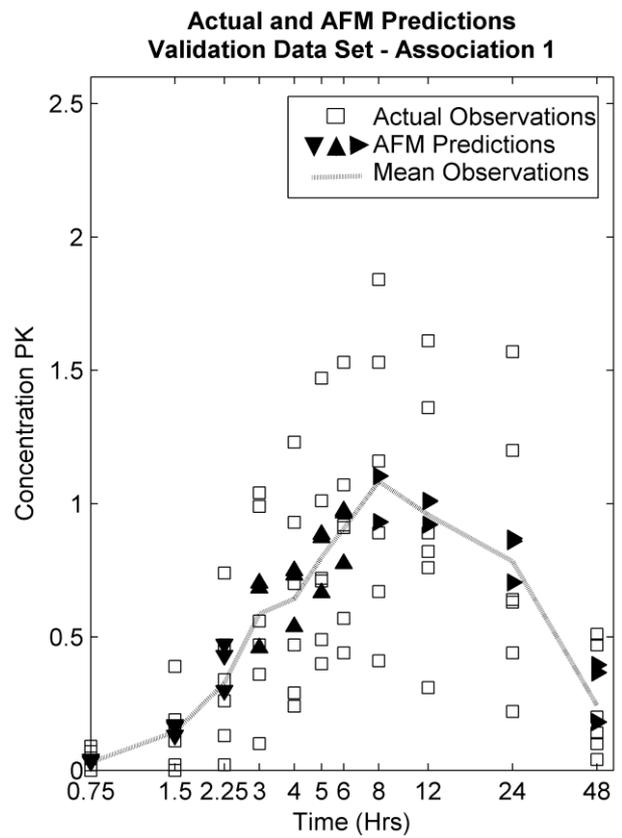


FIGURE 6



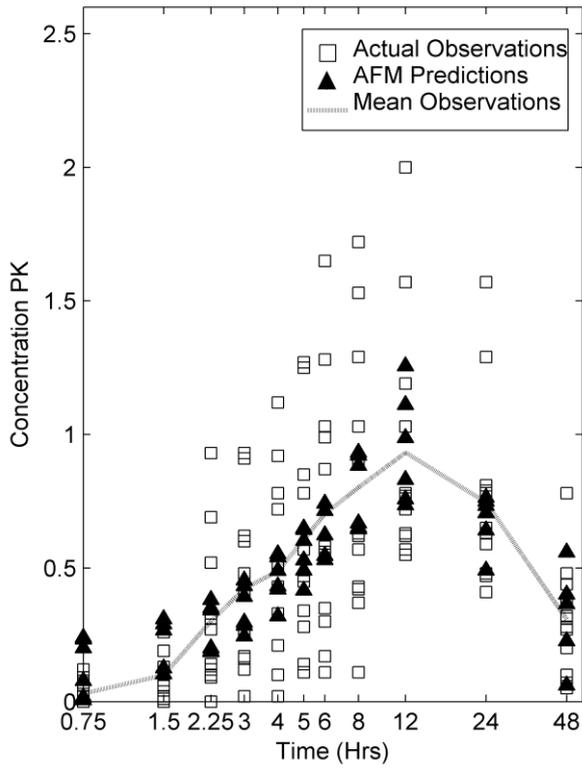
(a)



(b)

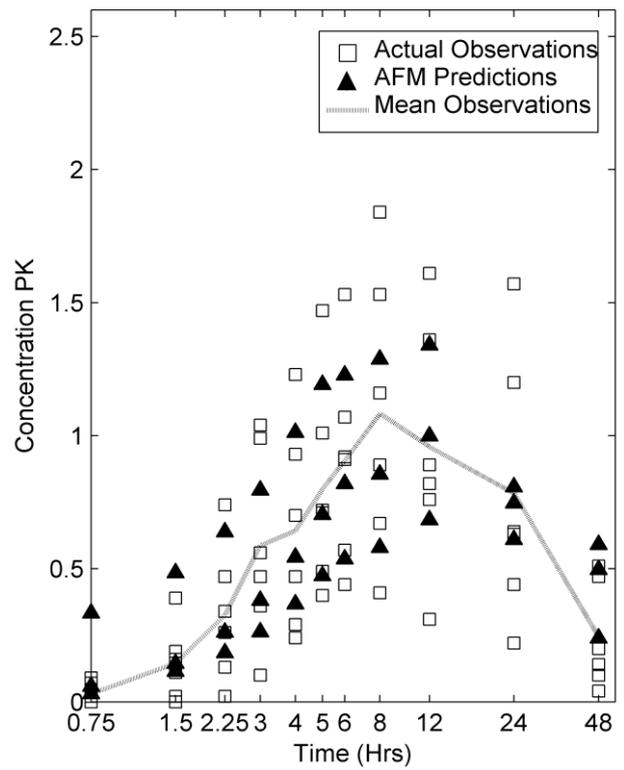
FIGURE 7

**Actual and AFM Predictions
Training Data Set - Association 2**



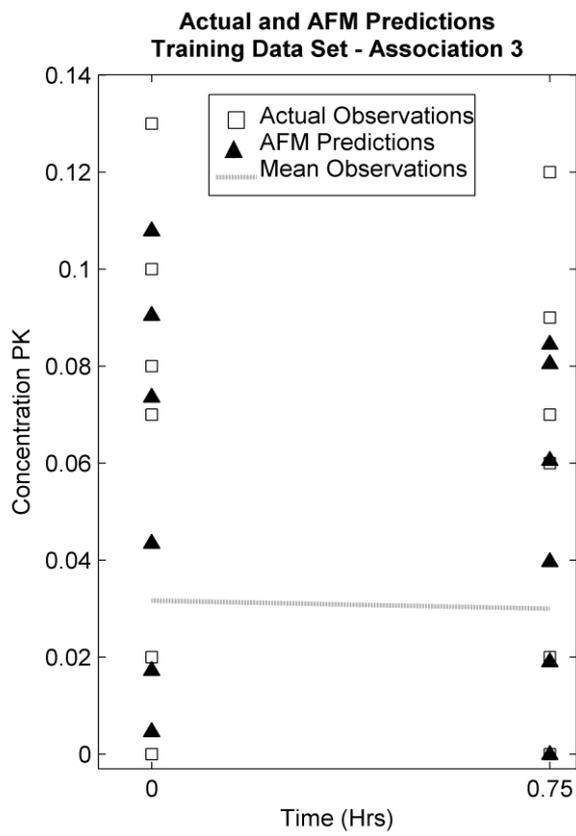
(a)

**Actual and AFM Predictions
Validation Data Set - Association 2**

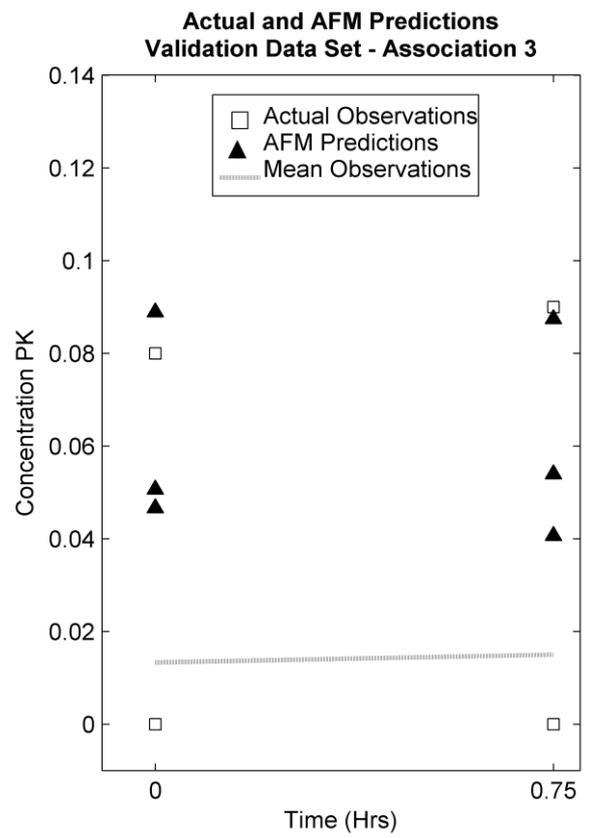


(b)

FIGURE 8

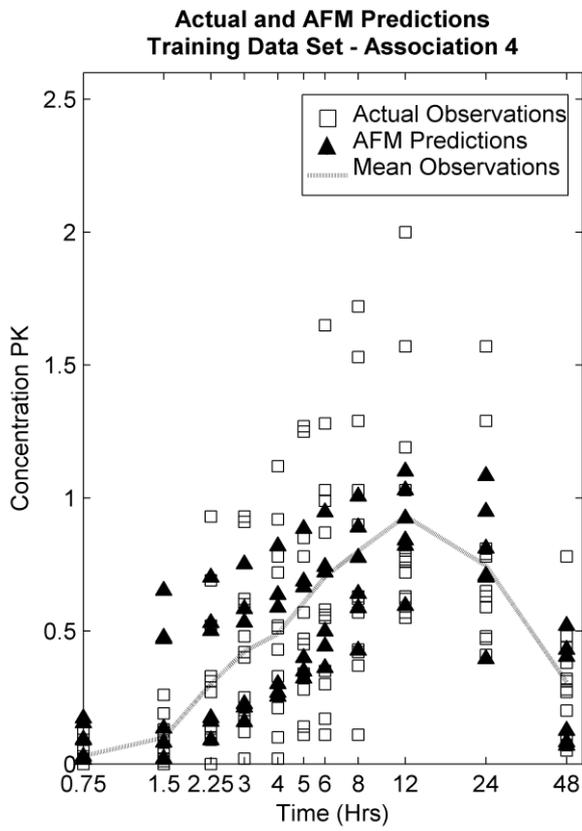


(a)

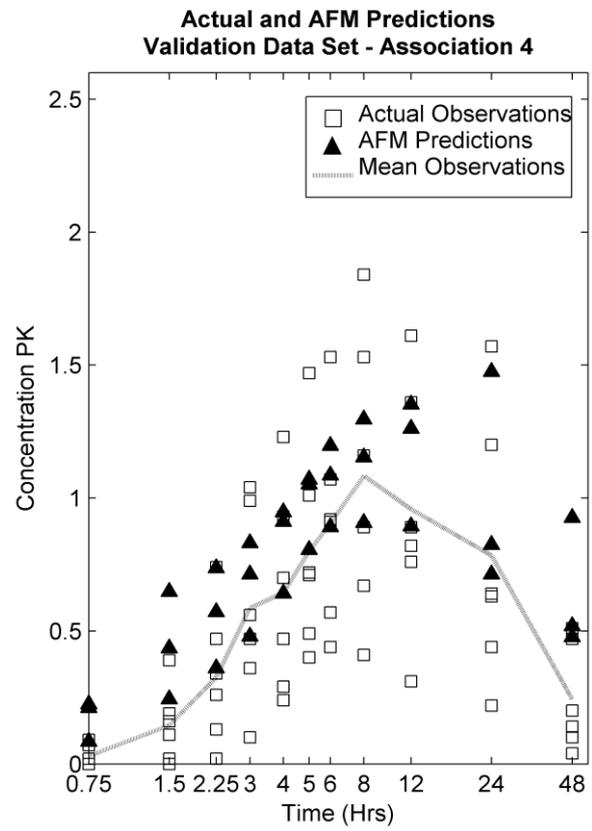


(b)

FIGURE 9

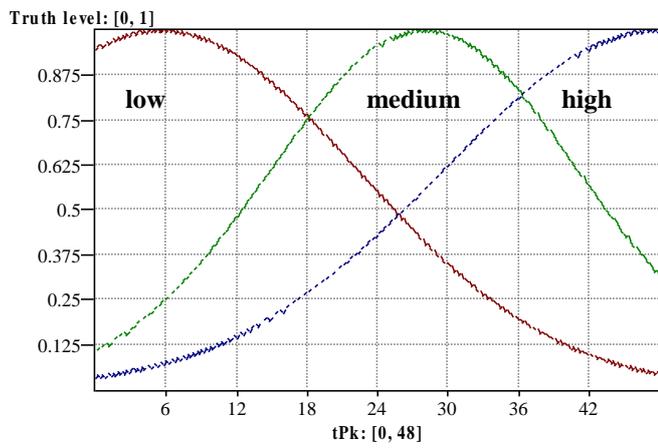


(a)

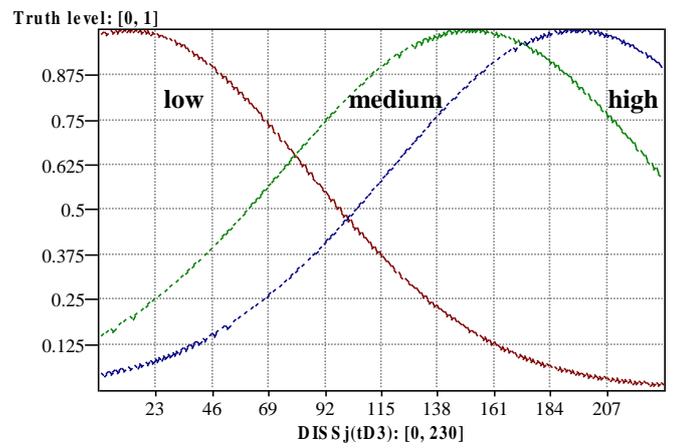


(b)

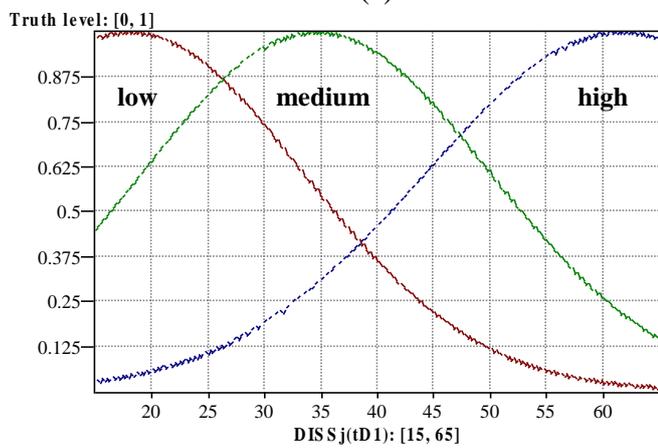
FIGURE 10



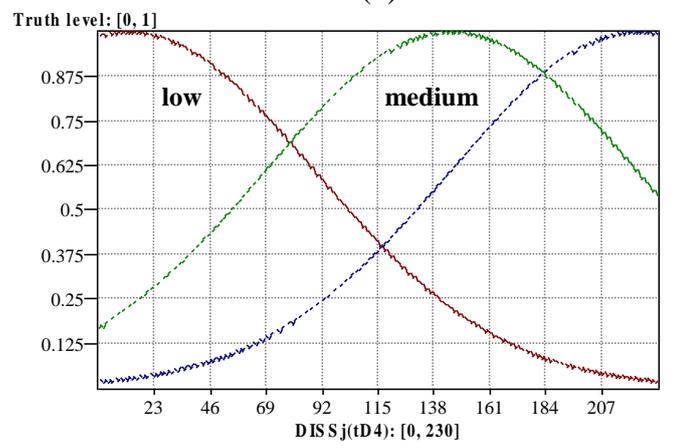
(a)



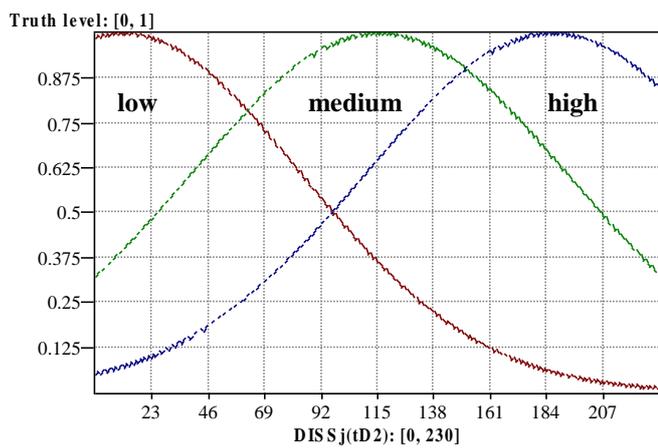
(d)



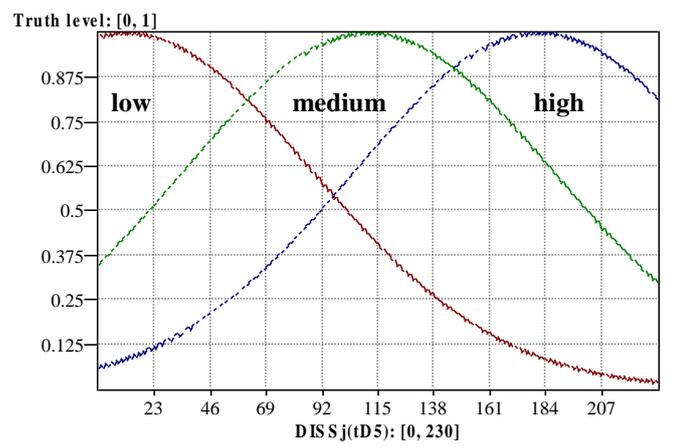
(b)



(e)



(c)



(f)