PREFACE

Cluster analysis has been used in a variety of fields. An algorithm is applied to the data with the purpose of dividing the elements in groups in such a way that objects in the same group are as similar as possible to each other and are as dissimilar as possible to the objects of different groups. Based on the observed values of some variables, the elements are compared using some similarity or dissimilarity measure. This new book explores the fields of clustering algorithms and mathematical modeling with research gathered from around the globe.

Invariably, the design of an effectual and cost-effectual oral drug delivery system (DDS) aims at delivering the drug over extended periods of time to a chronically ill patient. Paralleled with aggressive research in the domain of extended release (ER) oral DDS, its industrial demand has been escalating exponentially with an alarming number of marketed product launches lately. The concept of ER formulation involves drug release form a polymeric material in a precise and engineered manner. This drug release is usually controlled by several physical processes including polymer permeation by dissolution media, diffusion of dissolved drug from the matrix and erosion of the matrix material. Essential to the development of a successful ER DDS, therefore, is the comprehension of drug release kinetics as influenced by several physical, chemical and geometrical parameters of the formulation device. The quantitative interpretation of the drug dissolution profile is facilitated by a set of generic equations which mathematically translate the profile as a function of these multiple parameters. The quintessence of ER drug delivery to yield an impeccable therapeutic effect, accordingly, is modulation and modeling of its drug release kinetic potential. Mathematical modeling, coupled with numerical simulation, provides useful guidance and insight to the pharmaceutical product development scientist on ER, largely reducing the experimentation needed to obtain the desired drug release profile(s). In the last a few decades, a number of mathematical models have been hypothesized by various researchers to analyze drug release kinetics from different types of ER dosage forms, including matrices (insoluble, soluble or biodegradable), microcapsules, coated beads and osmotic pumps. Chapter 1 would embark upon these mathematical approaches to describe drug release kinetics from diverse types of oral DDS. Also, it would address the exigent need for numeric corrections required during a typical dissolution run of an ER product. The chapter will also endeavor to provide an overview on various popular model-independent and model-dependent kinetic approaches to compare the entire drug dissolution profiles with the standard ones along with the requisite algorithms employed.

The importance of understanding and predicting oral drug absorption is well recognised by both research facilities/institutions and the pharmaceutical industry. For this purpose, the use of mathematical models can represent a very profitable and indispensable tool. Indeed, mathematical models can verify the correctness of the mechanisms proposed to describe drug release, absorption, distribution and elimination and thus reduce the number of expensive and time consuming experiments that have to be performed. In this frame, Chapter 2 will deal with the mathematical modeling of the simultaneous drug release from polymeric particles and gastro-intestinal absorption, metabolism and excretion (ADME). Particular care will be devoted to simulate the ADME process in the case of amorphous and nanocrystalline drugs as it is well known that drug solubility increases with decreasing nanocrystals radius (amorphous phase can be thought as a crystal of vanishing dimensions). The model accounts for polymeric particles swelling, particles polydispersity, different initial drug distribution in the particles, solid drug dissolution and re-crystallisation, the simultaneous presence of drug in the amorphous, nano-crystalline and macro-crystalline state, drug absorption and elimination. The gastro-intestinal environment is described as a well-stirred compartment (tank model). All model equations are solved by means of the implicit control volume method. The aim of the model is to establish a theoretical correlation between the drug blood profile concentration after oral administration and the physical properties of the release system.

Drug release modeling and optimization has become an immensely important in the field of pharmaceutical technology, from both academic and industrial points of view. Different drug release modeling approaches are utilized for specific types of pharmaceutical dosage forms. This chapter is a comprehensive review of the current state of the art in mathematical modeling of drug release, and it deals with various types of drug delivery products. Application of mathematical modeling in studying the drug release kinetics allows for elucidation of mechanisms governing the drug release. It also enables accurate prediction of drug release profiles for novel dosage forms, without the need for experimentation. By this assessment, costs of the development of novel dosage forms are reduced and precious time for experimental work is saved.

Both simple (semi)empirical and more complex mechanistic realistic models have their advantages and shortcomings. Phenomenological models and Fickian equations, anomalous transport models and scaling law-based approaches are addressed in Chapter 3. Practical examples of usage of different models are demonstrated. Criteria for selection of appropriate model and its predictive power and precision as well as the easiness of application. Number of parameters to be determined can significantly vary between the models.

In Chapter 4 the authors present an originally developed mathematical model of paracrine regulation of proliferative activity in psoriatic epidermis involving T-lymphocytes (T-helpers and T-suppressors) in the form of a system of three ordinary differential equations describing mitosis, its inhibition, activation and apoptosis of the connected and free cell subpopulations. It is shown that the model will be well coordinated with current world literature data on cell phenotypes and cell kinetics in psoriasis. It shows the existence of ten qualitatively various modes, eight of which describe patterns of formation, long persistent preservation, resolution and also spasmodic transition (trigger) from a pattern of formation to a pattern of resolution of psoriatic lesion skin and vice versa.

One of the steady modes of behavior of the offered model is the self-oscillation of dense epidermal cells observed only in the chronic psoriatic lesions, but not in normal-appearing

uninvolved skin. The period of self-oscillations shown in the lead analysis is about three weeks. For finding out their role in pathogenesis of psoriasis, the authors have created a computing experiment that carries out research of laws of synchronization of self-oscillations of density epidermal cells by periodically changing factors of the environment (by hormones, atmospheric agents, climatic factors, etc.).

The received results have allowed for a number of conclusions about the reasons of dependence of curse of disease from these factors. The hypothesis that the existence of two types of nonpustular psoriasis with early and late onset is caused by two genetically determined steady conditions (oscillatory and stationary) of change of density epidermal cells in psoriatic lesion skin is offered.

Cluster analysis has been used in a variety of fields. Basically, an algorithm is applied to the data with the purpose of dividing the elements in groups in such a way that objects in the same group are as similar as possible to each other (internal cohesion) and are as dissimilar as possible to the objects of different groups (external isolation). Based on the observed values of some variables, the elements are compared using some similarity or dissimilarity measure. Most of the studies published in the literature focus on clustering algorithms for continuous variables. There are many papers proposing new algorithms or comparing the performance of the existing methods. However, in many applications the data consist of categorical and continuous variables and the researcher wants to use both types of variables simultaneously to cluster the data. Less attention has been given in the literature to this situation.

The main purpose of Chapter 5 is to show some of the results of a comparative study. Three algorithms; Average linkage, ROCK and K-Prototypes, which can be used for categorical and continuous variables, were examined.

The study was performed by using Monte Carlo simulation. Data were simulated considering factors such as cluster overlapping, number of groups, variables and categories, and the correlation between continuous variables.

Some of the results show that when the number of groups increase, independent of their structure, the performance of the clustering algorithms decreased. The effect of the increase of the number of variables and categories depends on the internal structure of the clusters. It was also noticed that the correlation between the continuous variables did not cause any effect on the percentage of correct classification and that the clustering methods had better results when more weight was given to the continuous variables than to the categorical on the similarity measure used to cluster the data. In terms of efficiency, ROCK algorithm presented better performance for the artificial data sets generated in the simulation. An example is shown using a real dataset.

For bioequivalence (BE) studies, measurement of the parent drug in single dose is generally recommended. However, for some situations, measurement of an active or inactive metabolite and/or steady state studies might be necessary. These situations are still today a controversial issue with different recommendations from the European Medicines Agency (EMEA) and Food and Drug Administration (FDA) guidance documents (EMEA 2001, EMA 2010, FDA 2002). Modeling and simulation approaches are useful tools for evaluating plasma concentration-time profiles of parent drug and metabolites; its use may be advantageous to assess the potential outcome of different scenarios in bioequivalence studies. Actually, the current recommendations on regulatory guidance documents come, from some simulation studies performed and published in the past. Chapter 6 will review some of the previous simulation studies of bioequivalence trials as well as their assumptions and the limitations of

their outcomes. The extent of pre-systemic metabolism and the non-linearity of the metabolic processes are controversial issues that require harmonization with regards to analyte selection. The lack of agreement in FDA and EMEA recommendations makes evident that the simulations on which those recommendations were based were performed in a different set of scenarios under a different set of assumptions leading to different answers to the same question. A new approach will be presented based on a semi-physiological pharmacokinetic model with a more systematic approach, based on BCS (Biopharmaceutic Classification System) to explore the relevance of the intrinsic clearance magnitude (high or low) and the intra-subject variability (high or low) on the bioequivalence outcome of Class I to IV drugs under linear and non-linear conditions. The ultimate goal is to address if the existence of pre-systemic (first-pass effect) gut wall and hepatic metabolism are reasons to require the measurement of the active metabolite, as requested apparently by the FDA. The second objective is to assess the differences in the outcome of the BE trials in both single dose and steady state study designs, in particular when the system is not linear. The final aim is to show how a modeling and simulation exercise could be useful in a regulatory environment and what are the consequences and limitations of these virtual trials.

Finding clusters in data is a challenging problem especially when the clusters are being of widely varied shapes, sizes, and densities. Herein a new scalable clustering technique which addresses all these issues is proposed. In data mining, the purpose of data clustering is to identify useful patterns in the underlying dataset. Within the last several years, many clustering algorithms have been proposed in this area of research. Among all these proposed methods, density clustering methods are the most important due to their high ability to detect arbitrary shaped clusters. Moreover these methods often show good noise-handling capabilities, where clusters are defined as regions of typical densities separated by low or no density regions. In Chapter 7, the authors aim at enhancing the well-known algorithm DBSCAN, to make it scalable and able to discover clusters from uneven datasets in which clusters are regions of homogenous densities. The authors achieved the scalability of the proposed algorithm by using the k-means algorithm to get initial partition of the dataset, applying the enhanced DBSCAN on each partition, and then using a merging process to get the actual natural number of clusters in the underlying dataset. This means the proposed algorithm consists of three stages. Experimental results using synthetic datasets show that the proposed clustering algorithm is faster and more scalable than the enhanced DBSCAN counterpart.

In Chapter 8, the authors investigate the feature weight selection based on the Pearson's coefficient of variation. They also discuss the effect of the feature weight for fuzzy c-means algorithm.

In Chapter 9, the authors consider some mathematical models of hydriding of metal powders. Their aim is to choose the most significant factors in order to construct adequate but rather simple models. We describe the process of non-metal hydride formation for a single particle and focus on the case of hydriding with formation of a hydride skin on the surface of a particle. Mathematical models are constructed for this case. The authors present the results of fitting the series of experimental curves in order to estimate kinetic parameters and study the influence of the powder particles' shape. It is shown that different shapes provide similar results and thus the shape is insignificant. Then they add some factors, previously not taken into account: changing pressure, heat releasing effects accompanied by changing temperature, different heat capacities of metal and hydride phases and hydrogen. Significant release of heat

during hydriding is an important problem when constructing hydride hydrogen storage systems. The model contains a non-classical boundary-value problem; the authors describe how to transform it so that it could be easily solved numerically. Finally, they describe an approach to modeling a real powder: in the case of quick sorption we have the heat-controlling hydriding with concentration redistribution between the particles. We construct the discrete 1-D, 2-D, and 3-D models that are convenient for parallel calculating.

Adapting behaviour to changes in the environment, as well as to memory of previous history, is typical for human communities and animal populations. One obvious example of such adaptation of behaviour is switching to a "safe mode of behaviour" when a danger, either from a predator species, or from an infectious disease, becomes apparent. In human society such switching to safe behaviour is particularly evident and often occurs in the case of epidemics.

Mathematically, such changes of behaviour in response to changes in conditions, can be described by models with switching. In most cases the switching is assumed to occur in a manner that depends on the system state, and, thus, the switching disregards history and memory. Memory can be introduced into a mathematical model using a phenomenon that is known as "hysteresis". In Chapter 10 we illustrate this idea, using a simple population model that arises in epidemiology as a case study. Our goal is to show why and how hysteresis can arise in this type of model, and how it may be applied to describe "memory effects". An ancillary objective is to introduce a new unified paradigm for mathematical modeling with memory effects in epidemiology and ecology.