European Journal of Pharmaceutical Sciences 41 (2010) 421–429

Contents lists available at ScienceDirect

European Journal of Pharmaceutical Sciences

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

Analysis of pellet properties with use of artificial neural networks

Aleksander Mendyk[a,*], Peter Kleinebudde[b], Markus Thommes[b], Angelina Yoo[b], Jakub Szlęk[a], Renata Jachowicz[a]

[a] Department of Pharmaceutical Technology and Biopharmaceutics, Jagiellonian University, Medical College, Faculty of Pharmacy, Medyczna 9 St, 30-688 Kraków, Poland
[b] Institute of Pharmaceutics and Biopharmaceutics, Heinrich-Heine University of Duesseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany

ARTICLE INFO

Article history:
Received 16 November 2009
Received in revised form 18 June 2010
Accepted 16 July 2010
Available online 24 July 2010

Keywords:
Artificial neural networks
Neuro-fuzzy systems
Pellets
Data-mining
Pharmaceutical technology
Computational intelligence

ABSTRACT

The objective was to prepare neural models identifying relationships between formulation characteristics and pellet properties based on algorithmic approach of crucial variables selection and neuro-fuzzy systems application. The database consisted of information about 227 pellet formulations prepared by extrusion/spheronization method, with various model drugs and excipients. Cheminformatic description of excipients and model drugs was employed for numerical description of pellet formulations. Initial numbers of neural model inputs were up to around 3000. The inputs reduction procedure based on sensitivity analysis allowed to obtain less than 40 inputs for each model. The reduced models were subjects of fuzzy logic implementation resulting in logical rules tables providing human-readable rule sets applicable in future development of pellet formulations. Neural modeling enhanced knowledge about pelletization process and provided means for future computer-guided search for the optimal formulation.

1. Introduction

Artificial neural networks (ANNs) are empirical modeling tools, which were proved to be reliable and flexible in many applications in science and technology (Żurada, 1992). They were constructed to mimic biological neural systems in their functional and topological design. The former is the key value of ANNs as these systems are capable of self-learning on the available data. This ability allows to create models without prior assumptions, which is a common pitfall of classical, statistical modeling procedures. Therefore, ANNs belong to the computational intelligence and are often classified as empirical modeling tools. Moreover, ANNs are built from non-linear data processing units (artificial neurons), thus allowing for effective identification of non-linear problems, which is sometimes difficult in statistical approaches. Another distinctive feature of ANNs is their ability to deal effectively with multidimensional problems including several thousands of features and cases as well. Since there were developed tailored procedures of sensitivity analysis (Żurada et al., 1997), ANNs were used effectively as tools for crucial variables selection. Superior abilities of ANNs as empirical modeling tools have attracted attention of scientists in the pharmaceutical field as well. ANNs were used as predictive models in pharmacokinetics and pharmacodynamics (Brier and Żurada, 1995; Brier and Aronoff, 1996; Chow et al., 1997; Hussain et al., 1993; Veng-Pedersen and Modi, 1992, 1993), in vitro–in vivo correlation (IVIVC) development (Dowell et al., 1999), and pharmaceutical technology as well (Hussain et al., 1994). Hussain et al. (1991) were the first to introduce ANNs to the field of pharmaceutical technology, pointing to the possible advantages of guided search for the optimal pharmaceutical formulation. Various dosage forms were subjects of neural analysis: tablets (Bourquin et al., 1998a,b,c; Türkoğlu et al., 1995), pellets (Peh et al., 2000), capsules (Mendyk et al., 2007), emulsions (Gašperlin et al., 1998, 2000) and microemulsions (Agatonovic-Kustrin and Alany, 2001), hydrogels (Takayama et al., 1999, 2003), transdermal delivery systems (Kandimalla et al., 1999), etc. An example of extensive review of ANNs predictive modeling abilities was provided by Bourquin et al. (1998a,b,c), who compared predictive performance of neural models and classical statistical approaches on the extensive database consisting of tablets direct compression studies. The authors concluded that ANNs were superior to statistical approaches especially when the data were not derived from statistical experiment design and contained outliers. Preparation technology was also the subject of studies conducted with ANNs. Inghelbrecht et al. (1997) and Mansa et al. (2008) studied roll compaction procedure with neural models integrating process parameters and granulate properties. ANNs were found to be effective modeling tools providing satisfactory predictive performance in the multidimensional analysis as well

* Corresponding author. Tel.: +48 12 6205 604; fax: +48 12 6205 619.
E-mail address: mfmendyk@cyf-kr.edu.pl (A. Mendyk).

0928-0987/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejps.2010.07.010

as possibility to optimize technology parameters in regard to the desired product characteristics (i.e. particles size). The latter paper (Mansa et al., 2008) introduced neuro-fuzzy systems as predictive and data-mining tools, which provided sets of linguistic rules allowing for easier understanding of analyzed process. The data-mining procedures usage in pharmaceutical product development were studied less extensively than predictive neural modeling. They were usually performed by use of neuro-fuzzy systems and literal presentation of fuzzy rules derived from the data after training of neuro-fuzzy systems. However, another approach involving crucial variables selection was also applied in this field. The crucial variables selection is done by ANN input vector reduction to the minimal number of inputs allowing to represent particular problem with appropriate accuracy. The examples of such approach could be applications of ANNs to the analysis of drug release from hydrodynamically balanced systems (HBS) (Mendyk et al., 2007) or prediction of nanoemulsions droplets sizes (Amani et al., 2008). Additionally, Mendyk and Jachowicz (2007) provided algorithmic solution for crucial variables selection based on the sensitivity analysis, 10-fold cross-validation procedure and knowledge-based search for minimal input vector size.

Pellets have been in the focus of pharmaceutical research for more than 30 years based on their particulate properties. The major advantages of pellets are the homogenous plasma profiles and the reduction of local irritations in the gastrointestinal tract which lead to less side effect for the patient. Pellets are defined as multiparticulate unit dosage forms characterized by spherical shape and a particle size of 0.5–2 mm diameter (Ghebre-Sellassie, 1989). There are several techniques for pellet manufacturing as there are layering and granulation. In this study pellets prepared by extrusion/spheronization technique were used based on their particulate spherical shape and narrow size distribution (Baert et al., 1993). Using this common manufacturing technique the pellet properties are defined by substance properties as well as process specifics (Hellen et al., 1993; Jerwanska et al., 1997). In all previous investigations just a few variables were investigated at the same time, therefore more comprehensive approach was required.

The objective of this work was to prepare neural models identifying relationships between formulation characteristics and pellet properties based on algorithmic approach of crucial variables selection and neuro-fuzzy systems application. Thus, the models would provide the knowledge about impact of manufacture properties as well as formulation composition on the resulting formulation properties.

2. Materials and methods

2.1. The data

The data set consisted of 227 formulations of pellets prepared with eight different model drugs (API) and several excipients (Bornhöft et al., 2005; Thommes and Kleinebudde, 2006, 2007a,b, 2008). The APIs were as follows: acetaminophen, theophylline, mesalamine, hydrochlorothiazide, phenacetin, chloramphenicol, dimenhydrinate and lidocain. The excipients were of polymeric nature like starch, cellulose derivatives and kappa-carrageenans, as well as small molecules like lactose, dicalcium phosphate and mannitol.

2.2. Input and output vector definitions

The input vector included all available information describing pellets quantitative and qualitative composition as well as technology of preparation. Every input vector was divided into the sections describing the above-mentioned classes of information. The number of inputs will be commented in Section 2.3 as it depends on the stage of modeling.

In order to overcome difficulties arising during the preparation of numerical representation of the available data, several approaches were applied regarding encoding techniques and modeling procedures (Section 2.3).

The most difficult to encode numerically was qualitative composition. There were two main approaches implemented:

(a) topological encoding where the position of non-zero value in the input cell was denoting particular drug or excipient presence and
(b) chemical descriptors provided by cheminformatics software.

There was a special procedure of polymer encoding implemented due to lack of cheminformatics software, which could compute descriptors for the whole polymers molecules. It was assumed that each of the polymers could be encoded as the dimer expressing statistical distribution of molecular weights, substitution etc. If the excipient was a mixture of particular chemical entities, the weighted average of available descriptors was finally introduced as the expression of qualitative composition.

The properties of pellets chosen as output variables included mean dissolution time (MDT) and median aspect ratio (AR). For each of mentioned above parameters separate models with one output were created. Thus, the neural model objective was to analyze pellet properties like aspect ratio (AR) and mean dissolution time (MDT) in regard to their qualitative, quantitative composition and preparation technology.

Depending on the analyzed problem different numbers of data records were chosen for the analysis. For AR there were all 227 formulations included, whereas for mean dissolution time MDT there were 89 formulations taken into the consideration due to the measurements results availability.

2.3. Modeling procedure

Two-stage modeling procedure was applied. First stage included analysis of output variable vs. cheminformatic representation of drug substance and topological representation of excipients, whereas second stage was the analysis of the output variable vs. cheminformatic representation of drug substance and carrageenans with topological representation of other excipients. This mixture of continuous and topological encoding of excipients and drugs was enforced by cheminformatics software restrictions mainly due to the presence of inorganic salt (dicalcium phosphate). The detailed discussion of continuous vs. topological encoding for ANNs was given elsewhere (Mendyk and Jachowicz, 2010). The above-mentioned evolution of methodology resulted in the input vector definitions. In the stage 1 for MDT there were initially 1494 input variables whereas for AR there were 1495. In the stage 2 there was added cheminformatic representation of carrageenans, therefore the respective input numbers increased to 3023 and 3024. The discrepancy between input numbers for AR and MDT was a consequence of the fact that not all the formulations were characterized by MDT, thus one of the inputs was unnecessary.

Neural modeling was carried out according to the following algorithm:

1. input vector definition
2. training set of ANNs on the whole data set
3. sensitivity analysis
4. input vector reduction