International Journal of Pharmaceutics 434 (2012) 35–42

Contents lists available at SciVerse ScienceDirect

International Journal of Pharmaceutics

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

Establishing and analyzing the design space in the development of direct compression formulations by gene expression programming

Mariana Landin[a,*], Raymond C. Rowe[b], Peter York[c]

[a] Departamento de Farmacia y Tecnología Farmacéutica, Facultad de Farmacia, Campus Vida, Universidad de Santiago de Compostela, Santiago de Compostela, Spain
[b] Intelligensys Ltd., Springboard Business Centre, Stokesley Business Park, Stokesley, North Yorkshire TS9 5JZ, UK
[c] Institute of Pharmaceutical Innovation, University of Bradford, Bradford, West Yorkshire BD7 1DP, UK

ARTICLE INFO

Article history:
Received 30 November 2011
Accepted 26 April 2012
Available online 19 May 2012

Keywords:
Gene expression programming
Quality by design
Design space
Direct compression tablets
Response Surface Methodology

ABSTRACT

Purpose: In this paper we have evaluated the gene expression programming (GEP) methodology for modeling the effect of different variables (continuous and nominal) and their interactions on the properties of direct compression formulations.
Methods: The effect of four variables was studied; variety of diluents, type and percentage of drug and maximum compression force, on the mechanical and drug release properties of direct compression tablets. The generated database (36 formulations) was used for mathematical and GEP modeling.
Results: GEP has been shown to have a high accuracy in prediction for four out five outputs studied including friability which had no replicate measurements. Compared to the traditional statistical treatment GEP is less time consuming and gives equations which are extremely helpful in understanding the interactions of the different variables and for establishing the design space in the development of direct compression formulations.
Conclusions: GEP allows similar conclusions than traditional statistical treatment. The helpfulness of this methodology in establishing the design space has been demonstrated. The knowledge derived from GEP can easily be increased by including additional information or new inputs, such as additional drugs or combinations of excipients in the data set.

1. Introduction

In 2002, the FDA announced a new initiative (cGMPs for the 21st century: A risk-Based Approach) intending to modernize its regulations of pharmaceutical quality for human drugs and establishing a new regulatory framework focused on quality by design (QbD), risk management, and quality systems (FDA, 2003; Jiang and Yu, 2009).

The International Conference on Harmonization guideline (ICH Q8) states that QbD is a systemic approach to development that starts predefining objectives and emphasizes product and process understanding and process control, based on sound science and quality risk management. QbD requires an understanding of how formulation and process variables influence product quality and a definition of the design space inside the knowledge space (Garcia et al., 2008).

ICH Q8 defines the design space as "the multidimensional combination and interaction of input variables (e.g. material attributes) and process parameters that have been demonstrated to provide assurance of quality". Operating within the design space is not considered as a change in a formulation and does not require regulatory oversight; however, movements out of the design space are considered changes and need regulatory approvals. This approach should reduce cost, time and improve process efficiency and quality of the formulations (Zomer et al., 2010). Therefore, for the pharmaceutical industry, adopting the QbD methodology represents an opportunity, but also a great challenge.

Response Surface Methodology (RSM) including statistical experimental designs and multiple linear regression analysis under a set of constrained equations is one recommended method for establishing "the design space" with the inconvenience that nominal factors cannot be included in those designs (Lunney et al., 2008). In those cases, a valid alternative strategy would be to repeat the response surface design for each level of the nominal factors.

Abbreviations: FDA, Food and Drug Administration; GMP, Good Manufacturing Practice; QbD, quality by design; ICH, International Conference on Harmonization; RSM, Response Surface Methodology; GEP, gene expression programming; MSE, Mean Squared Error; SRM, Structural Risk Minimization; AIC, Akaike's Information Criteria; GCV, Generalized Cross Validation; MDL, Minimum Descriptor Length; WVC, weight variation coefficient; CS, crushing strength; *F*, friability; DT, disintegration time; $D_{30\,min}$, drug dissolved after 30 min.

* Corresponding author at: Departamento de Farmacia y Tecnología Farmacéutica, Facultad de Farmacia, Campus Vida, Universidad de Santiago de Compostela, Santiago de Compostela 15782, Spain. Tel.: +34 881815044; fax: +34 981547148.
E-mail address: m.landin@usc.es (M. Landin).

0378-5173/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ijpharm.2012.04.078

In the recent years, together with traditional statistical techniques, soft computing methods, such as neural networks, fuzzy logic and genetic algorithms, are being introduced offering novel solutions of improving control and modeling in pharmaceutics (Colbourn and Rowe, 2009). There are many examples in the literature about the use of neural networks in modeling and optimizing processes and formulation parameters (Ibrić et al., 2003; Subramanian et al., 2004; Arulsudar et al., 2005; Singh et al., 2006; Wei et al., 2008; Ali et al., 2009; Woolfson et al., 2010). Despite its great utility models generated are "black box" and their interpretation is not always easy, especially when the number of variables is large.

Novel approaches such as evolutionary computing can offer even more possibilities and challenges. Gene expression programming (GEP) is an extension of genetic programming, the soft-computing method that simulates the biological evolution process through an algorithm.

One important application of GEP is symbolic regression or function finding, where the goal is to find an expression (equation) that performs well for all fitness cases within a certain error of the correct value (Ferreira, 2006).

The "function finding" application of GEP can be extremely important within the pharmaceutical field. In general, the relationships between response variables and causal factors are not simple and the prediction of response variables on the basis of mathematical expressions using empirically observed values or measurements is a common and important problem to be solved. Moreover, most data-driven systems, modeling or system identification techniques, are developed on the basis of an a priori *known* model structure and focus mainly on the calculation of the model parameter values (Kremer and Hancock, 2006; Siepmann and Siepmann, 2008). GEP is a domain-independent problem-solving technique and its applications have not been fully explored. However, it has been applied successfully in solving some problems within the engineering and food industry fields in the development of new and better materials (Eskil and Kanca, 2008), the prediction of material properties (Antoniou et al., 2010) and the improvement of food processing (Kahyaoglu, 2008). Recently it has been applied to modeling pharmaceutical formulations (Colbourn et al., 2011) where the GEP approach has been compared to neural networks.

Using a desktop computer, researchers can handle by GEP, a large number of variables (inputs and outputs) simultaneously. Thus it can be trained with experimental data to model the interactions between inputs and outputs and obtain a function or an equation (transparent model) relating them. Moreover, it can be combined with other artificial intelligence techniques such as genetic algorithms to perform optimization processes.

This paper investigates the utility of GEP methodology as a tool to model the effect of different variables (continuous and nominal) and their interactions on the properties of direct compression formulations. We compare the GEP results with conventional methodology, ANOVA and stepwise multiple linear regression, and we discuss the helpfulness of this methodology in establishing a design space within the framework of quality by design.

1.1. The gene expression programming algorithm

GEP was invented by Candida Ferreira in 2001 (Ferreira, 2001). A complete description of this algorithm and extensive literature on the subject can be found in her book "Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence" (Ferreira, 2006).

Ferreira (2001) has proposed a new evolutionary algorithm that evolves complex computer programs (neural networks, decision trees, polynomial constructs, mathematical or logical expressions) and encodes them into linear forms named chromosomes. The linear chromosomes can also be expressed or translated into branched structures.

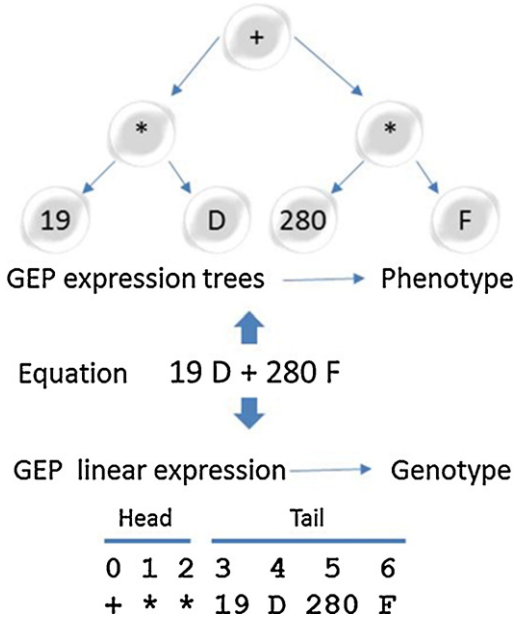

Fig. 1. Example of genotype and phenotype of an equation by GEP.

The principle of genetic algorithm was developed by Koza (1992) in the 90s. Any equation (e.g. $19D+280F$) can be represented by an expression tree and can also be linearized as can be seen in Fig. 1.

In the GEP, linear structures, the chromosomes, represent the genotype and the branched structures, the expression trees, represent the phenotype. They are different entities, structurally and functionally, the discrimination of both, genotype and phenotype, being the fundamental difference between GEP and other technologies like tree algorithms (Ferreira, 2001).

When using GEP, chromosomes are usually composed by more than one gene of equal length. Every gene has a head and a tail (Fig. 1). The head contains symbols that represent functions and terminals, whereas the tail contains only terminals. The set of functions involve any mathematical or Boolean function that the user believes is appropriate to solve the problem (+, −, /, *, exp, etc.). The head length (h) is chosen by the user, whereas the tail length (t) is given by the expression:

$$t = (n-1)h + 1$$

where n is the number of arguments of the function with most arguments.

Each gene is able to code expression trees (phenotypes) of different shapes and sizes, the simplest being composed for only one node and the largest one composed for the number of nodes as the length of the gene.

When modeling, GEP uses populations of individuals (populations of models and solutions for a specific problem); it selects and reproduces them according to fitness, and introduces genetic variations in one or more genetic operators simulating mutations, transpositions or recombinations (Ferreira, 2006). The final result is a mathematical expression or an equation that expresses the relationship between variables (inputs) and results (outputs), which can be used to predict outcomes.

2. Materials and methods

2.1. Data set

The materials used were: two microcrystalline celluloses, Avicel PH 101 (lot 852, C. Barcia, Spain) and Avicel PH 102 (lot 831, C. Barcia, Spain), micronized prednisone (lot 035, J. Escuder, Spain), hydrochlorothiazide (lot 014, J. Escuder, Spain) and magnesium stearate B.P. (lot 548, C. Barcia, Spain). Tablets were produced by direct compression of the mixtures prepared in a Turbula T2C mixer at 30 rpm for 15 min. In all cases, 0.5% (w/w) magnesium stearate as the lubricant was added. For the tabletting process an eccentric tablet machine (Korsch EKO) was used. The machine was equipped with 9 mm flat punches and piezoelectric transducers to measure the upper punch compression force. The lower punch was adjusted to produce 200 mg tablets. The production rate was 33 tablets/min.

The database (Table 1) includes 36 formulations following a factorial design for four variables: variety of microcrystalline cellulose (Avicel PH101 and Avicel PH102 with different nominal particle size), drug (prednisone and hydrochlorothiazide), percentage of drug (0, 5 and 10%, w/w) and maximum compression force (900, 1700 and 2500 N).

Five tablet properties were measured to access the quality of the tablet following US Pharmacopoeia specifications: weight variation coefficient (WVC, %), crushing strength (CS, kg), friability (*F*, %), disintegration time (DT, s), and drug dissolved (%) after 30 min.

2.2. Factorial design and multiple linear regression

Mathematical modeling was carried out to obtain a second order polynomial equation, introducing the terms pointed out as significant by the analysis of variance corresponding to the experimental design employed (Cochram and Cox, 1957) which had been performed previously.

The magnitude of the influences of the terms with the greatest contribution was obtained as regression equations by stepwise multiple regression using the BMDP.P2R package (Dixon, 1983). There were no replicate measurements for the parameters coefficient of variation of weight and percentage of friability, hence the ANOVA was not possible and the effects of the variables were assessed by introducing all the terms, both independently and as interactions, into the regression.

2.3. Gene expression programming

A commercial computer program INForm v4.11, supplied by Intelligensys Ltd., 2010, UK, was used in this study. A separate model was developed for each property. The accuracy of the GEP model is assessed using the correlation coefficient (R^2) for each output and the ANOVA parameters (*f* ratio, degrees of freedom).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}$$

where $\bar{y}$ is the mean of the dependent variable and $\hat{y}$ is the predicted value from the model.

The larger the value of the Train Set R^2, the more the model captured the variation in the training data. Values between 70 and 99.9% are indicative of reasonable model predictabilities (Colbourn and Rowe, 2005).

Type of microcrystalline cellulose (expressed by the commercial brand, Avicel PH 101 and Avicel PH 102), type of drug (expressed

Table 1
Differential characteristics of the formulations studied and mean values of the parameters used to characterize them.

Drug	Excipient	% drug	Compression, force (N)	WVC (%)	*F* (%)	CS (kg)	DT (s)	$D_{30\,min}$ (%)
Prednisone	Avicel PH101	0	900	0.48	0.19	3.18	12	0.00
		0	1700	1.40	0.00	6.50	16	0.00
		0	2500	0.97	0.00	9.17	17	0.00
		5	900	1.35	1.60	2.67	4	100.00
		5	1700	1.41	0.44	5.42	11	77.92
		5	2500	1.99	0.23	8.58	16	37.66
		10	900	0.49	2.88	2.33	10	100.00
		10	1700	0.77	0.51	5.83	13	99.62
		10	2500	0.78	0.25	8.42	12	69.57
	Avicel PH102	0	900	0.67	0.45	4.50	16	0.00
		0	1700	0.77	0.19	9.00	37	0.00
		0	2500	0.67	0.10	11.42	85	0.00
		5	900	0.32	0.81	3.08	14	92.75
		5	1700	0.32	0.54	6.78	23	43.83
		5	2500	0.22	0.49	9.00	32	13.30
		10	900	0.63	1.03	3.33	9	98.03
		10	1700	0.48	0.67	6.33	17	72.72
		10	2500	0.63	0.37	7.42	21	52.27
HCT	Avicel PH101	0	900	0.48	0.19	3.18	12	0.00
		0	1700	1.40	0.00	6.50	16	0.00
		0	2500	0.97	0.00	9.17	17	0.00
		5	900	0.35	0.53	3.25	16	96.97
		5	1700	0.43	0.16	6.17	21	85.91
		5	2500	1.15	0.02	10.00	28	75.45
		10	900	0.48	0.74	2.67	9	93.71
		10	1700	0.48	0.18	6.50	16	92.50
		10	2500	0.58	0.04	9.83	20	78.48
	Avicel PH102	0	900	0.67	0.45	4.50	16	0.00
		0	1700	0.77	0.19	9.00	37	0.00
		0	2500	0.67	0.10	11.42	85	0.00
		5	900	1.29	0.45	4.58	14	98.18
		5	1700	1.09	0.38	8.08	29	69.98
		5	2500	1.14	0.24	10.67	55	33.18
		10	900	0.43	0.33	5.75	21	96.89
		10	1700	0.37	0.29	9.17	35	51.29
		10	2500	0.29	0.27	11.00	62	25.76

Table 2
GEP training parameters setting with Inform v4.11 for modeling each parameter.

	WVC (%)	*F* (%)	CS (kg)	DT (s)	$D_{30\,min}$ (%)
General					
Number of populations	10	10	10	10	10
Population size	1000	1000	1000	1000	1000
No. generations	1000	1000	200	200	1000
Head length	7	7	7	7	7
No. genes	4	4	4	4	4
Random seed	1	1	1	1	10
Fitness type	MDL	SRM	MDL	MDL	SRM
C1	–	1	–	–	1
C2	–	4.8	–	–	4.8
Node weighting factor	0.1	0.1	0.1	0.1	0.1
Proportion of elite	0.05	0.05	0.05	0.05	0.05
Genetic operators					
One point recombination			0.3		
Two point recombination			0.3		
Gene recombination			0.1		
Transposition of IS			0.1		
Transposition of root			0.1		
Gene transposition			0.1		
Mutation			2		
Constants mutation			1		
Using random number constants? Yes					
Function set	+ −/*	+ −/* exp	+ −/*	+ −/*	+ −/* exp
Input connected					
Excipient			Yes		
Drug			Yes		
% drug			Yes		
Compression force			Yes		

by the name, prednisone and hydrochlorothiazide), percentage of drug in the formulation (%) and compression force (N) were introduced as ingredients or process conditions (inputs) and weight variation coefficient, crushing strength, percentage of friability, disintegration time and percentage of drug dissolved at 30 min were selected as properties (outputs).

INForm v4.1 contains various statistical fitness criteria including Mean Squared Error (MSE), Structural Risk Minimization (SRM), Akaike's Information Criteria (AIC), Generalized Cross Validation (GCV) and Minimum Descriptor Length (MDL). All were investigated to obtain the model that gave the best R^2, for each property measured and additionally the simplest and statistical significant equations. The best results for each output were found with different training parameters which are presented in Table 2.

3. Results and discussion

3.1. Models predictability

Table 3 shows the GEP model equations and the corresponding model statistics. It is interesting to note that the GEP approach allows most of the outputs with a reduced number of terms to be predicted and GEP performs well in some cases such as for the friability parameter, which has been reported as a difficult parameter for obtaining good models in making predictions (Plumb et al., 2005; Shao et al., 2006). Most predictions are similar or even better than the regression equations obtained by stepwise linear regression (Landín et al., 1992, 1993).

Fig. 2 shows the correlations between experimental values and predicted values using both GEP models and stepwise linear regression models (Landín et al., 1992, 1993). As can be seen the correlation coefficients are quantitatively similar or even higher for the GEP, even for the parameter weight variation coefficient despite the model being not statistically significant for this parameter.

3.2. Mechanical properties

As two variables, type of drug and binder are nominal, the GEP equations for the parameters related to the mechanical properties of tablets (Figs. 3 and 4) allow four GEP predicted response surfaces for crushing strength and friability to be represented.

Using the GEP equations and the 3D plots for these parameters and interpreting both simultaneously the following features can be easily deduced:

1. The main effect on the mechanical properties is compression force followed by type of Avicel and percentage of drug. An increment in compression force allows harder tablets with low friability to be obtained.

Table 3
Equations and statistical parameters for the derived GEP models, where E is the variety of microcrystalline cellulose, M is the type of drug, D is the percentage of drug in the formulation and F is the maximum compression force.

Outputs	Model equations	Training R^2	d.f.	*f* ratio
WVC (%)	$= 0.52 + 0.17M + (0.02/(0.07 + E - D)) + (0.27/(6.40 - M - D + E))$	68.07	22 and 35	1.24*
CS (kg)	$= 3.72 \times 10^{-3}F - 0.4983 + ((E + 0.62)/(1.7 - E + (M \times D)))$	96.33	18 and 35	23.70**
F (%)	$= 19\,DF^{-1} + 160\,MDF^{-1}10^{-E} + 280\,F^{-1}$	81.58	19 and 35	3.27**
DT (s)	$= 2.7 \times 10^{-3} + D + 7.1 \times 10^{-3}\,F - MD - EMD + ((ED + (0.1\,EF))/(4.8 + D))$	86.84	24 and 35	2.81**
$D_{30\,min}$ (%)	$= (EF/(4.16\,D - 96)) + (0.23\,F/(D - 22)) + 20.4\,D + 20 + 15\,E - E\,D - D^2$	91.20	26 and 35	2.25**

* Not statistically significant.
** Statistically significant $\alpha < 0.01$.

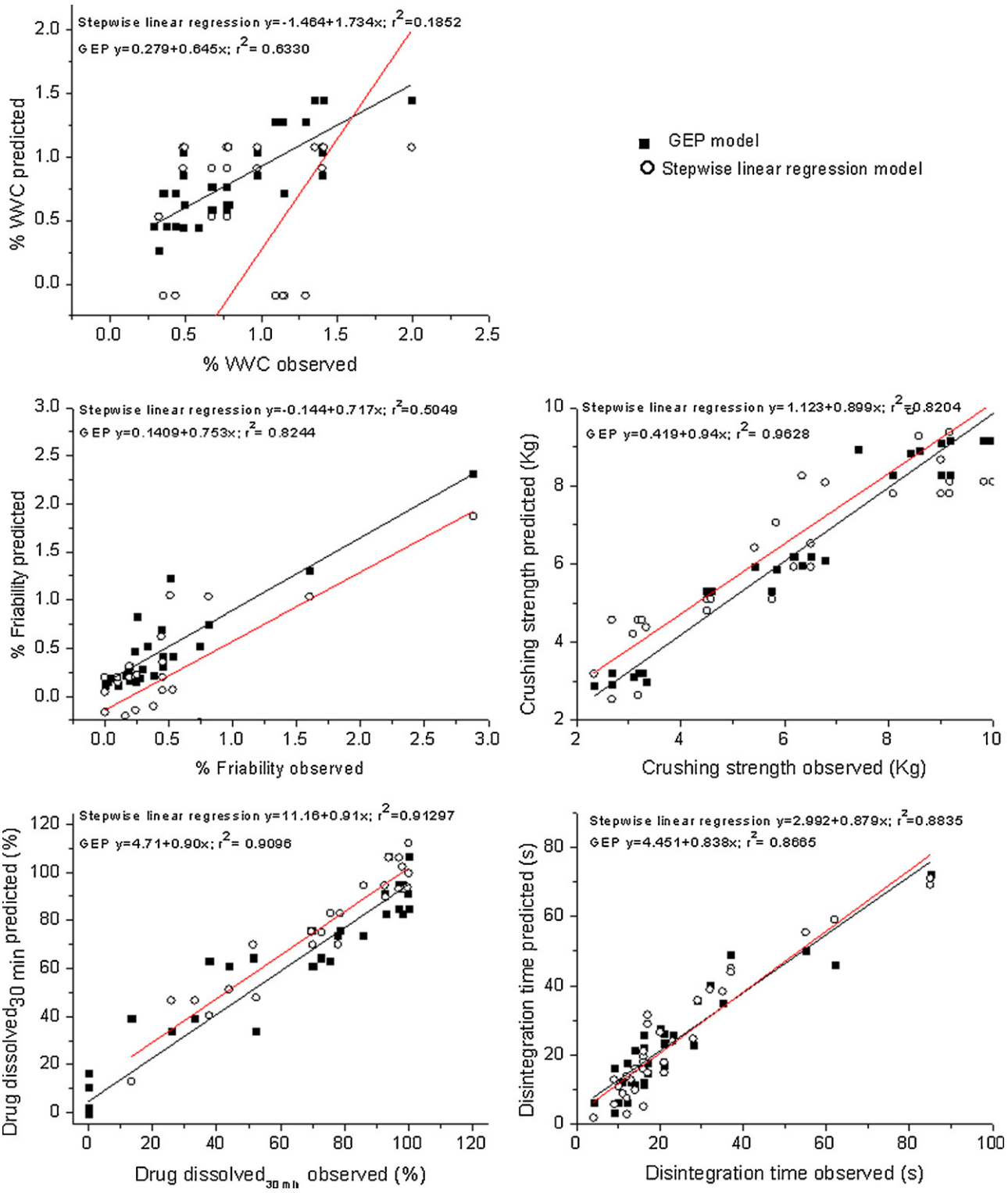

Fig. 2. Experimental against predicted values by stepwise linear regression equations (Landín et al., 1992, 1993) and GEP models for the different parameters studied.

2. Mechanical properties of Avicel PH 102 based tablets are better than mechanical properties of Avicel PH 101 whatever the formulated drug.
3. Prednisone affects negatively tablet mechanical properties, the percentage of drug interacting with compression force, in such a way that for high percentages of prednisone the compression force should be high to avoid tablets being out of USP limit for friability (<1%), especially when Avicel PH 101 is used.

3.3. Tablet disintegration time and drug dissolution profile

With regard to those parameters, it should be pointed out that GEP equations include a high number of terms for both the disintegration time and the percentage of drug dissolved (Table 3), but surprisingly the type of drug is not included in the GEP equation for $D_{30\,min}$. Differences in their hydrosolubility (both poorly soluble) or their mechanisms of compression (plastic deformation or particle

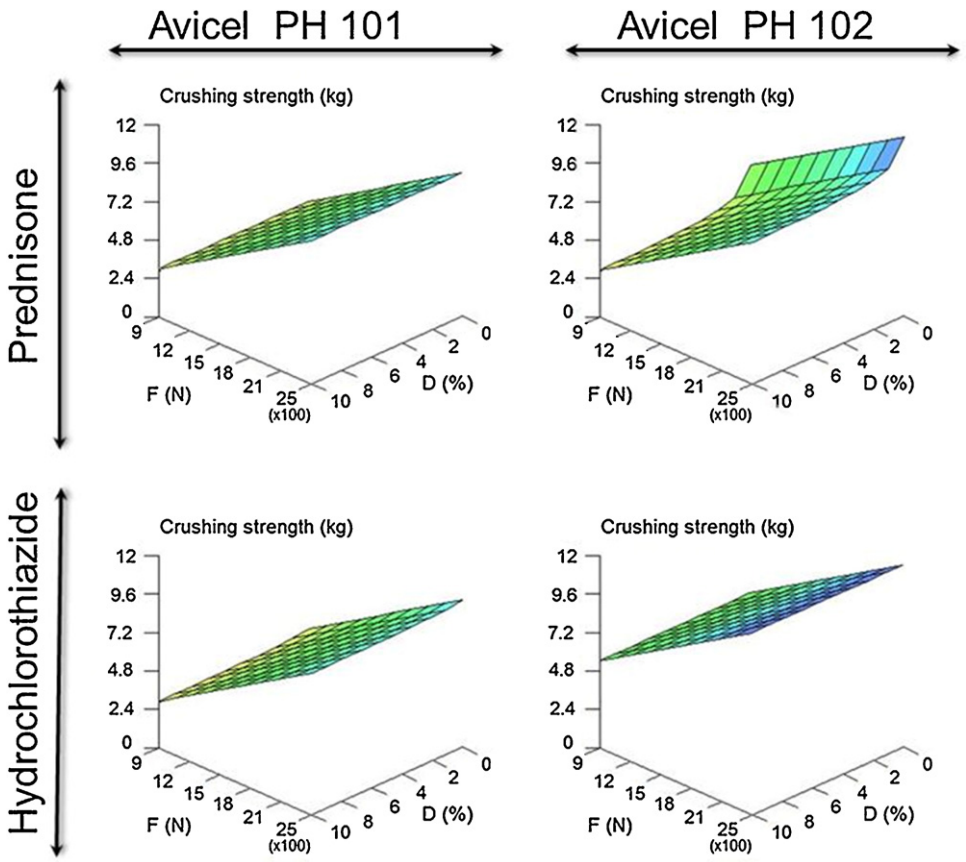

Fig. 3. Effect of percentage of drug and compression force on the crushing strength predicted by GEP.

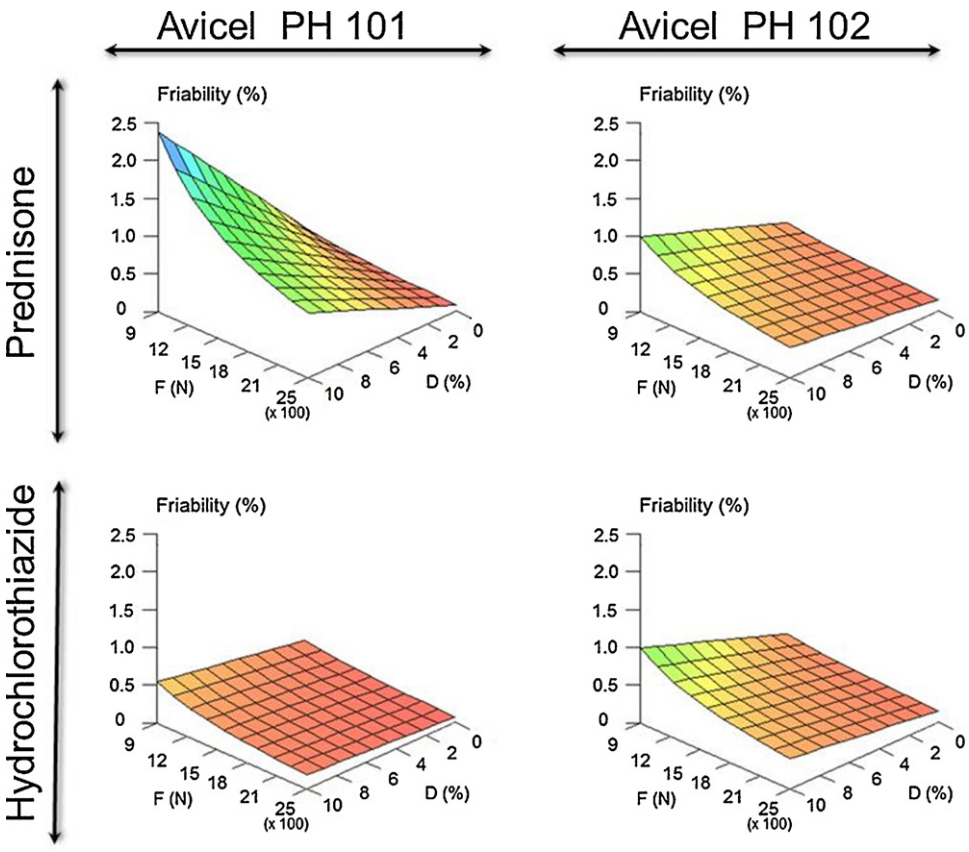

Fig. 4. Effect of percentage of drug and compression force on the loss of weight by friability predicted by GEP.

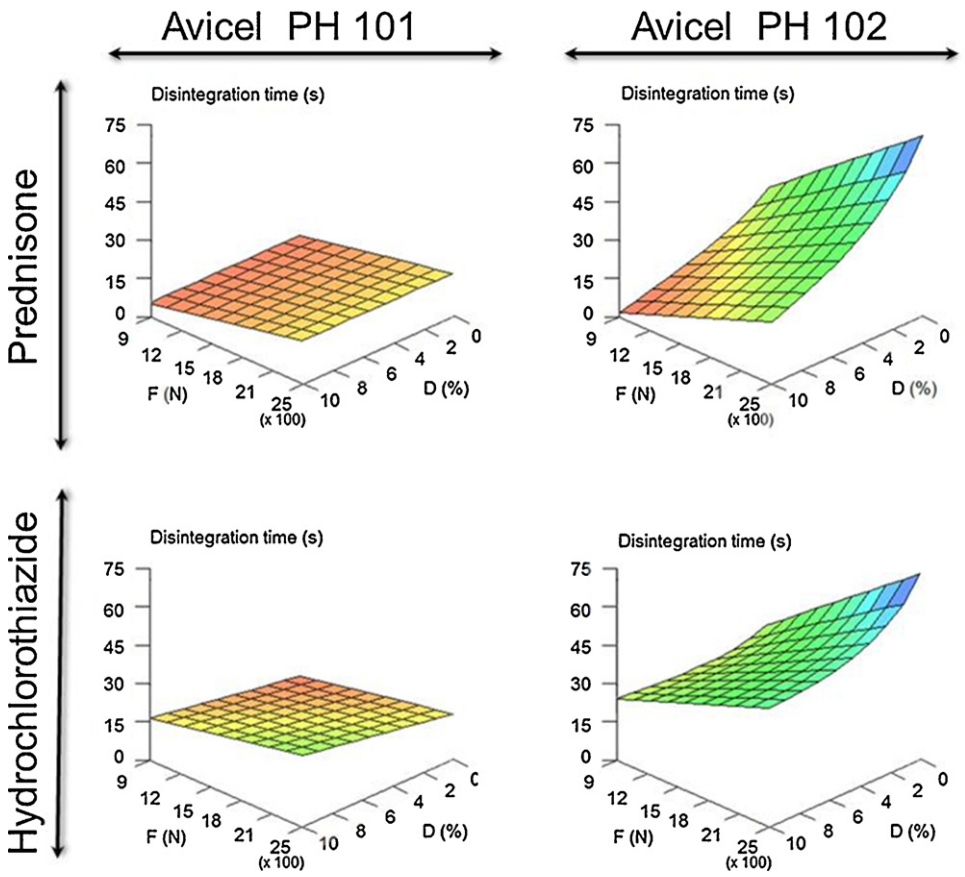

Fig. 5. Effect of percentage of drug and compression force on the tablet disintegration time predicted by GEP.

fragmentation) do not explain the variability in $D_{30\,min}$, this being a parameter dependent on the characteristics of the binder and other process variables.

The interpretation of 3D plots (Figs. 5 and 6) allowed the deductions that:

1. A linear effect of compression force on both DT and especially on $D_{30\,min}$ is observed.
2. Avicel PH101 gives tablets of shorter disintegration time and quicker drug release than Avicel PH102 whatever the formulated drug.
3. There is an interaction between the compression force and percentage of drug in the formulation. At high compression forces, tablets, especially Avicel PH102-based ones, can be out of limits for the parameter $D_{30\,min}$ (the USP limits for $D_{30\,min}$ are >60% and >80% for hydrochlorothiazide and prednisone-based tablets respectively).

The knowledge generated from GEP modeling for both parameters is completely in agreement with the findings generated by the traditional approach (Landín et al., 1992, 1993). Moreover, the GEP equations can be used to define the design space for a formulation as they describe the relationships between variables. As an example, Fig. 7 represents the superimposed contour graphs from GEP for the parameters friability and $D_{30\,min}$ for prednisone-Avicel

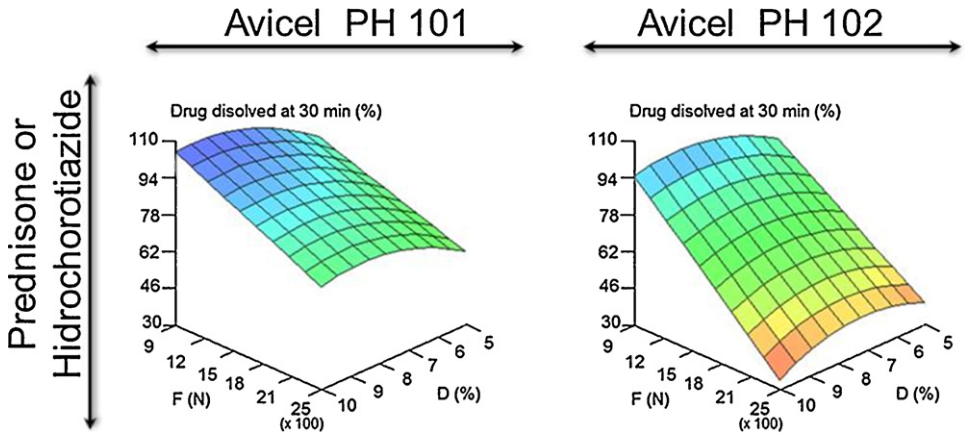

Fig. 6. Effect of percentage of drug and compression force on the percentage of drug dissolved at 30 min predicted by GEP.

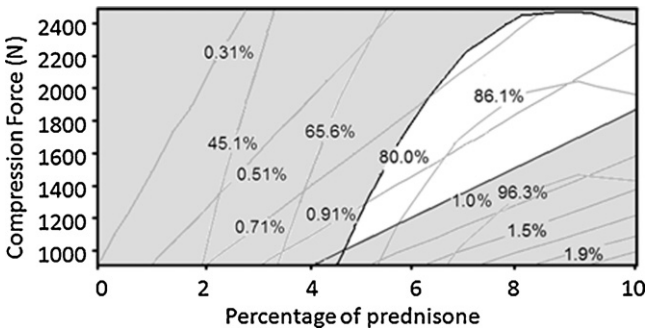

Fig. 7. Example of the design space predicted by GEP for prednisone-based tablets when Avicel 101 is used as binder (the white area represents tablets within the USP limits; $D_{30\,min} > 80\%$ and $F < 1\%$).

PH101 based tablets. The white area represents the common region of successful operating ranges.

4. Conclusions

In this study gene expression programming technology has been applied within the pharmaceutical field, to a direct compression tablet formulation data set. Results show that GEP allows similar conclusions to the obtained traditional statistical treatment. GEP was shown to have a high accuracy in prediction for the four out five outputs studied (Table 3) including one which had no replicate measurements like friability.

GEP is less time consuming than ANOVA or stepwise linear regression and gives equations (transparent model) which are extremely helpful in understanding the interactions of the different variables and for establishing the design space in the development of direct compression formulations.

The knowledge derived from GEP can easily be increased by including additional information or new inputs, such as additional drugs or combinations of excipients in the data set.

Conflict of interest

The authors declare no conflicts of interest.

Acknowledgments

Authors thank Ms. J. Menis for her help in the correction of the English version of the work. This work was supported by the Regional Government of Xunta de Galicia [exp.2007/097]. ML thanks the Spanish MEC [PR2010-0460] for her financial support during her sabbatical at Faculty of Science, University of Utrecht, Netherlands.

References

Ali, H.S.M., Blagden, N., York, P., Amani, A., Brook, T., 2009. Artificial neural networks modelling the prednisolone nanoprecipitation in microfluidic reactors. Eur. J. Pharm. Sci. 37, 514–522.

Antoniou, M.A., Georgopoulos, E.F., Theofilatos, K.A., Vassilopoulos, A.P., Likothanassis, S.D., 2010. A gene expression programming environment for fatigue modeling of composite materials. Lect. Notes Comput. Sci. 6040, 297–302.

Arulsudar, N., Subramanian, N., Murthy, R.S.R., 2005. Comparison of artificial neural network and multiple linear regression in the optimization of formulation parameters of leprolide acetate loaded liposomes. J. Pharm. Pharm. Sci. 8, 243–258.

Dixon, W.J., 1983. BMDP Statistical Software. University of California Press, Berkeley.

Cochram, W.G., Cox, G.W., 1957. Experimental Designs, second ed. J Willey, London.

Colbourn, E.A., Rowe, R.C., 2005. Neural computing and pharmaceutical formulation. In: Swarbrick, J., Boylan, J.C. (Eds.), Encyclopedia of Pharmaceutical Technology. , third ed. Marcel Dekker.

Colbourn, E.A., Rowe, R.C., 2009. Novel approaches to neural and evolutionary computing in pharmaceutical formulation: challenges and new possibilities. Future Med. Chem. 1, 713–726.

Colbourn, E.A., Roskilly, S.J., Rowe, R.C., York, P., 2011. Modelling formulations using gene expression programming—a comparative analysis with artificial neural networks. Eur. J. Pharm. Sci. 44, 366–374.

Eskil, M., Kanca, E., 2008. A new formulation for martensite start temperature of Fe–Mn–Si shape memory alloys using genetic programming. Comput. Mater. Sci. 43, 774–784.

FDA, 2002. Final Report on Pharmaceutical CGMPS for the 21st Century—A Risk-Based Approach (2003) http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/Manufacturing/QuestionsandAnswersonCurrentGoodManufacturingPracticescGMPforDrugs/UCM176374.pdf (accessed March 2011).

Ferreira, C., 2001. Gene expression programming: a new adaptive algorithm for solving problems. Complex Syst. 13, 87–129.

Ferreira, C., 2006. Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence, second ed. Springer-Verlag, Germany.

Garcia, T., Cook, G., Nosal, R., 2008. PQLI key topics-criticality, design space and control strategy. J. Pharm. Innov. 3, 60–68.

Ibrić, S., Jovanović, M., Djurić, Z., Parojčić, Petrovič, S.D., Solomun, L., Stupar, B., 2003. Artificial neural networks in the modeling and optimization of aspirin extended release tablets with Eudragit L 100 as matrix substance. AAPS PharmSciTech 4, 1–9.

Jiang, W., Yu, L.X., 2009. Modern pharmaceutical quality regulations: question based review. In: Qiu, Y., Chen, Y., Zhang, G.G.Z. (Eds.), Developing Solid Oral Dosage Forms. Pharmaceutical Theory and Practice. Academic Press, Burlington, pp. 885–901.

Kahyaoglu, T., 2008. Optimization of the pistachio nut roasting process using response surface methodology and gene expression programming. LWT 41, 26–33.

Koza, J.R., 1992. Genetic Programming: On the Programming of Computers by Means of Natural Selection. The MIT Press, Cambridge, MA.

Kremer, D.M., Hancock, B.C., 2006. Process simulation in the pharmaceutical industry: a review of some basic physical models. J. Pharm. Sci. 95, 517–529.

Landín, M., Vazquez, M.J., Souto, C., Concheiro, A., Gómez-Amoza, J.L., Martínez-Pacheco, R., 1992. Comparison of two varieties of microcrystalline cellulose as filler-binders I. Prednisone tablets. Drug Dev. Ind. Pharm. 18, 355–369.

Landín, M., González, M.P., Souto, C., Concheiro, A., Gómez-Amoza, J.L., Martínez-Pacheco, R., 1993. Comparison of two varieties of microcrystalline cellulose as filler-binders II. Hydrochlorothiazide tablets. Drug Dev. Ind. Pharm. 19, 1211–1220.

Lunney, P.D., Gogdill, R.P., Drennen, J.K., 2008. Innovation in pharmaceutical experimentation. Part 1: review of experimental designs used in industrial pharmaceutics research and introduction to Bayesian D-optimal experimental design. J. Pharm. Innov. 3, 188–203.

Plumb, A.P., Rowe, R.C., York, P., Brown, M., 2005. Optimisation of the predictive ability of artificial neural network (ANN) models: a comparison of three ANN programs and four classes of training algorithm. Eur. J. Pharm. Sci. 25, 395–405.

Shao, Q., Rowe, R.C., York, P., 2006. Comparison of neurofuzzy logic and neural networks in modeling experimental data of an immediate release tablet formulation. Eur. J. Pharm. Sci. 28, 394–404.

Siepmann, J., Siepmann, F., 2008. Mathematical modeling of drug delivery. Int. J. Pharm. 364, 328–343.

Singh, B., Chakkal, S.K., Ahuja, N., 2006. Formulation and optimization of controlled release mucoadhesive tablets of atenolol using response surface methodology. AAPS PharmSciTech 7, E1–E10.

Subramanian, N., Yajnik, A., Murthy, R.S.S., 2004. Artificial neural network as an alternative to multiple regression analysis in optimizing formulation parameters of Cytarabine liposomes. AAPS PharmSciTech 5, 1–9.

Wei, H., Zhong, F., Ma, J., Wa, Z., 2008. Formula optimization of emulsifiers for preparation of multiple emulsions based on artificial neural networks. J. Dispersion Sci. Technol. 29, 319–326.

Woolfson, A.D., Umrethia, M.L., Kett, V.L., Malcolm, R.K., 2010. Freeze-dried, mucoadhesive system for vaginal delivery of the HIV microbicide, dapivirine: optimisation by an artificial neural network. Int. J. Pharm. 388, 136–143.

Zomer, S., Guptam, M., Scott, A., 2010. Application of multivariate tools in pharmaceutical product development to bridge risk assessment to continuous verification in a quality by design environment. J. Pharm. Innov. 5, 109–118.