Genetic fuzzy system predicting contractile reactivity patterns of small arteries

Tang, J; Sheykhzade, Majid; Clausen, B F; Boonen, H C M

www.nature.com/psp

ORIGINAL ARTICLE

Genetic Fuzzy System Predicting Contractile Reactivity Patterns of Small Arteries

J Tang[1], M Sheykhzade[1], BF Clausen[1] and HCM Boonen[1]

Monitoring of physiological surrogate end points in drug development generates dynamic time-domain data reflecting the state of the biological system. Conventional data analysis often reduces the information in these data by extracting specific data points, thereby discarding potentially useful information. We developed a genetic fuzzy system (GFS) algorithm that is capable of learning all information in time-domain physiological data. Data on isometric force development of isolated small arteries were used as a framework for developing and optimizing a GFS. GFS performance was improved by several strategies. Results show that optimized fuzzy systems (OFSs) predict contractile reactivity of arteries accurately. In addition, OFSs identified significant differences that were undetectable using conventional analysis in the responses of arteries between groups. We concluded that OFSs may be used in clustering or classification tasks as aids in the objective identification or prediction of dynamic physiological behavior.

CPT Pharmacometrics Syst. Pharmacol. (2014) **3**, e108; doi:10.1038/psp.2014.3; pubished online 2 April 2014

INTRODUCTION

Pharmacodynamic effects of drugs are often measured using time-domain series of physiological data, e.g., vascular reactivity, blood pressure, electrocardiograms, or electroencephalograms. It is not uncommon that large amounts of data are reduced to but a few data points expressing, e.g., a drug effect at a certain dose. The dynamic component and the effect of time are hereby lost. However, high-resolution physiological data include potentially considerably more information about the biological system's behavior, which can be relevant for assessing drug efficacy or safety. The purpose of this work was to develop a hypothesis-free mathematical algorithm, capable of predicting the dynamic behavior of time-dependent contractile reactivity of isolated blood vessels. An advantage of a hypothesis-free approach is that no assumptions are made with regard to the underlying mechanism of the dynamic process. This is of importance in the development of new medicines for which understanding of mechanisms of action, or in particular, potential off-target effects of the new drug, may be incomplete. Vascular reactivity is central in normal physiology, as also in the etiology of pathologies of the cardiovascular system. It serves as a surrogate end point for pharmacological treatment aimed toward normalization of altered vascular function. Classic pharmacological characterization of *in vitro* isolated blood vessels or mathematical modeling of blood vessels has clarified molecular mechanisms[1–3] and biophysical characteristics controlling vascular smooth muscle contractile behavior.[4] However, these methods are often based on assumptions of steady state and simplifications of the underlying mechanism. Principally, the dynamics of contractile behavior, an emergent property of blood vessels, is too complex to explain on the basis of a few parameters. Adaptation, feedback mechanisms, and structural remodeling, among others, may alter initial assumptions and can therefore not be accounted

for by classic methods of investigation.[5–7] For example, time-domain data of force development in isolated blood vessels is reduced to a few extracted data points and analyzed in order to understand artery dynamics. Hence, most available data are discarded as nonuseful information. Although waveform analysis–fast Fourier transform has revealed the chaotic dynamic behavior of isolated blood vessels,[8,9] the patterns of tension generated over time of specific isolated blood vessels are often characteristic for the type and origin of arteries and may be indicative of underlying system changes due to pathology.[10,11]

Therefore, artificial intelligence or machine learning methods may help in describing dynamic patterns of system behavior by learning all information in the data. In mathematical modeling, ordinary differential equations could describe a mechanistic model of systems. However, these approaches are incapable of describing all parameters of the system or uncertainties due to changes in either the system or its parameters.[4] To address this problem, fuzzy logic may be an alternative, providing rule-based modeling in which the phenotypic expression of the system is under control (i.e., tension), rather than its individual components. A fuzzy logic system is, by nature, capable of capturing the behavior and unexpected events of nonlinear complex dynamic systems within certain limits. It has been used previously in pharmacology for different purposes.[12–15]

This study describes the development of a genetic fuzzy system (GFS) as a method to study and predict the dynamic behavior of contractile responses of isolated arteries. Briefly, GFS is an artificial intelligence–based computer algorithm that can recognize patterns. GFS has two constituents, a genetic algorithm (GA), which is a search algorithm that is used to find the most optimal solution to a fuzzy logic system, which is the second component of the algorithm that enables prediction of dynamic behavior or

[1]Department of Drug Design and Pharmacology, Systems Pharmacology Section, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. Correspondence: HCM Boonen (hcmb@sund.ku.dk)

patterns. The principle behind GFS is to explore information from experimental data, in cases where there is only partial or no mechanistic knowledge of the complex biological system.[16,17]

Thus, the purpose of this study is to optimize a fuzzy control system using a GA and to use the optimized fuzzy system (OFS) to predict contractile reactivity data of arteries as accurately as possible. OFSs are thus shown to be superior to conventional analysis in identifying differences in vascular behavior.

RESULTS

This study comprised two parts: development of the GFS algorithm, and application of an OFS.

GFS development
The experimental data (data set 1) used for development of the GFS algorithm consisted of recordings of isometric force development in isolated resistance-sized branches of rat femoral arteries (**Figure 1**). In **Figure 1**, vertical lines represent additions of vasoactive compounds at various time points and at different concentrations. The nature of these compounds is, in principle, not relevant for the purpose of this study, but details of the experimental protocols are provided in the **Supplementary Data**.

GFS design and implementation. The GFS design had two objectives that would collectively improve GFS performance (i) to obtain an OFS yielding the best prediction for available data and (ii) to improve computational efficiency. A detailed description of the GFS principles is provided in the **Supplementary Data**. Principally, a GA is used to optimize a fuzzy system that can predict the behavior of the dynamic response, based on a fitness value. The fitness value is defined as the difference between the output and the desired output, i.e., the cumulative sum of the squared errors and can be considered similar to goodness of fit (see **Supplementary Data**).

Variation in the number of membership functions per fuzzy set, from two (*GFS (222)*) to six (*GFS (666)*) for each of two
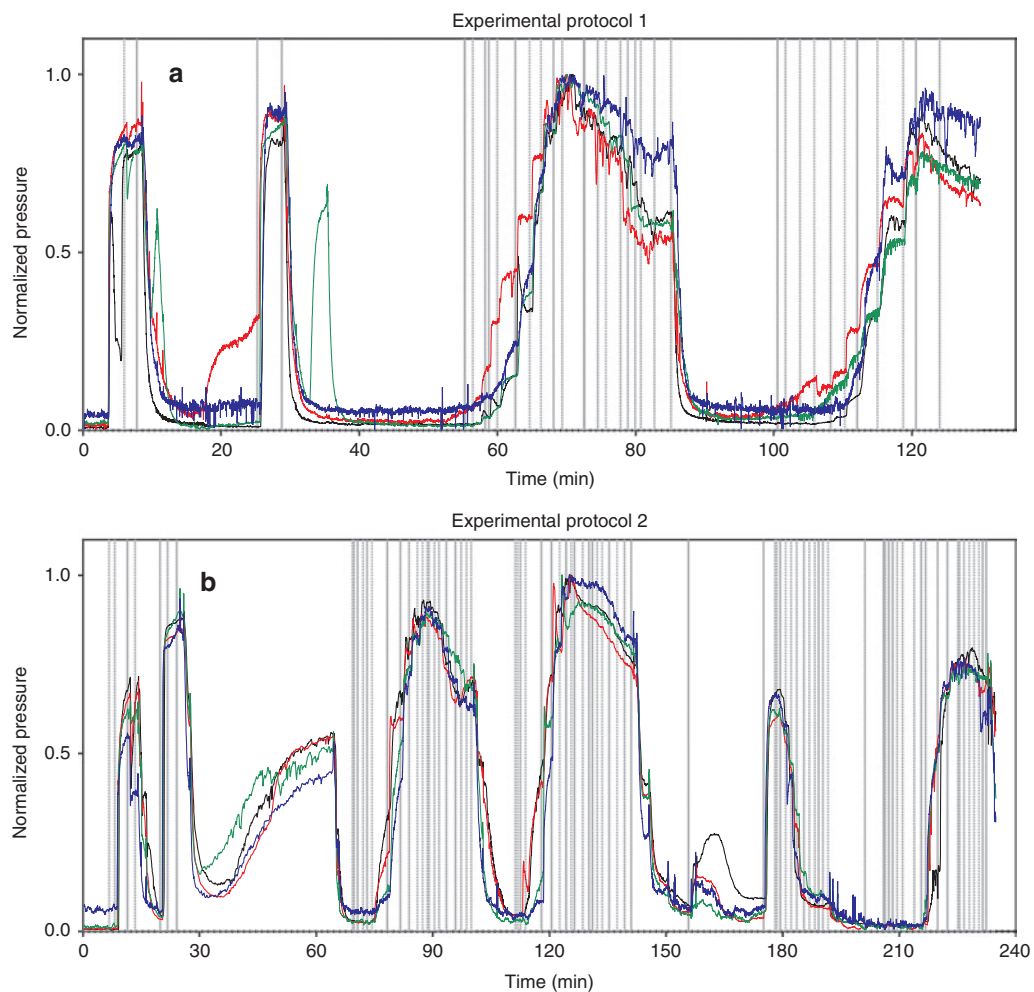

Figure 1 Recordings of isometric tension development in isolated rat femoral arteries over time, expressed as normalized pressure. These data sets were used as training and test sets for genetic fuzzy system (GFS) development. Vertical lines represent additions of vasoactive substances in various concentrations (see **Supplementary Data** for complete experimental protocol). (**a**) Data set 1; (**b**) data set 2. Each panel displays recordings of four different arteries, represented by different colors.

inputs and one output, showed no correlation with GFS performance (**Figure 2**). Therefore, the simplest GFS design with two fuzzy sets per variable (*GFS (222)*) yielded the best computation-to-fitness ratio. Strategies were implemented to improve either the search heuristic for fuzzy systems or
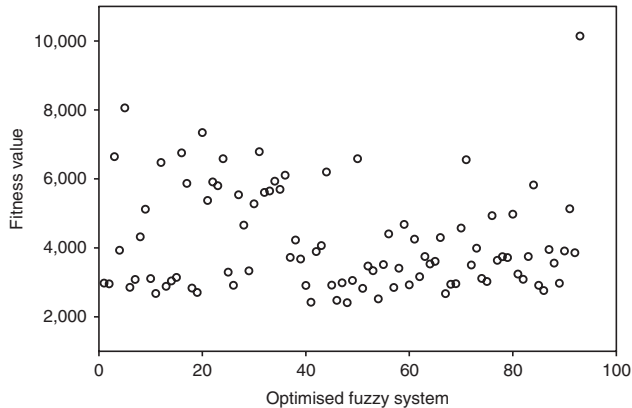
to improve performance. The results of these strategies are described below.

GFS performance. Briefly, four strategies were performed either singly or in combination affecting (i) type and weight of membership functions, (ii) generation of membership function variables, (iii) population strategy, and (iv) fuzzy system range and defuzzification method. The impact of these strategies was assessed using system *GFS(222)*. **Figure 3a** displays evolution of fitness over 1,000 generations using fuzzy sets with six possible membership functions as suggested by Shi *et al.*[18] Repetition yielded comparable fitness values, indicating that searches ended in the same minimum. Implementing strategies 1 and 2 by increasing the number of membership functions and how these were generated resulted in an evolution of fitness shown in **Figure 3b**,**c**. With these two strategies, the system becomes more flexible and adaptive to dynamic changes in data. Nevertheless, the same algorithm design resulted in different fitness values that were either similar (**Figure 3a**) or worse with strategy 1 alone. In addition, a large variation in mean fitness during the optimization process indicated that this design not necessarily converge to a global minimum in just one run. Addition of strategy 3 to the algorithm design resulted in an evolution of fitness as shown in **Figure 3d**, suggesting that a combination of all three strategies gave the largest diversity in the population, further improving the chance to obtain individuals with good fitness.
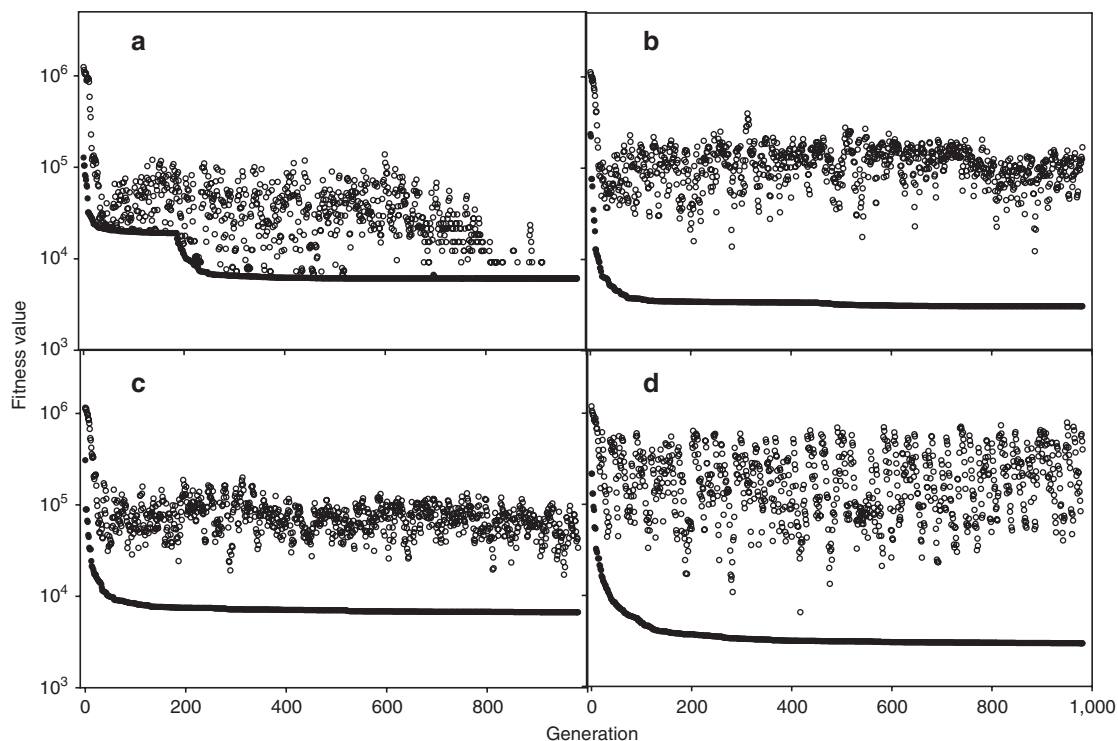

Figure 2 The impact of different numbers of fuzzy sets per input or output on optimized fuzzy system (OFS) fitness values. Displayed are best fitness values for different OFSs from *OFS(222)* containing any combination ranging from two fuzzy sets per input or output up to *OFS(456)* containing four and five fuzzy sets, or six fuzzy sets, for the two inputs and one output, respectively. There was no correlation with number of fuzzy sets and fitness values in OFSs.

Figure 3 Different strategies were investigated in order to improve the genetic algorithm search for the most optimal fuzzy system. This figure depicts the evolution of fitness values over 1,000 generations of *GFS(222)* with these strategies. (**a**) *GFS(222)* evolution of fitness using fuzzy sets with six possible membership functions, as suggested by Shi *et al.*[23] (final best fitness: 6,097.33; final mean fitness: 6,097.46; range: 6,097.30–6,102.20); (**b**,**c**) different runs of *GFS(222)* with strategies 1 plus 2; ((**b**) final best fitness: 3,071.02; final mean fitness: 169,377; range: 3,071.00–1,760,600; (**c**) final best fitness: 6,708.43; final mean fitness: 34,097; range: 6,708.40–1,314,400); (**d**) *GFS(222)* with strategies 1, 2, and 3. The training data are from the nonnormalized data of BV1 (final best fitness: 2,979.78; final mean fitness: 399,836; range: 2,979.80–3,397,800). Solid circles represent the best fitness at each generation, and open circles represent mean fitness per generation.

Although strategies 1–3 improved the performance of the GFS dramatically, the prediction of OFSs appeared worst at the extremes of the data range (**Figure 4a**). Two possible reasons could be responsible for this. First, some membership functions are rarely being used in fuzzy sets of the OFS at either ends of each input or output, e.g., Gaussian, triangle, or trapezoid-shaped functions (**Supplementary Figures S1 and S2**). Additionally, because the default defuzzification function in Matlab is centroid, the prediction at the borders of the output was underestimated. With strategy 4, these drawbacks were overcome and predictions further improved (**Figure 4b**).

OFS test and specificity. Prediction of data using *OFS(222)* of a small population of arteries is illustrated in **Figure 5**. It is evident that, by this method, the averaged prediction of the population of arteries can almost be superimposed on the averaged original data and variation in the prediction is well within that of the original data.

The difference in fitness between an OFS training artery and others using the same OFS reflects specificity. Specificity was investigated by increasing the number of inputs (see **Supplementary Data**) and by comparing fitness of these OFSs for recordings of arteries with the same (protoc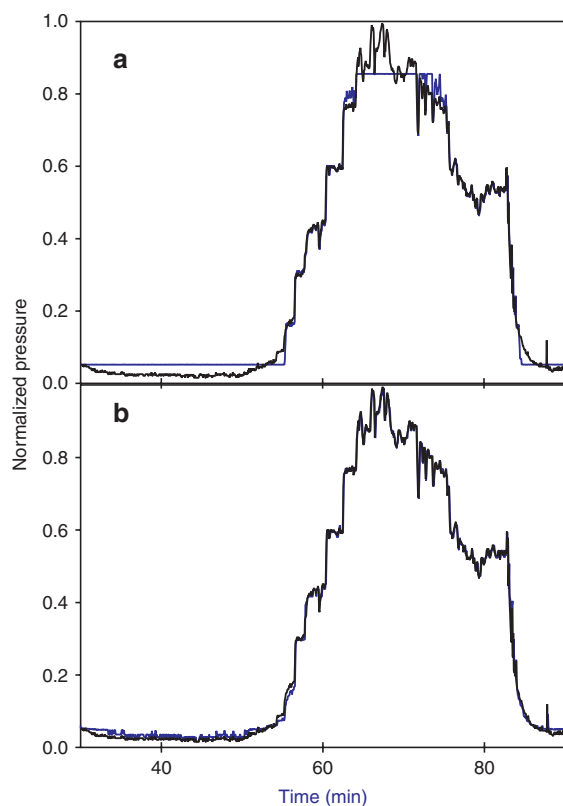ol 1) or different (protocol 2) experimental protocols. Four OFSs with different inputs were optimized using the same training data from blood vessel 3 (BV3) (**Table 1**). Absolute fitness values and the percentage difference between training data fitness and test data fitness are shown. The higher the percentage difference, the more specific the OFS is for predicting a certain data set. In general, the training artery had the best fitness with different OFSs. Additionally, for every increase in inputs, the fitness for arteries in protocol 1 was better than that for recordings in protocol 2. Furthermore, it is evident that increasing the number of inputs improved OFS specificity because the difference in fitness for training artery and test data increased proportionally with more inputs.

Application of GFSs

Comparison of OFSs with conventional analysis of data was performed using another data set consisting of two experimental groups of isolated rat small mesenteric arteries that had either been freshly isolated (Group A) or been incubated for 24 h in tissue culture medium (Group B). The dynamic pattern of norepinephrine cumulative concentration–response relationships in these arteries was visually very different between groups (**Figure 6a**). In the continuous presence of increasing concentrations of norepinephrine, contractions in freshly isolated arteries were relatively stable, whereas in cultured arteries, contractions were transient and returned to baseline state, despite the presence of norepinephrine. Conventional analysis of norepinephrine potency and maximal responses, either as maximum response or as averaged responses per concentration, did not reveal this difference (**Figure 6b**). On the other hand, fitness values that were generated by an OFS, trained on a random artery of either Group A or Group B, revealed statistically significant differences between groups (**Figure 6c**). OFSs trained on either Group A or Group B showed differences in absolute fitness values and variance of fitness, illustrating (i) a larger variation in dynamic behavior of arteries in Group B as compared with that in Group A and (ii) greater specificity of OFSs trained on Group B data. Nevertheless, differences between groups were evident regardless of the training set. Fuzzy rules of these OFSs are provided in the **Supplementary Table S1**.

Figure 4 Representative part of original data from data set 1 visualizing improvement in prediction at the bottom and top ranges of data with the addition of strategy 4. (**a**) Prediction using *OFS(222)* without strategy 4; (**b**) prediction using *OFS(222)* with strategy 4. The training data are from the normalized data of BV3 (black line), whereas the prediction is represented by the blue line.

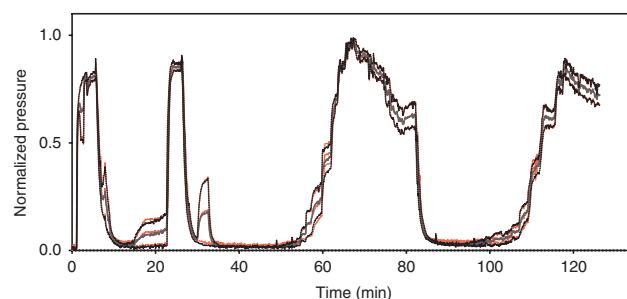**Figure 5** Test of optimized fuzzy system. The black solid line shows the averaged normalized pressure data trace of the population of blood vessels from data set 1, with standard error indicated by the dotted black lines. The average and standard error of the prediction of the population are shown as solid and dotted red lines, respectively. The prediction is not always visible because it is superimposed on the original data.

Table 1 Average fitness values of OFSs with increasing number of inputs

	Protocol 1				Protocol 2			
	BV1	**BV2**	*BV3*	**BV4**	**BV5**	**BV6**	**BV7**	**BV8**
OFS1	0.3911 (1)	0.3961 (3)	*0.3859*	0.3459 (−10)	0.4870 (26)	0.5331 (38)	0.4597 (19)	0.5443 (41)
OFS2	0.4369 (62)	0.2841 (6)	*0.2689*	0.3153 (17)	0.3973 (48)	0.6562 (144)	1.0092 (275)	1.2236 (355)
OFS3	1.1628 (179)	0.4886 (17)	*0.4166*	0.4690 (3)	1.2680 (204)	0.5382 (29)	0.5885 (41)	0.6491 (56)
OFS4	1.8900 (696)	0.3491 (47)	*0.2373*	0.3713 (56)	0.5887 (149)	1.3367 (464)	1.1468 (384)	0.4332 (83)

The table shows the impact of changes in GFS design with respect to the number of inputs on the performance of OFS. OFS1(*OFS(222)*) uses error and change of error as inputs; OFS2 (*OFS(2222)*) adds one more input, average of integration of all previous errors from current time point; OFS3(*OFS(2222)*) adds one more input, average of integration of five previous neighboring errors from current time point; OFS4 (*OFS(222222)*) adds three more inputs, time points before current time point t, such as $t − 4$, $t − 7$, and $t − 10$. Shown are average fitness (fitness value/length of data given in ‰) and percentage difference from the training data fitness (numbers in brackets) for each respective OFS. This percentage difference does not reflect error, but a large difference of fitness values between training and test data indicates a more specific OFS for the training set. Differences in fitness can therefore be indicative of differences in dynamic behavior of the original data.

GFS, genetic fuzzy system; OFS, optimized fuzzy system.

DISCUSSION

This study describes the development of a GFS algorithm, which is an artificially intelligent system, capable of predicting time-domain dynamic physiological data. The resulting OFS predicted the complex behavior of isometric contractions of isolated arteries with an accuracy level beyond what, to our knowledge, has been shown with alternative mathematical modeling methods. We have shown that OFSs could be useful for prediction of the behavior of complex physiological parameters and thereby can objectively assist with clustering or classification tasks. This method could therefore be useful in assessing the effect of drug intervention or perturbation of time-domain physiological parameters.

GFS can be designed in various ways depending on the purpose of the study and whether previous knowledge of the system is available or not.[16,17] In this study, we assumed that there is no precise knowledge about the contractile behavior of an isolated artery, which can form the basis for an expert-designed fuzzy logic system or mechanistic model. Hence, a GA was used to randomly generate parameters of FSs, which were then coded as sequences of genes, in analogy with biological genetic principles.[18] Briefly, individuals in a population of fuzzy systems were tested for their ability to predict a training set of physiological data and individuals survived, based on precision of the prediction (fitness). Several strategies were introduced with the purpose of improving the performance of GFSs. It was shown that this process was essential because data sets were very dynamic, noncyclic, and complex (**Figure 1**). Strategies in which type and weight of membership functions and the generation of membership function variables were optimized (strategies 1 and 2) improved the adaptivity of GFSs to the dynamic data dramatically (**Figure 3a**–**c**). Nevertheless, because the search space is multidimensional and multimodal, it appeared likely that the GFS was trapped within a bad local minimum of the search space. This situation was circumvented by increasing population diversity (strategy 3). Although this strategy improved predictions, searching in some runs preconverged at better local minima. Moreover, random replacement of individuals in a population increased the probability of escaping local minima. When strategies 1–3 were implemented, the predictions at the top and the bottom were still suboptimal

(**Figure 4a**). Strategy 4 addressed this problem by extending the fuzzy set variable ranges and changing the defuzzification method (**Figure 4b**). It can be concluded that the additional strategies prevented premature optimization of the GA.[19–21]

Generally, GFSs are capable of finding an OFS that fits the data set well. The advantage is that GFSs are hypothesis free and therefore they do not require assumptions with regard to the biological system that is being evaluated, which is the case for deterministic mechanism–based models such as ordinary differential equations. Furthermore, GFSs can learn information from a real dynamic complex system. Neural networks or genetic neural networks could likewise be used to learn dynamic data.[22] However, neural networks behave as a black box, whereas fuzzy systems provide a description of the data by their rule and knowledge base, offering an advantage as compared with neural network approaches. Although GFSs seem powerful in modeling, there are some drawbacks that should be taken into account. First, as mentioned earlier, the behavior of the GA is unknown and the optimization could become trapped in local minima when the search space is highly dimensional, multimodal, or discrete.[23] The size of the data set for training and the length of the chromosome describing the fuzzy system proportionally increase the complexity of the search space. Therefore, second, GA may be computationally intensive in case the structure of the GFS is large and complicated. Finally, many parameters, such as population size, crossover rate, mutation rate, range for fuzzy sets, and number of fuzzy rules, which control the GFS could affect the results. Any of these might affect the outcome in an unpredictable and nonlinear fashion and it may be a challenge finding the best parameters. This problem could be addressed by implementing fuzzy systems, GA, or other optimization methods, finding optimal parameter values.[24] However, this complicates computation further.

The use of GFSs in pharmacological studies

Behavior of a biological system, e.g., as a result of disease progression or as effect in response to therapy, can be monitored by measuring surrogate end points such as key physiological time series responses. Analyzing these responses and interpreting their consequent or causative action within the biological system can be very challenging due to the
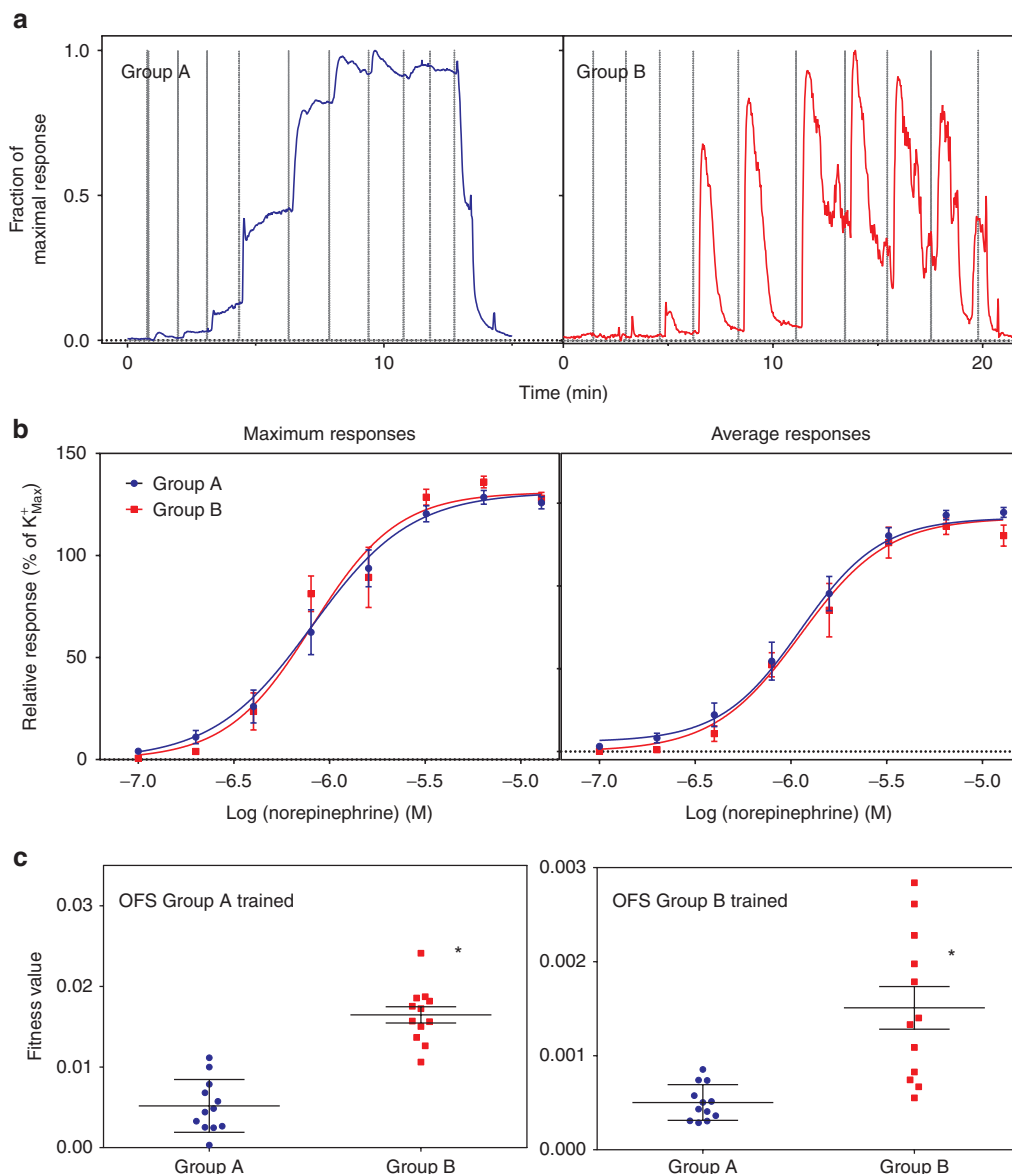
Figure 6 Comparison of OFSs with conventional analysis of data. (**a**) Representative recordings of cumulative concentration–response curves constructed by the addition of increasing concentrations of noradrenaline to isolated arteries that had been mounted in a wire myograph. Small rat mesenteric arteries in Group A had been freshly isolated, whereas those in Group B had been subjected to tissue culture for a period of 24 h (see **Supplementary Data** for details). (**b**) These recordings were analyzed by extracting maximum responses (left) or by averaging all data points per concentration of noradrenaline (right) and expressed as a percentage of contractions induced by a depolarizing high potassium solution. Calculation of noradrenaline potency ($-\text{Log}EC_{50}$) and maximal effects (E_{max}) using nonlinear logistic regression did not show significant differences between groups (for **b** (left panel): noradrenaline – $\text{Log}EC_{50}$ 6.06 ± 0.05 and 6.12 ± 0.05; E_{max} $129.8 \pm 5.4\%$ and $131.5 \pm 5.5\%$, for Groups A and B, respectively; for right panel: noradrenaline – $\text{Log}EC_{50}$ 5.96 ± 0.05 and 5.96 ± 0.05; E_{max} $147.9 \pm 6.1\%$ and $135.3 \pm 6.2\%$, for Groups A and B, respectively). (**c**) fitness values for all artery recordings (per group) that were obtained by running an OFS with five inputs and one output, which was trained on a random artery of either Group A (left panel) or Group B (right panel). The difference between fitness values for these groups was highly significant, indicated by the asterisk (unpaired *t*-test, $P = 0.0001$ and $P = 0.0003$, respectively, for left and right panels).

dynamic nature and inherent adaptive properties of the biological system toward perturbations. As suggested previously by others,[25] fuzzy systems may offer an attractive method to describe these complex systems. The final result of an OFS demonstrates that fuzzy control systems can adapt to the dynamic changes of, as shown in this example, contractile behavior of isolated arteries exposed to external perturbations and can identify differences in effect patterns that are not detectable using conventional analysis.

More importantly, the results shown in **Table 1** and **Figure 6** demonstrate that OFSs, on the basis of their fitness levels, could distinguish between behaviors from different experimental protocols and, therefore, could be applied for advanced data analysis, such as clustering and classification. In cases where conventional statistical methods may fall short, e.g., during preclinical safety assessment of new drugs, OFSs could contribute with objective evaluation of results.

Finally, future research could address whether the rules in OFS may form the basis for mechanistic mathematical modeling of complex high-level physiological parameters in biological systems. Some studies hypothesize that biological systems are managed by rules organized in a hierarchical structure.[26] Fuzzy systems may provide a means to generate these rules by learning from experimental data. Furthermore, OFSs could be applied in other rule-based modeling, such as agent-based modeling.[27]

Future perspective
The herein-presented results show that OFSs can predict experimental data accurately. However, the OFSs described in this study are type 1 fuzzy systems, and these do not handle uncertainty in experimental data optimally. Type 2 fuzzy systems could solve this problem.[28–30] Another shortcoming in the current approach could be that OFSs predict one step ahead, although with high accuracy. Hence, long-term prediction or increasing memory of mode in time could be another feature of this approach because this would be particularly useful in predicting pharmacodynamics of drugs. There are several examples of long-term prediction,[22,31] and the principles behind these methods could be implemented in future GFS models.

Taken together, hypothesis-free GFSs appear very effective in predicting complex time-domain physiological data and can be used in identification, clustering, or classification problems. As such, OFSs could aid in identifying minute but important differences in disease- or drug-induced changes in pivotal physiological parameters, which may not be possible with conventional analysis methods.

METHODS
Data sets
Different archived data sets of isometric tension development of isolated small femoral or mesenteric arteries (of male rats) mounted in a wire myograph were used. One set (data set 1) was used for development of the GFS, and the other (data set 2) for testing the OFS. Arteries had been normalized to their individual optimal diameter for active force development as previously described.[32] Tension data were collected using Myodaq (DMT, Aarhus, Denmark). The data range encompasses a baseline resting tension of fully relaxed arteries and maximal contractile responses, not necessarily the absolute maximum contractile response obtainable in these preparations. Apparent transmural pressure was calculated using the Laplace relation.[32] The data sets represent an evolution of transmural pressure of isolated resistance-sized arteries over time. Furthermore, original data were preprocessed and normalized (see **Supplementary Data**). Data set 1 consisted of two experimental protocols containing four arteries each. These protocols consisted of consecutive single concentrations or cumulative concentration–response curves to various vasoactive substances (see **Supplementary Data**). Data set 2 consisted of two groups, one containing recordings of norepinephrine cumulative concentration–response curves constructed in freshly isolated rat small mesenteric arteries, and the other containing recordings of the same protocol in arteries that

had been kept in tissue culture medium for 24 h. Concentration–response curves were conventionally analyzed either by measuring maximum responses per agonist concentration or by averaging all data points per agonist concentration, taking into account differences in concentration time intervals (Myodata, DMT A/S, Aarhus, Denmark). These data were further analyzed by four-parameter nonlinear logistic regression (Prism, Graphpad, La Jolla, CA), yielding norepinephrine potency (half-maximal effective concentration or EC_{50}) and maximum responses. These values were compared using the *F*-test (Prism, Graphpad, La Yolla, CA). In addition, *OFS(222222)* with five inputs and one output, was trained on whole recordings of a random artery from either Group A or Group B and fitness values of these OFSs for all arteries in both groups were obtained. Fitness values of groups were compared using the unpaired *t*-test.

Genetic fuzzy system
Learning and predicting artery contractile behavior using GFS: Autoregulation and feedback are pivotal in complex biological systems, from the molecular level[33] to the organ and system levels.[34] Autoregulation can maintain a dynamic equilibrium of biological systems even after a sudden or long-time disturbance. Therefore, in our assumption, arteries autoregulate their dimensions around an optimal set point or attractor following perturbations by physical or chemical interventions, and this control can be simulated by a fuzzy control system. In this study, the symbol $GFS(n_1 n_2 \ldots n_k)$(Genetic Fuzzy System) and training result $OFS(n_1 n_2 \ldots n_k)$ (Optimized Fuzzy System) are used to illustrate the structure of fuzzy systems. $n_1 n_2 \ldots n_k$ means the number of fuzzy sets for inputs or outputs. Here, only one output is used; therefore, n_k designates output and all previous numbers' inputs. The first input is the change of error (CE): $CE = P_2 - P_1$; and the second is error (E): $E = S - P_1$, according to Ying.[35] P_1 denotes current pressure, P_2 denotes the next pressure, and S indicates the set point. Moreover, to test the specificity of the fuzzy system, $GFS(n_1 n_2 n_3 n_4)$ and $GFS(n_1 n_2 n_3 n_4 n_5 n_6)$ with additional inputs were implemented (these are described in the **Supplementary Data**). In all implementations, the set point, range of fuzzy sets, membership functions, and rules were unknown and needed to be optimized with training data. Furthermore, the search space for optimization was limited by defining the range for each unknown parameter according to the training data[18] and by experience.

Strategies improving performance: The overall design of the GFS was derived from the study of Shi *et al.*[18] and is further described in the **Supplementary Data**. To improve fitness, four strategies were applied to this algorithm.

Strategy 1: Assigning weight to each rule and increasing number of membership functions.

In theory, rules are not equally predicting the output. To address this difference between rules, a rule weight was assigned between 0 and 1. Twelve membership functions (**Supplementary Figure S1**) were used to ascertain optimal conditions for the fuzzy system to fit the training data. Different functions may have a different number of relevant parameters (two to four parameters). Therefore, to ensure that all individuals in the same population had the same length, all

membership functions were assigned four parameters even though one or two of the four parameters may be considered nonsense data. The first parameter is the start point of the fuzzy set and the fourth is the end point of the same fuzzy set. Furthermore, all parameters were arranged in an ascending order.

Strategy 2: Random generation of membership function parameters and sorting in ascending order.

The varying range of each fuzzy set was defined in the same way as previously described.[18] Each membership function had four parameters, and these generated randomly within the range of the fuzzy set. Then, the lowest and the highest determined the fuzzy set range. Ultimately, each fuzzy set consisted of five parameters, four randomly generated parameters, and a number defining its membership function.

Strategy 3: Increasing population diversity by randomly replacing similar individuals.

Although GAs are powerful in global searching, there is a risk that they become trapped in a local minimum because of the complex search space. Increasing the diversity of the population reduces this risk. First, one of two or more individuals who had the same fuzzy sets and rules with similar or different weights for rules, survived. Then, some individuals died randomly, except elite individuals, who had the best fitness. In order to maintain the size of the population constant, new individuals were generated randomly.

Strategy 4: Extending fuzzy set variable ranges and changing defuzzification.

In this approach, if the first and last fuzzy sets were defined according to the exact range of the data set, some membership functions, e.g., **Supplementary Figure S1d–j** would be mostly ignored. As a result, prediction of data points at the extremes of the output would not be optimal. To circumvent this, extension of the first and last fuzzy sets with two steps[18] into directions beyond the range of inputs or outputs was expected to yield better predictions. In this way, the minimum or maximum value would be at the center of the fuzzy set and could be predicted using the centroid method in a fuzzy system. To further enhance the prediction of data points at the extremes of outputs, a new defuzzification method was introduced by summarizing the results from different defuzzification methods available in Matlab. The reason is that the results from these methods vary over a wide range and a compromise between these methods was expected to better predict results. Finally, multiple threads executed in parallel using the Matlab "Parallel computing toolbox," enhancing speed of computation.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

✓ Genetic fuzzy systems (GFSs) have been applied in many technical fields, such as engineering, computer science, and other control systems. However, their use in pharmaceutical sciences, in particular, predictive pharmacokinetic/pharmacodynamic modeling or safety assessment, has largely been unexplored.

WHAT QUESTION DID THIS STUDY ADDRESS?

✓ This study addresses the question whether a GFS algorithm can be developed that can predict the behavior of a dynamic time-domain physiological parameter by learning information from the data without knowing underlying mechanisms and therefore without predefined assumptions.

WHAT THIS STUDY ADDS TO OUR KNOWLEDGE

✓ GFSs may provide a systems pharmacological approach to the analysis and prediction of complex time-dependent pharmacodynamic effects.

HOW THIS MIGHT CHANGE CLINICAL PHARMACOLOGY AND THERAPEUTICS

✓ GFSs may enable objective evaluation of complex biological system behavior, which in pharmaceutical sciences is often still based on subjective expert interpretation of very few surrogate end points. GFS clustering and classification might enable identification of patient subpopulations and preclinical/clinical safety issues that may not be evident using conventional analysis methods.

1. Jacobsen, J.C., Aalkjaer, C., Nilsson, H., Matchkov, V.V., Freiberg, J. & Holstein-Rathlou, N.H. A model of smooth muscle cell synchronization in the arterial wall. *Am. J. Physiol. Heart Circ. Physiol.* **293**, H229–H237 (2007).
2. Chen, R.R., Silva, E.A., Yuen, W.W. & Mooney, D.J. Spatio-temporal VEGF and PDGF delivery patterns blood vessel formation and maturation. *Pharm. Res.* **24**, 258–264 (2007).
3. Op den Buijs, J., Musters, M., Verrips, T., Post, J.A., Braam, B. & van Riel, N. Mathematical modeling of vascular endothelial layer maintenance: the role of endothelial cell division, progenitor cell homing, and telomere shortening. *Am. J. Physiol. Heart Circ. Physiol.* **287**, H2651–H2658 (2004).
4. Jacobsen, J.C., Mulvany, M.J. & Holstein-Rathlou, N.H. A mechanism for arteriolar remodeling based on maintenance of smooth muscle cell activation. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **294**, R1379–R1389 (2008).
5. Whitacre, J.M. Degeneracy: a link between evolvability, robustness and complexity in biological systems. *Theor. Biol. Med. Model.* **7**, 6 (2010).
6. Edelman, G.M. & Gally, J.A. Degeneracy and complexity in biological systems. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13763–13768 (2001).
7. Weng, G., Bhalla, U.S. & Iyengar, R. Complexity in biological signaling systems. *Science* **284**, 92–96 (1999).
8. Mager, D.E. & Abernethy, D.R. Use of wavelet and fast Fourier transforms in pharmacodynamics. *J. Pharmacol. Exp. Ther.* **321**, 423–430 (2007).
9. Griffith, T.M. & Edwards, D.H. Complexity of chaotic vasomotion is insensitive to flow and pressure but can be regulated by external control. *Am. J. Physiol.* **269**, H656–H668 (1995).

10. Nilsson, H. & Aalkjaer, C. Vasomotion: mechanisms and physiological importance. *Mol. Interv.* **3**, 79–89, 51 (2003).
11. Boonen, H.C. & De Mey, J.G. Increased calcium sensitivity in isolated resistance arteries from spontaneously hypertensive rats: effects of dihydropyridines. *Eur. J. Pharmacol.* **179**, 403–412 (1990).
12. Naranjo, C.A., Bremner, K.E., Bazoon, M. & Turksen, I.B. Using fuzzy logic to predict response to citalopram in alcohol dependence. *Clin. Pharmacol. Ther.* **62**, 209–224 (1997).
13. Cleophas, E.P. & Cleophas, T.J. Fuzzy modeling, a novel approach to studying pharmacodynamics. *Am. J. Ther.* **21**, 20–25 (2014).
14. Sproule, B.A., Naranjo, C.A. & Türksen, I.B. Fuzzy pharmacology: theory and applications. *Trends Pharmacol. Sci.* **23**, 412–417 (2002).
15. Sproule, B.A., Bazoon, M., Shulman, K.I., Turksen, I.B. & Naranjo, C.A. Fuzzy logic pharmacokinetic modeling: application to lithium concentration prediction. *Clin. Pharmacol. Ther.* **62**, 29–40 (1997).
16. Cordon, O. A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems. *Int. J. Approx. Reason.* **52**, 894–913 (2011).
17. Herrera, F. Genetic fuzzy systems: taxonomy, current research trends and prospects. *Evolutionary Intelligence* **1**, 27–46 (2008).
18. Shi, Y., Eberhart, R. & Chen, Y. Implementation of evolutionary fuzzy systems. *IEEE Trans. Fuzzy Syst.* **7**, 109–119 (1999).
19. Lozano, M., Herrera, F. & Cano, J.R. Replacement strategies to preserve useful diversity in steady-state genetic algorithms. *Information Sciences* **178**, 4421–4433 (2008).
20. Sareni, B. & Krahenbuhl, L. Fitness sharing and niching methods revisited. *IEEE Trans. Evol. Comput.* **2**, 97–106 (1998).
21. Zhang, J., Huang, D.S., Lok, T.M. & Lyu, M.R. A novel adaptive sequential niche technique for multimodal function optimization. *Neurocomputing* **69**, 2396–2401 (2006).
22. Nguyen, H.H. & Chan, C.W. Multiple neural networks for a long term time series forecast. *Neural. Comput. Appl.* **13**, 90–98 (2004).
23. Someya, H. & Yamamura, M. A robust real-coded evolutionary algorithm with toroidal search space conversion. *Soft Computing* **9**, 254–269 (2005).
24. Cordón, O., Gomide, F., Herrera, F., Hoffmann, F. & Magdalena, L. Ten years of genetic fuzzy systems: current framework and new trends. *Fuzzy. Set. Syst.* **141**, 5–31 (2004).
25. Mahfouf, M., Abbod, M.F. & Linkens, D.A. A survey of fuzzy logic monitoring and control utilisation in medicine. *Artif. Intell. Med.* **21**, 27–42 (2001).
26. McElroy, M.W. Integrating complexity theory, knowledge management and organizational learning. *J Knowl Manag* **4**, 195–203 (2000).
27. Grimm, V. *et al.* A standard protocol for describing individual-based and agent-based models. *Ecol. Model.* **198**, 115–126 (2006).
28. Mendel, J.M. *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions* (Prentice Hall PTR: New Jersey, USA, 2001).
29. Liang, Q. & Mendel, J.M. Interval type-2 fuzzy logic systems: theory and design. *IEEE Trans. Fuzzy Syst.* **8**, 535–550 (2000).
30. Sepulveda, R., Melin, P., Díaz, A.R., Mancilla, A. & Montiel, O. Analyzing the effects of the footprint of uncertainty in type-2 fuzzy logic controllers. *Engineering Letters* **13**, 138–147 (2006).
31. Herrera, L.J. *et al.* Recursive prediction for long term time series forecasting using advanced models. *Neurocomputing* **70**, 2870–2880 (2007).
32. Mulvany, M.J. & Halpern, W. Contractile properties of small arterial resistance vessels in spontaneously hypertensive and normotensive rats. *Circ. Res.* **41**, 19–26 (1977).
33. Davidson, E. & Levin, M. Gene regulatory networks. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4935 (2005).
34. Fakheri, R.J. & Javitt, N.B. Autoregulation of cholesterol synthesis: physiologic and pathophysiologic consequences. *Steroids* **76**, 211–215 (2011).
35. Ying, H. *Fuzzy Control and Modeling: Analytical Foundations and Applications.* (IEEE Press: Washington DC, USA, 2000).

Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website (http://www.nature.com/psp)