Journal of Pharmaceutical Sciences xxx (2016) 1-9

Contents lists available at ScienceDirect

Journal of Pharmaceutical Sciences

journal homepage: www.jpharmsci.org

Pharmaceutics, Drug Delivery and Pharmaceutical Technology

Predicting the Fine Particle Fraction of Dry Powder Inhalers Using Artificial Neural Networks

Joanna Muddle [1, 2], Stewart B. Kirton [3, *], Irene Parisini [3], Andrew Muddle [4], Darragh Murnane [3], Jogoth Ali [3], Marc Brown [3, 4], Clive Page [2], Ben Forbes [1]

[1] Institute of Pharmaceutical Science, King's College London, 150 Stamford Street, London SE19NH, UK
[2] Sackler Institute of Pulmonary Pharmacology, Institute of Pharmaceutical Science, King's College London, 150 Stamford Street, London SE19NH, UK
[3] Department of Pharmacy, University of Hertfordshire, Hatfield, Herts AL109AB, UK
[4] MedPharm Ltd, R&D Centre, Units 1 and 3 Chancellor Court, 50 Occam Road, Surrey Research Park, Guildford GU27AB, UK

ARTICLE INFO

Article history:
Received 9 July 2016
Revised 26 September 2016
Accepted 6 October 2016

Keywords:
artificial neural networks
in silico modeling
dry powder inhaler
next-generation impactor
in vitro performance
fine particle fraction

ABSTRACT

Dry powder inhalers are increasingly popular for delivering drugs to the lungs for the treatment of respiratory diseases, but are complex products with multivariate performance determinants. Heuristic product development guided by *in vitro* aerosol performance testing is a costly and time-consuming process. This study investigated the feasibility of using artificial neural networks (ANNs) to predict fine particle fraction (FPF) based on formulation device variables. Thirty-one ANN architectures were evaluated for their ability to predict experimentally determined FPF for a self-consistent dataset containing salmeterol xinafoate and salbutamol sulfate dry powder inhalers (237 experimental observations). Principal component analysis was used to identify inputs that significantly affected FPF. Orthogonal arrays (OAs) were used to design ANN architectures, optimized using the Taguchi method. The primary OA ANN r^2 values ranged between 0.46 and 0.90 and the secondary OA increased the r^2 values (0.53-0.93). The optimum ANN (9-4-1 architecture, average r^2 0.92 $\pm$ 0.02) included active pharmaceutical ingredient, formulation, and device inputs identified by principal component analysis, which reflected the recognized importance and interdependency of these factors for orally inhaled product performance. The Taguchi method was effective at identifying successful architecture with the potential for development as a useful generic inhaler ANN model, although this would require much larger datasets and more variable inputs.

Introduction

Dry powder inhalers (DPIs) are increasingly popular for delivering drugs to the lungs for the treatment of respiratory diseases. Over the last 2 decades, device and formulation technologies have developed considerably, along with the scientific understanding of the determinants of inhaler performance.[1-7] *In vitro* impactor deposition is the principal pharmaceutical performance assay and is an essential assessment during the development and registration of new originator or generic DPI formulations.[8] The fine particle fraction (FPF) determined using *in vitro* deposition techniques has been used as a DPI performance characteristic for mechanistic modeling,[6] *in vitro–in vivo* correlation,[9] and to make estimations of clinical relevance.[5] FPF is defined as the proportion of the particles that are <5 μm in diameter, that is, the respirable dose. Modeling the influence of formulation and device variables on FPF could facilitate the development of the next generation of inhaled medicines[10] and the matching of test inhalers to reference products during the development of generic products.[8,11]

As the patents for inhaled drugs expire, many companies are expending much effort in developing generic versions of innovator inhaled medicines. The U.S. Food and Drug Administration and European Medicine Agency have clear processes by which a generic orally inhaled product can be developed for the market,[12] both of which include extensive *in vitro* studies.[13,14] The U.S. Food and Drug Administration recently issued specific guidance for salmeterol-fluticasone combinations aiming for registration as a generic Advair[15,16] including the requirement to match FPF. For generic product development, the matching of *in vitro* performance is the obvious first stage in product development. However, heuristic

* *Correspondence to:* Stewart B. Kirton (Telephone: +44 (0)1707 285203; Fax: +44 (0) 1707 284115).
E-mail address: s.b.kirton3@herts.ac.uk (S.B. Kirton).

http://dx.doi.org/10.1016/j.xphs.2016.10.002
0022-3549/

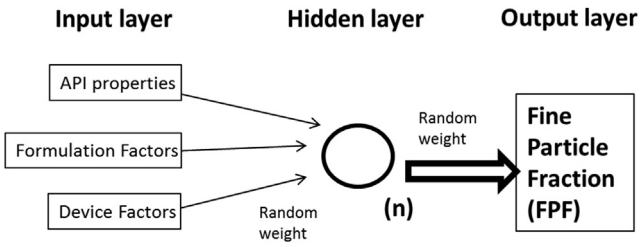

Figure 1. A simplified structure for a supervised ANN showing the feedforward links from the input layer to one neuron of the hidden layer which subsequently feeds onto the output layer (prediction of FPF). Input layer elements are the properties of API, formulation factors, and device factors. The output layer is product performance in terms of FPF.

development of a product guided by *in vitro* FPF is a costly and time-consuming process. There has already been interest in using *in silico* models to streamline this process and there is a belief that modeling may become increasingly crucial in the design and development of inhaled drug products.[17] Linear regression models have been used to understand the complex formulation characteristics of DPIs,[6,18] for example, Kinnunen et al.[19] showed a best r^2 value of 0.861 between the physical properties of the carrier (amount of fines which are <4.5 μm in size) and the *in vitro* performance of the formulation, for example, FPF. Regression models make the assumption that relationships between physiochemical factors are linear. As such, prediction of pharmaceutical response based on these polynomial relationships is limited and can provide an incomplete estimate of the observed response in experimental determinations. Therefore, there is a need to investigate non-linear models as an alternative approach to determine whether or not this improves model quality.[20] The use of artificial neural networks (ANNs) is one such technique that could be exploited.

ANNs are computational models that replicate the way the human brain processes information, by linking input variables to a desired output via the selective activation of artificial "neurons" in a complex network (Fig. 1).[21] The ANN is typically separated into 3 sections: the input layer, the hidden layer, and the output layer. The input layer contains information that will be used as stimuli for the ANN, such as experimental variables. The elements of the input layer are connected to the neurons of the hidden layer via links. Each link has a weight value (w) associated with it, and each neuron receives numerical input from each of the elements of the input layer (element value × link weight). The hidden layer neuron can exist in one of the 2 states: inactive or active. Whether or not a neuron is active or inactive is determined by an activation threshold function. If the sum of the numerical inputs from the elements of the input layer exceeds the threshold value then the neuron is activated. Activated neurons then communicate with the output layer. Each neuron in the hidden layer is connected via a link (with an associated weight, w) to each element in the output layer. The elements in the output layer define one of a possible number of permissible outcomes from the input. The output value(s) selected by an ANN is a function of the neurons activated in the hidden layer.

Supervised ANNs, such as the ones used in this investigation, work by training the network to generate a desired output from input variables using an iterative process of calculating and minimizing the error between the generated and expected output value.[22] Error is minimized by manipulating the weights associated with the links between the 3 layers of the ANN via a process known as backpropogation.[23]

ANNs are powerful pattern recognition tools and have been utilized for analyzing inhaled drug delivery. Nazir et al.[24,25] produced ANNs for predicting the regional and total aerosol deposition in the human lungs and De Matas et al.[26-28] produced ANNs that predicted deposition and clinical effects for pulmonary drug delivery (improvements in the forced expiratory volume in 1 s and urinary excretion of the drug in 24 h). To create an optimal ANN architecture, many models need to be created and tested, which can be challenging and time-consuming. Hence, any techniques that can help to reduce the number of architectures that need to be investigated are useful with respect to streamlining the development process. As such, the Taguchi method has been widely used in formulation development[29,30] and in designing the architecture of ANNs.[31-33] Design of experiments has also been used to evaluate the effect of lactose size fractions on the performance of the dry powder formulation,[2] but has not been used to design an optimum ANN predicting the *in vitro* performance of a DPI. The objective of this study was to investigate the feasibility of using ANNs to predict FPF based on formulation device variables. The Taguchi method, which has not previously been applied to ANN to predict the FPF of DPI, was used to develop an optimized ANN.

Methods

Materials

Unscrambler™ was obtained from CAMO Software (Oslo, Norway), while Minitab™ was from On Line Computers (York, UK). Neurosolutions™ was from Neuro-dimension (Gainsville, FL).

Experimental Data for Modeling

The self-consistent dataset used in this study was assembled from studies into the effect DPI formulation factors on FPF reported by Hassoun et al.,[34] Muddle et al.,[35] and Parisini[36] (Table 1) where we define a self-consistent dataset as experimental data for a range of compounds and/or formulation types that has been obtained using identical experimental procedures. Two of these studies, investigating the effect of the carrier lactose particle, total fine lactose content and device resistance on FPF of salmeterol xinafoate[34] and salbutamol sulfate,[35] were performed in the same manner to allow direct comparison. In brief, the active pharmaceutical ingredient (API) was blended with 3 coarse lactose grades (Respitose® ML001, Respitose ML006, and Lactohale® LH200) and different amounts of fine lactose were added (0%-20%) (Lactohale LH300). These formulations were tested using a next-generation pharmaceutical impactor to measure their *in vitro* deposition with 3 different inhalers (Aeroliser®, Handihaler®, and Rotahaler®). The third study[36] investigated salmeterol xinafoate and salbutamol sulfate blended with a different coarse lactose (Respitose SV003) ± addition of 5% fine lactose. These 4 formulations were then tested using a Cyclohaler® at 2 different pressure drops (2 and 4 kPa). Thus, the combined dataset had 2 APIs (salmeterol xinafoate and salbutamol sulfate) tested with 4 different inhalers (Aeroliser, Handihaler, Rotahaler, and Cyclohaler) and 4 different coarse lactoses blended with a variety of different fine lactose ratios to allow a large range of total fines content to be examined (Table 1).

To allow the datasets to be combined as described above, the device factors, resistance, test flow rate (Q), and pressure drop (P) were converted to a value for the power generated by simulated "inhalation" through the inhaler at each flow rate using an established method[37] (Eq. 1). This enabled the 3 original datasets to be combined by overcoming the problem that the 3 inhalers were originally tested at different flow rates.

$$Power = P.Q. \quad (1)$$

The records in the combined dataset (237 records) were split into a training set (60% of the records), a cross validation (CV, 20%