Interdisciplinary Applied Mathematics 30

Panos Macheras
Athanassios Iliadis

Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics

Homogeneous and Heterogeneous Approaches

Second Edition

Springer

Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics

Interdisciplinary Applied Mathematics

Volume 30

Problems in engineering, computational science, and the physical and biological sciences are using increasingly sophisticated mathematical techniques. Thus, the bridge between the mathematical sciences and other disciplines is heavily traveled. The correspondingly increased dialog between the disciplines has led to the establishment of the series: *Interdisciplinary Applied Mathematics*.

The purpose of this series is to meet the current and future needs for the interaction between various science and technology areas on the one hand and mathematics on the other. This is done, firstly, by encouraging the ways that mathematics may be applied in traditional areas, as well as point towards new and innovative areas of applications; and, secondly, by encouraging other scientific disciplines to engage in a dialog with mathematicians outlining their problems to both access new methods and suggest innovative developments within mathematics itself.

The series will consist of monographs and high-level texts from researchers working on the interplay between mathematics and other fields of science and technology.

More information about this series at http://www.springer.com/series/1390

Panos Macheras • Athanassios Iliadis

Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics

Homogeneous and Heterogeneous Approaches

Second Edition

Springer

Panos Macheras
Faculty of Pharmacy
National and Kapodistrian University
of Athens
Athens, Greece

Athanassios Iliadis
Faculty of Pharmacy
Aix-Marseille University
Marseille, France

ISSN 0939-6047 ISSN 2196-9973 (electronic)
Interdisciplinary Applied Mathematics
ISBN 978-3-319-27596-3 ISBN 978-3-319-27598-7 (eBook)
DOI 10.1007/978-3-319-27598-7

Library of Congress Control Number: 2015957958

Mathematics Subject Classification (2010): 92C45, 62P10, 74H65, 60K20

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

- To our ancestors who inspired us
- To those of our teachers who guided us
- To our families

Panos Macheras and Athanassios Iliadis

Preface to the First Edition

Η μεγάλη τέχνη βρίσκεται οπουδήποτε ο άνθρωπος κατορθώνει ν'αναγνωρίζει τον εαυτόν του και να τον εκφράζει με πληρότητα μες στο ελάχιστο.
Great art is found wherever man achieves an understanding of self and is able to express himself fully in the simplest manner.

Odysseas Elytis (1911–1996)
1979 Nobel Laureate in Literature
The Magic of Papadiamantis

Biopharmaceutics, pharmacokinetics, and pharmacodynamics are the most important parts of pharmaceutical sciences because they bridge the gap between the basic sciences and the clinical application of drugs. The modeling approaches in all three disciplines attempt to:

- Describe the functional relationships among the variables of the system under study.
- Provide adequate information for the underlying mechanisms.

Due to the complexity of the biopharmaceutic, pharmacokinetic, and pharmacodynamic phenomena, novel physically physiologically based modeling approaches are sought. In this context, it has been more than ten years since we started contemplating the proper answer to the following complexity-relevant questions: Is a solid drug particle an ideal sphere? Is drug diffusion in a well-stirred dissolution medium similar to its diffusion in the gastrointestinal fluids? Why should peripheral compartments, each with homogeneous concentrations, be considered in a pharmacokinetic model? Can the complexity of arterial and venular trees be described quantitatively? Why is the pulsatility of hormone plasma levels ignored in pharmacokinetic–dynamic models? Over time we realized that questions of this kind can be properly answered only with an intuition about the underlying heterogeneity of the phenomena and the dynamics of the processes. Accordingly, we borrowed geometric, diffusional, and dynamic concepts and tools from physics and mathematics and applied them to the analysis of complex biopharmaceutic, pharmacokinetic, and pharmacodynamic phenomena. Thus, this book grew out of

our conversations with fellow colleagues, correspondence, and joint publications. It is intended to introduce the concepts of fractals, anomalous diffusion, and the associated nonclassical kinetics and stochastic modeling, within nonlinear dynamics, and illuminate with their use of the intrinsic complexity of drug processes in homogeneous and heterogeneous media. In parallel fashion, we also cover in this book all classical models that have direct relevance and application to the biopharmaceutics, pharmacokinetics, and pharmacodynamics.

The book is divided into four sections, with Part I, Chapters 1–3, presenting the basic new concepts: fractals, nonclassical diffusion-kinetics, and nonlinear dynamics; Part II, Chapters 4–6, presenting the classical and nonclassical models used in drug dissolution, release, and absorption; Part III, Chapters 7–11, presenting empirical, compartmental, and stochastic pharmacokinetic models; and Part IV, Chapters 12 and 13, presenting classical and nonclassical pharmacodynamic models. The level of mathematics required for understanding each chapter varies. Chapters 1 and 2 require undergraduate-level algebra and calculus. Chapters 3–8, 12, and 13 require knowledge of upper undergraduate- to graduate-level linear analysis, calculus, differential equations, and statistics. Chapter 11 requires knowledge of probability theory.

We would like now to provide some explanations in regard to the use of some terms written in italics below, which are used extensively in this book starting with *homogeneous* vs. *heterogeneous* processes. The former term refers to kinetic processes taking place in well-stirred, Euclidean media where the classical laws of diffusion and kinetics apply. The term *heterogeneous* is used for processes taking place in disordered media or under topological constraints where classical diffusion-kinetic laws are not applicable. The word *nonlinear* is associated with either the kinetic or the dynamic aspects of the phenomena. When the kinetic features of the processes are nonlinear, we basically refer to Michaelis–Menten-type kinetics. When the dynamic features of the phenomena are studied, we refer to nonlinear dynamics as delineated in Chapter 3.

A *process* is a real entity evolving, in relation to time, in a given environment under the influence of internal mechanisms and external stimuli. A *model* is an image or abstraction of reality: a mental, physical, or mathematical representation or description of an actual process, suitable for a certain purpose. The model need not be a true and accurate description of the process, nor need the user have to believe so, in order to serve its purpose. Herein, only mathematical models are used. Either processes or models can be conceived as boxes receiving inputs and producing outputs. The boxes may be characterized as gray or black, when the internal mechanisms and parameters are associated or not with a physical interpretation, respectively. The *system* is a complex entity formed of many, often diverse, interrelated elements serving a common goal. All these elements are considered as *dynamic processes* and *models*. Here, deterministic, random, or chaotic real processes and the mathematical models describing them will be referenced as *systems*. Whenever the word "system" has a specific meaning like *process* or *model*, it will be addressed as such.

For certain processes, it is appropriate to describe globally their properties using numerical techniques that extract the basic information from measured data. In

the domain of linear processes, such techniques are correlation analysis, spectral analysis, etc. and in the domain of nonlinear processes, the correlation dimension, the Lyapunov exponent, etc. These techniques are usually called *nonparametric models* or, simply, indices. For more advanced applications, it may be necessary to use models that describe the functional relationships among the system variables in terms of mathematical expressions like difference or differential equations. These models assume a prespecified parameterized structure. Such models are called *parametric models*.

Usually, a mathematical model simulates a process behavior, in what can be termed a *forward problem*. The *inverse problem* is, given the experimental measurements of behavior, what is the structure? A difficult problem, but an important one for the sciences. The inverse problem may be partitioned into the following stages: hypothesis formulation, i.e., model specification, definition of the experiments, identifiability, parameter estimation, experiment, and analysis and model checking. Typically, from measured data, nonparametric indices are evaluated in order to reveal the basic features and mechanisms of the underlying processes. Then, based on this information, several structures are assayed for candidate parametric models. Nevertheless, in this book we look only into various aspects of the forward problem: given the structure and the parameter values, how does the system behave?

Here, the use of the term "model" follows Kac's remark, "models are caricatures of reality, but if they are good they portray some of the features of the real world" [1]. As caricatures, models may acquire different forms to describe the same process. Also, Fourier remarked, "nature is indifferent toward the difficulties it causes a mathematician"; in other words the mathematics should be dictated by the biology and not vice versa. For choosing among such competing models, the "parsimony rule," Occam's "razor rule," or Mach's "economy of thought" may be the determining criteria. Moreover, modeling should be dependent on the purposes of its use. So, for the same process, one may develop models for process identification, simulation, control, etc. In this vein, the tourist map of Athens and the system controlling the urban traffic in Marseille are both tools associated with the real life in these cities. The first is an identification model and the second, a control model.

Over the years we have benefited enormously from discussions and collaborations with students and colleagues. In particular we thank P. Argyrakis, D. Barbolosi, A. Dokoumetzidis, A. Kalampokis, V. Karalis, K. Kosmidis, C. Meille, E. Rinaki, and G. Valsami. We wish to thank J. Lukas whose suggestions and criticisms greatly improved the manuscript.

Piraeus, Greece P. Macheras
Marseille, France A. Iliadis
August 2005

Preface to the Second Edition

The objectives and scope of this book remain the same as in the first edition: to present the homogeneous and heterogeneous approaches used in the modeling work of biopharmaceutics, pharmacokinetics, and pharmacodynamics. However, in the ensuing ten years from the first publication of the book, significant changes took place in the modeling and simulation work in the field of drug development and research. The prominent drug agencies US FDA and EMA have taken specific steps toward the utility of modeling and simulation in drug development and registration of medicinal products. Moreover, the latest EMA guideline on investigation of bioequivalence introduced several new concepts including alternative possibilities for the clinical design, the statistical analysis, the moiety to be analyzed, and the application of classification of drugs using the biopharmaceutic classification system (BCS) as well as the biopharmaceutic drug disposition classification system (BDDCS). As a result of all these developments, this edition contains two new chapters "Fractional Pharmacokinetics" (Chapter 9) and "Modeling and Simulation in Bioequivalence" (Chapter 10) written by Dr. Aristides Dokoumetzidis and Dr. Vangelis Karalis, respectively. Chapter 9 presents the mathematical formalism based on fractional calculus for the analysis of pharmacokinetics of drugs following anomalous kinetics. Chapter 10 presents an overview of the modeling and simulation methods, which are applied to many areas of bioequivalence assessment. Besides, new material has been added in Chapters 5 and 6 based on recent developments in reaction-limited dissolution models and supersaturated dissolution data as well as the recent advances related to BCS and BDDCS. Also, new sections regarding the time-varying models and analysis of nonlinear mechanisms were also added to Chapters 7 and 13, respectively. Almost all models and procedures were implemented within MATLAB, the commonly used software for numerical analysis purposes.

The first edition of the book was awarded a prize in "Sciences" of the Academy of Athens in 2007. The authors wish to thank Athanassios S. Fokas, Academician and Professor in Nonlinear Mathematical Science at the University of Cambridge, UK, for his support in proposing the book for this award.

Finally, we thank the readers of the first edition for the excellent feedback and stimulus to produce an updated version. We would also like to thank S. Benay, R. Bies, N. Frances, and N. Pispa for their helpful discussions and suggestions.

Piraeus, Greece P. Macheras
Marseille, France A. Iliadis
October 2015

Contents

List of Figures

Part I
Basic Concepts

Biopharmaceutics, pharmacokinetics, and pharmacodynamics have been developed and expanded, as have many other scientific subjects, based on the concepts of *homogeneity* and *linearity*.

Homogeneity is a presupposition in almost all research in these fields. Some typical examples can be quoted. The shape of a drug particle is assumed to be an ideal sphere and its surface smooth. The permeability of the intestinal complex membrane is considered constant along the gastrointestinal tract while the concentration of drug is postulated to be homogeneous in the gastrointestinal fluids. Homogeneous conditions are assumed for each one of the hypothetical compartments of pharmacokinetic models and the effect compartment of pharmacokinetic-dynamic models. However, common intuition and scientific knowledge tell us that the drug particle is not an ideal sphere, and its surface is not smooth: the permeability of the gastrointestinal wall is position- and time-dependent. Moreover, the assumed concentration homogeneity of drug in (1°) the gastrointestinal tract, (2°) the peripheral compartments of compartmental systems, and (3°) the biophase at receptor's level, is synonymous with a well-mixed system.

Linearity is the basic assumption behind classical biopharmaceutics, pharmacokinetics, and pharmacodynamics. Formally, a system is linear if the output of an operation is proportional to the input. This property of proportionality along with the property of independence, i.e., the response of the system to an input is equal to the sum of the outcomes, is the fundamental feature of linear systems. In contrast, nonlinear kinetic phenomena are frequently observed in pharmacokinetic and pharmacodynamic studies, when the output (e.g., the area under the time–concentration curve) is nonlinearly related to the input (e.g., the dose). However, we deal with complicated biological systems consisting of a large number of interrelated components and processes. Linear approaches neglect these intrinsic relations and therefore are not complete.

In various fields of research, e.g., physics, chemistry, and physiology, scientists are increasingly finding that at the research level it is the nonlinear phenomena that control the game; physical or physiological heterogeneity is everywhere, while heterogeneous conditions prevail in numerous physical, physiological, and biochemical

Table 1 Classical and nonclassical considerations of the in vitro and in vivo drug processes.

Fields	Classical	Nonclassical
Geometry	Euclidean	Fractal
Topology	Ordered media	Disordered media
Diffusion	Regular	Anomalous
Kinetics	Deterministic	Stochastic
Dynamics	Linear	Nonlinear

processes. Today, science shows that the real world is relentlessly nonlinear, and therefore the techniques of nonlinear dynamics are required to analyze the nonlinear phenomena. In parallel, structural and functional heterogeneities can be described and understood with the concept of fractals.

Kinetic processes in various scientific fields are traditionally treated with classical kinetics. The latter is quite satisfactory for reactions and processes in well-stirred media, i.e., under homogeneous conditions. In fact, the kinetics of diffusion-controlled processes and reactions in three-dimensional homogeneous systems obey the classical laws of diffusion where the rate constant of the process is linearly proportional to the diffusion coefficient. However, this proportionality is not valid for systems with smaller dimensions, fractal spaces, or disordered systems since the laws of transport are different in these media. Accordingly, fractal kinetics has been developed since classical kinetics was found to be unsatisfactory under dimensional constraints, e.g., phase boundaries, understirred media, or membrane reactions.

The concepts delineated above are quoted in Table 1 in a comparative manner. The clear message of the middle column of Table 1 is that classical behavior is expected when the phenomena are taking place in Euclidean spaces and/or ordered media (e.g., well-stirred systems) since the classical laws of diffusion and kinetics are valid. According to the third column of Table 1, deviations from the normal behavior appear when the processes take place in fractal spaces and/or under topological constraints since neither diffusion nor kinetics follow the classical pattern in that case. Finally, the dynamics of the systems are linear when the variable of interest is considered to be detached from the remainder of the system. On the contrary, nonlinear dynamic behavior exists when the interaction of the studied variable with the other variables constituting the system is taken into account.

We do believe that it is only through the understanding of underlying principles that adequate exploration in the fields of biopharmaceutics, pharmacokinetics, and pharmacodynamics can be carried out. To this end, in this first, introductory part of the book we deal with the basic nonclassical concepts. Chapter 1 gives the essentials of fractals such as self-similarity, scaling laws, and fractal dimension. Chapter 2 deals with the basics of diffusion-kinetics in either Euclidean or disordered media. The concepts of nonlinear dynamics are discussed in Chapter 3, where reference is made to a number of characteristics of dynamic systems, e.g., phase space, attractor, bifurcation.

Chapter 1
The Geometry of Nature

The proper route to an understanding of the world is an examination of our errors about it.

Euclid (325-265 BC)

Our understanding of nature has been based on the classical geometric figures of smooth line, triangle, circle, cube, sphere, etc. Each of these regular forms can be determined by a characteristic scale. For example, the length of a straight line can be measured with a ruler that has a finer resolution than the entire length of the line. In general, each Euclidean object has a unique value for its characteristics (length, area, or volume). It is also known that when these objects are viewed at higher magnification they do not reveal any new features.

In the real world, however, the objects we see in nature and the traditional geometric shapes do not bear much resemblance to one another. Mandelbrot [2] was the first to model this irregularity mathematically: clouds are not spheres, mountains are not cones, coastlines are not circles, and bark is not smooth, nor does lightning travel in a straight line. Mandelbrot coined the word *fractal* for structures in space and processes in time that cannot be characterized by a single spatial or temporal scale. In fact, the fractal objects and processes in time have multiscale properties, i.e., they continue to exhibit detailed structure over a large range of scales. Consequently, the value of a property of a fractal object or process depends on the spatial or temporal characteristic scale measurement (ruler size) used.

The physiological implications of the fractal concepts are serious since fractal structures and processes are ubiquitous in living things, e.g., the lung, the vascular system, neural networks, the convoluted surface of the brain, ion channel kinetics, and the distribution of blood flow through the blood vessels. Besides, many applications of fractals exist for the morphology of surfaces, e.g., the surface area of a drug particle, surface reactions on proteins. Thus, fractal geometry allows scientists to formulate alternative hypotheses for experimental observations, which lead to more realistic explanations compared to the traditional approaches. These fractal hypotheses can be expressed in terms of quantifying the fractal properties of the system under study as delineated below.

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_1

1.1 Geometric and Statistical Self-Similarity

The most interesting property of fractals is geometric self-similarity, which means that the parts of a fractal object are smaller exact copies of the whole object. Replacement algorithms are used to generate geometric fractals. For example, the Koch curve shown in Figure 1.1 can be produced after four successive replacements according to the following replacement rule: two lines of the same length replace the middle third of the length of the line at each step. Infinite recursions can be applied resulting in a continuous increase of the "line" length by a factor of $4/3$ at each successive step. This continuous ramification of the Koch curve leads to a surprising result if one attempts to measure the length of its perimeter: the length is dependent on the ruler size used for its measurement. In fact, the smaller the ruler size used, the longer the perimeter. Accordingly, when we deal with fractal objects or processes we say that their characteristics (length in this case) "scale" with the measurement resolution.

Similar algorithms for area and volume replacement can be used to create fractals from two- or three-dimensional objects. The fractals shown in Figure 1.2 are called the Sierpinski triangle (gasket) and Menger sponge. They have been generated from an equilateral triangle and a cube, respectively, by applying the following replacement algorithms:

- Sierpinski triangle: At each step an equilateral triangle with area equal to one-quarter of the remaining triangle is removed.
- Menger sponge: At each step one-third of the length of the side of each cube is removed taking care to apply this rule in 3 dimensions and avoiding removal

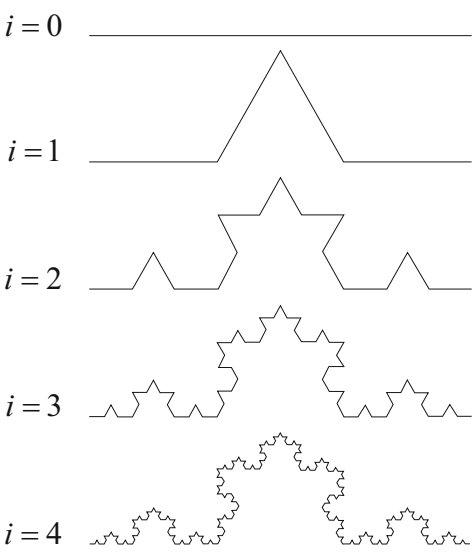


Fig. 1.1 The first four iterations of the Koch curve

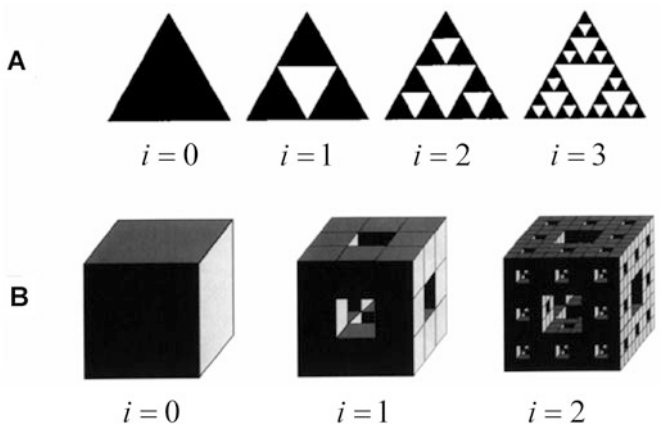


Fig. 1.2 Generation of the (**A**) Sierpinski triangle (gasket) (the first three iterations are shown), (**B**) Menger sponge (the first two iterations are shown) from their Euclidean counterparts

of corner cubes. This means that if the original cube has been constructed from $3 \times 3 \times 3 = 27$ small cubes, after the first iteration 20 small cubes are remaining (6 are removed from the center of the faces and one is removed from the center of the cube).

These line, area, and volume replacement rules give fractal structures (Figures 1.1 and 1.2), which are quite different from the original Euclidean objects. This obvious difference in shape has implications when one considers physical measurements or (bio)chemical processes taking place in Euclidean vs. fractal spaces. For example, surface and/or surface/volume ratios are extremely important for reactions or transport processes taking place at interfaces of different phases like liquid–solid boundaries, e.g., drug dissolution, drug uptake from the gastrointestinal mucosa. In general, objects with fractal surfaces are very efficient for surface reactions.

Replacement rules are expressed mathematically by difference equations, which can be used to generate the fractal structures. These equations are usually called maps and have the form

$$z_{i+1} = g\left(z_i\right), \tag{1.1}$$

where z_i and z_{i+1} are the input and output, respectively, at two successive steps, while the functional form of g in (1.1) depends on the exact features of the recursion process. The discrete nature of (1.1) allows for a recursive creation of the fractal object utilizing the output z_{i+1} as the next input z_i. In this respect, (1.1) operates like a copy machine, which produces the self-similar object in accord with the rule imposed on g.

The replacement rules used for the generation of fractal objects ensure the geometric self-similarity discussed above. However, the fractal objects or processes we encounter in nature are not generated by exact mathematical rules. For example, some biological objects with fractal structure like the venular and arterial tree cannot be characterized by geometric self-similarity; rather they possess statistical self-similarity. The fractal is statistically self-similar since the characteristics (such as the average value or the variance or higher moments) of the statistical distribution for each small piece are proportional to the characteristics that concern the whole object. For example, the average rate at which new vessels branch off from their parent vessels in a physiological structure can be the same for large and small vessels. This is due to the fact that portions of fractal biological objects resemble the whole object instead of being exact copies of the whole. The term *random fractal* is used for these fractal structures to underline their statistical character. Also, statistical self-similarity can be observed when time series data are recorded for physiological processes, e.g., the electroencephalogram or the electrocardiogram. In this case, we speak of statistical self-similarity in time and not in space.

At this point, a distinction should be made between geometrically and statistically self-similar fractals. The pure mathematical basis of geometric fractals does not impose any restriction on the range of application of their scaling laws. In contrast, scaling laws for statistically self-similar fractals adhering to biological objects or processes are subject to the limitations imposed by the physiology and/or the resolution of the measurement technique. In other words, experimental data usually obey scaling laws over a finite range of resolution measurements. This important aspect of scaling laws, with regard to the range of their application, should be carefully considered when one is applying scaling principles for the analysis of experimental data.

1.2 Scaling

The issue of scaling was touched upon briefly in the previous section. Here, the quantitative features of scaling expressed as scaling laws for fractal objects or processes are discussed. Self-similarity has an important effect on the characteristics of fractal objects measured either on a part of the object or on the entire object. Thus, if one measures the value of a characteristic $\theta(\omega)$ on the entire object at resolution ω, the corresponding value measured on a piece of the object at finer resolution $\theta(r\omega)$ with $r < 1$ will be proportional to $\theta(\omega)$

$$\theta(r\omega) = k\theta(\omega), \tag{1.2}$$

where k is a proportionality constant that may depend on r. When statistical self-similarity in time for recordings of an observable is examined, the scale $r\omega$ is a finer time resolution than scale ω. Relation (1.2) reveals that there is a constant ratio k

between the characteristic $\theta(\omega)$ measured at scale ω and the same characteristic $\theta(r\omega)$ measured at scale $r\omega$.

The above-delineated dependence of the values of the measurements on the resolution applied suggests that there is no true value of a measured characteristic. Instead, a scaling relationship exists between the values measured and the corresponding resolutions utilized, which mathematically may have the form of a scaling power law

$$\theta(\omega) = \beta\omega^{\alpha}, \tag{1.3}$$

where β and a are constants for the given fractal object or process studied. Equation (1.3) can be written as

$$\ln\theta(\omega) = \ln\beta + \alpha\ln\omega.$$

This equation reveals that when measurements for fractal objects or processes are carried out at various resolutions, the log–log plot of the measured characteristic $\theta(\omega)$ against the scale ω is linear. Such simple power laws, which abound in nature, are in fact self-similar: if ω is rescaled (multiplied by a constant), then $\theta(\omega)$ is still proportional to ω^{a}, albeit with a different constant of proportionality. As we will see in the rest of this book, power laws, with integer or fractional exponents, are one of the most abundant sources of self-similarity characterizing heterogeneous media or behaviors.

1.3 Fractal Dimension

The objects considered are sets of points embedded in a Euclidean space. The dimension of the Euclidean space that contains the object under study is called the *embedding dimension*, d_e, e.g., the embedding dimension of the plane is $d_e = 2$ and of three-dimensional space is $d_e = 3$.

One is accustomed to associating topological dimensions with special objects: dimension 1 with a curve, dimension 2 with a square, and dimension 3 with a cube. Because there are severe difficulties for the definition of the topological dimension d_t, it is convenient to associate the topological dimension of an object with its *cover dimension* d_o.

A curve in the plane is covered with three different arrangements of disks (Figure 1.3 center). In the right part of the figure there are only pairs of disks with nonempty intersections, while in the center part there are triplets and in the left part even quadruplets. Thus, one can arrange coverings of the curve by only one intersection of each disk with another, and the cover dimension of a line is defined as $d_o = d_t = 1$.

A set of points (Figure 1.3 top) can be covered with disks of sufficiently small radius so that there is no intersection between them. Their covering dimension is

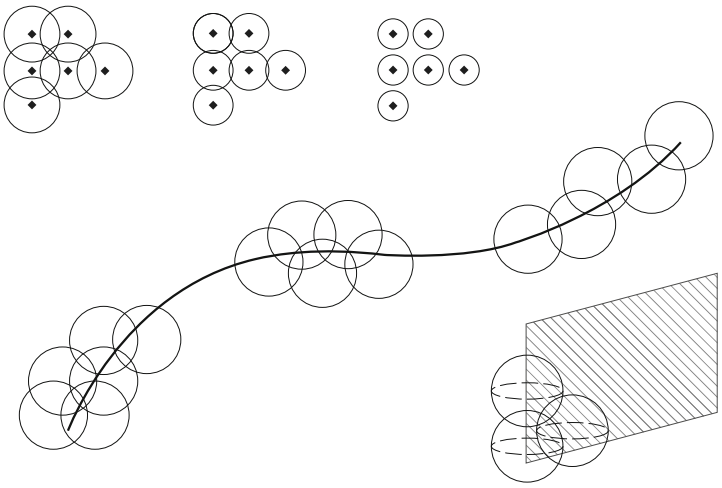


Fig. 1.3 The cover dimension

$d_o = d_t = 0$. A surface (Figure 1.3 bottom) has covering dimension $d_o = d_t = 2$, because one needs at least two overlapping spheres to cover the surface. The same ideas generalize to higher dimensions.

Similarly, the degree of irregularity of a fractal object is quantified with the *fractal dimension*, d_f. This term is used to show that apart from the Euclidean integer dimensions (1 or 2 or 3) for the usual geometric forms, fractal objects have noninteger dimensions. The calculation of d_f using the concept of self-similarity requires in essence the knowledge of the replacement rule, which dictates how many similar pieces m are found when the scale is reduced by a given factor r at each step. Thus, if we count the number m of the exact copies of the entire geometric fractal that are observed when the resolution of scale is changed by a factor of r, the value of d_f can be derived from

$$d_f = \frac{\ln m}{\ln r} \tag{1.4}$$

after logarithmic transformation of

$$m = r^{d_f}. \tag{1.5}$$

For example, the fractal dimension of the Koch curve is 1.2619 since four ($m = 4$) identical objects are observed (cf. levels $i = 0$ and $i = 1$ in Figure 1.1) when the length scale is reduced by a factor $r = 3$, i.e., $d_f = \ln 4 / \ln 3 \approx 1.2619$. What does this noninteger value mean? The Koch curve is neither a line nor an area since its (fractal) dimension lies between the Euclidean dimensions, 1 for lines and 2 for areas. Due to the extremely ramified structure of the Koch curve, it covers a portion

of a two-dimensional plane and not all of it and therefore its "dimension" is higher than 1 but smaller than 2.

Similarly, the first iteration in the generation of the Sierpinski gasket (Figure 1.2 A) involves the reduction of the scale by a factor $r = 2$ and results in 3 identical black equilateral triangles ($m = 3$); thus, $d_f = \ln 3/\ln 2 \approx 1.5815$. For the Menger sponge (Figure 1.2B), the reduction of the scale by a factor $r = 3$ results in $m = 20$ identical cubes, i.e., $d_f = \ln 20/\ln 3 \approx 2.727$. Both values of d_f are consistent with their dimensions since the Sierpinski gasket lies between 1 and 2, while the Menger sponge lies between 2 and 3.

Equations (1.4) and (1.5) are also valid for Euclidean objects. For example, if one creates $m = 16$ identical small squares in a large square by reducing the length scale by one-fourth, $r = 4$, the value of d_f is $\ln 16/\ln 4 = 2$, which is the anticipated result, i.e., the topological dimension $d_t = 2$ for a plane.

1.4 Estimation of Fractal Dimension

Irrespective of the origin of fractals or fractal-like behavior in experimental studies, the investigator has to derive an estimate for d_f from the data. Since strict self-similarity principles cannot be applied to experimental data extracted from irregularly shaped objects, the estimation of d_f is accomplished with methods that unveil either the underlying replacement rule using self-similarity principles or the power-law scaling. Both approaches give identical results and they will be described briefly.

1.4.1 Self-Similarity Considerations

In principle, the object under study is covered with circles for one- and two-dimensional objects or spheres for three-dimensional objects. This process is repeated using various sizes ω for circles or spheres, while overlapping may be observed. Then, the minimum number of "balls" (circles or spheres) $m(\omega)$ of size ω needed to cover the object is calculated. Finally, the fractal dimension, which in this case is called the *capacity dimension*, d_c is calculated from the relationship

$$d_c = \lim_{\omega \to 0} \frac{\ln m(\omega)}{\ln (1/\omega)}. \tag{1.6}$$

Note that (1.6) relies on the self-similarity concept since the number of identical objects m and the scale factor r in (1.5) have been replaced by the number of "balls" $m(\omega)$ and the reciprocal of the size $1/\omega$, respectively. The limit ($\omega \to 0$) is being used to indicate the estimation of d_c at the highest possible resolution, i.e., as the "ball" size ω decreases continuously.

The reference situation implied in this definition is that at $\omega = 1$, one "ball" covers the object. A clearer definition of d_c is

$$d_c = \frac{\ln\left[m\left(\omega\right)/m\left(1\right)\right]}{\ln\left(1/\omega\right)},$$

or in general, if at $\omega = 1$, k "balls" cover the object,

$$d_c = \frac{\ln\left[m\left(k\omega\right)/m\left(k\right)\right]}{\ln\left(k/k\omega\right)}$$

and

$$d_c = -\frac{\mathrm{d}\ln\left[m\left(\omega\right)\right]}{\mathrm{d}\ln\omega}. \tag{1.7}$$

The capacity dimension tells us how the number of "balls" changes as the size of the "balls" is decreased. This method is usually called *box counting* since the method is implemented in computers with a variety of algorithms utilizing rectangular grids instead of "balls." Dimensions d_f and d_c are quite similar, and the differences between them are usually much smaller than the error of estimates [3].

1.4.2 Power-Law Scaling

When the scaling law (1.3) of the measured characteristic θ can be derived from the experimental data (ω, θ), an estimate of the fractal dimension d_f of the object or process can be obtained as well. In order to apply this method one has first to derive the relationship between the measured characteristic θ and the function of the dimension $g(d_f)$, which satisfies

$$\theta \propto \omega^{g(d_f)}, \tag{1.8}$$

where ω represents the various resolutions used. Then, the exponents of (1.3) and (1.8) are equated,

$$g(d_f) = \alpha, \tag{1.9}$$

and (1.9) is solved in terms of d_f to derive an estimate for d_f.

The form of the function $g(d_f)$ in (1.9) depends on the measured characteristic θ [4]. For instance:

- When the characteristic is the mass of the fractal object, the exponent of (1.8) corresponds to the value of d_f, $d_f = \alpha$.

- When the characteristic is the average density of a fractal object, $d_f = d_e + \alpha$, where d_e is the embedding dimension.
- For measurements regarding lengths, areas, or volumes of objects, a simple equation can be derived using scaling arguments, $d_f = d_e - \alpha$.

Apart from the estimation of d_f from experimental data for mass, density, and purely geometric characteristics, the calculation of d_f for a plethora of studies dealing with various characteristics like frequency, electrical conductivity, and intensity of light is also based on the exact relationship that is applicable in each case between d_f and the scaling exponent α (1.9).

1.5 Self-Affine Fractals

The replacement rule we have used so far to generate geometric fractals creates isotropic fractals. In other words, the property of geometric self-similarity is the same in all directions. Thus, a unique value for the fractal dimension d_f is being used to quantify an irregular structure. When either the replacement algorithm or the actual physical object exhibits an asymmetry in different directions, then the anisotropic fractal is characterized as a *self-affine fractal*. For example, if one divides a large square into 6 identical small parallelograms and discards 3 of them in an alternate series at each iteration, the result is a disconnected self-affine fractal. Obviously, the unequal horizontal and vertical sides of the parallelograms produced with the successive replacements follow different scaling laws in accord with the dimensions of the sides. The basic difference between self-similarity and self-affinity lies in the fact that self-similar fractals become identical upon simple magnification (classical scaling), while to become identical, self-affine fractals should be scaled by different amounts of the spatial directions. Accordingly, there is no single value of d_f for self-affine fractals; it varies with the ruler size used for measurements. Usually, the box-counting method is applied in conjunction with (1.6) with limits $\omega \to 0$ and $\omega \to \infty$; two estimates for d_f are derived, namely, $d_{f,local}$ and $d_{f,global}$, respectively, and used to characterize a self-affine fractal. Both values indicate limiting values of the fractal dimension: the former is relevant when the size of the boxes decreases infinitely, while the latter corresponds to the largest length scale used for measurements.

1.6 More About Dimensionality

The concept of fractals has helped us to enrich the notion of dimensionality. Apart from the classical systems with dimensions 1, 2, and 3 there are disordered systems with noninteger dimensions.

In the simplest case, a system is called Euclidean or nonfractal if its topological dimension d_t is identical to the fractal dimension d_f. This means $d_t = d_f = 1$ for a curve, $d_t = d_f = 2$ for a surface, and $d_t = d_f = 3$ for a solid. The following relationship holds for the three expressions of dimensionality

$$d_t \leq d_f \leq d_e.$$

Although we have used the value of the fractal dimension d_f as a means to quantify the degree of disorderliness, it is the magnitude of the difference $d_f - d_t$ that in essence reflects how irregular (disordered) the system is. Geometrically speaking, this difference $d_f - d_t$ allows the disordered system to accommodate structure within structure, and the larger this difference is, the more disordered the system.

The above-defined d_f and d_t are *structural parameters* characterizing only the geometry of a given medium. However, when we are interested in processes like diffusion or reactions in disordered media, we need *functional parameters*, which are associated with the notion of time in order to characterize the dynamic behavior of the species in these media. The spectral or fracton dimension d_s and random-walk dimension d_w are two such parameters, and they will be defined in Section 2.2.

1.7 Percolation

The origins of percolation theory are usually attributed to Flory and Stockmayer [5–8], who published the first studies of polymerization of multifunctional units (monomers). The polymerization process of the multifunctional monomers leads to a continuous formation of bonds between the monomers, and the final ensemble of the branched polymer is a network of chemical bonds. The polymerization reaction is usually considered in terms of a lattice, where each site (square) represents a monomer and the branched intermediate polymers represent clusters (neighboring occupied sites), Figure 1.4 A. When the entire network of the polymer, i.e., the cluster, spans two opposite sides of the lattice, it is called a *percolating cluster*, Figure 1.4B.

In the model of bond percolation on the square lattice, the elements are the bonds formed between the monomers and not the sites, i.e., the elements of the clusters are the connected bonds. The extent of a polymerization reaction corresponds to the fraction of reacted bonds. Mathematically, this is expressed by the probability p for the presence of bonds. These concepts can allow someone to create randomly connected bonds (clusters) assigning different values for the probability p. Accordingly, the size of the clusters of connected bonds increases as the probability p increases. It has been found that above a critical value of $p_c = 0.5$ the various bond configurations that can be formed randomly share a common characteristic: a cluster percolates through the lattice. A more realistic case of a percolating cluster can be obtained if the site model of a square lattice is used with probability $p = 0.6$, Figure 1.5. Notice that the critical value of p_c is 0.593 for the

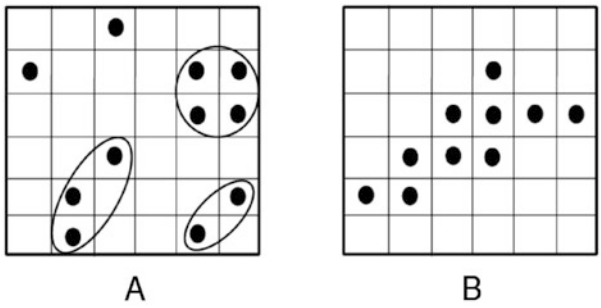


Fig. 1.4 A 6×6 square lattice site model. The dots correspond to multifunctional monomers. (**A**) The encircled neighboring occupied sites are clusters (branched intermediate polymers). (**B**) The entire network of the polymer is shown as a cluster that percolates through the lattice from left to right

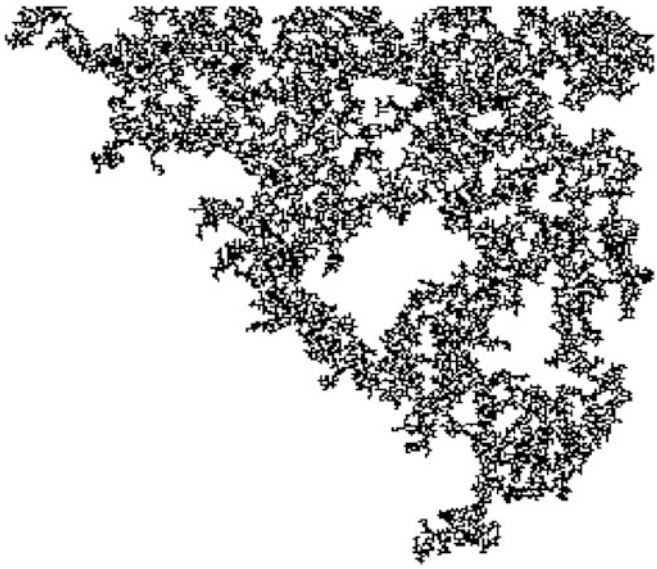


Fig. 1.5 A percolation cluster derived from computer simulation in a 300×300 square site model with $p = 0.6$. Only the occupied sites that belong to the percolating cluster are shown

two-dimensional site model. Also, the percolation thresholds vary according to the type of model (site or bond) as well as with the dimensionality of the lattice (2 or 3).

The most remarkable properties of percolation clusters arise from their sudden inception when the bond concentration (probability) reaches the critical threshold value $p = p_c$. At this specific value the emerged cluster spans two opposite sides of the lattice and if one conceives of the bonds as channels, the cluster allows a fluid to flow through the medium from edge to edge. Accordingly, the terms *percolation* and *percolation transition* have been coined in an attempt to capture the sudden

change in the geometry and the phase transition. In the same vein, the probability p_∞ that a bond belongs to the percolating cluster undergoes a sharp transition, i.e., $p_\infty = p = 0$ for $p_\infty = p < p_c$, while p_∞ becomes finite following a power law when $p > p_c$

$$p_\infty \propto (p - p_c)^\lambda ,$$

where λ is an exponent usually called the *critical exponent*. According to the findings in this field of research the critical exponent λ depends exclusively on the dimensionality of the system. This independence from other factors is characterized as *universality*.

Important characteristics of the clusters like the mass q and the typical length ξ of the clusters, usually called the *correlation length*, obey power laws too

$$q \propto |p - p_c|^{-\mu} , \qquad \xi \propto |p - p_c|^{-\nu} ,$$

where μ and ν are also critical exponents. These laws allow reconsideration of the fractal properties of the clusters. According to the last equation the clusters are self-similar as long as the length scale used for measurements is shorter than ξ. For example, the giant cluster shown in Figure 1.5 is a random fractal and as such has a characteristic value for its fractal dimension d_f. However, the calculation of the fractal dimension for the percolating cluster of Figure 1.5 should be performed with radii ρ shorter than ξ. In other words, when $\rho < \xi$ the self-similar character of the cluster is kept and the scaling law holds. Indeed, when the box-counting method is applied, the scaling law $q \propto \rho^{1.89}$ between the mass q (calculated from the mass of ink or equivalently from the number of dots) and the radius ρ of the box is obtained. This means that $d_f = 1.89$ for the percolating cluster of Figure 1.5 since the characteristic measured is the mass for various radii ρ, and no further calculations are required in accord with (1.8). On the contrary, for measurements with $\rho > \xi$, self-similarity no longer exists.

Chapter 2
Diffusion and Kinetics

Everything changes and nothing stands still.

Heraclitus of Ephesus (544-483 BC)

The principles of physical and chemical laws are essential for the understanding of drug kinetics in mammalian species. This also applies to pharmacodynamics since the interaction of drug with the receptor(s) relies on the physicochemical principles of the law of mass action. In reality one can consider the entire course of drug in the body as consecutive and/or concurrent processes of diffusion and convection. For example, the oral administration of a drug may include, among many others, the following processes:

- dissolution in the gastrointestinal fluids (diffusion),
- transport in the chyme by intestinal peristalsis (convection),
- transcellular uptake (diffusion),
- transport with the blood to organs (convection),
- transfer from the bloodstream into the interstitial and intracellular spaces (diffusion),
- interaction with receptors at the effect site (diffusion),
- transfer from tissues back into blood (diffusion),
- glomerular filtration (convection),
- transport with the urine into the efferent urinary tract (convection),
- reabsorption from the tubular lumen to the peritubular capillary (diffusion).

The above convection processes are the result of the movement of a liquid in bulk, i.e., the flow of the biological fluid. Consequently, convection processes are particularly dependent on physiology. For example, the glomerular filtration of a drug is extremely important from a therapeutic point of view, but it is solely determined by the physiological condition of the patient, e.g., the glomerular filtration rate. This is so, since a common translational velocity is superposed on the thermal motions of all drug molecules in any element of volume. On the other hand, convection processes for the dissolved and undissolved drug in the gastrointestinal tract are much more complicated. Here, physiology still plays a major role but dietary conditions and the type of formulation are important too. The picture becomes even more complicated

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_2

if one takes into account the oscillatory nature of intestinal motility, which is related to the food intake. Despite the complexity involved, the term convection implies that both dissolved drug molecules and undissolved drug particles along with the gastrointestinal fluid molecules are transported together without separation of individual components of the solution/suspension.

On the other hand, diffusion is the random migration of molecules or small particles arising from motion due to thermal energy. Here, drug diffusive fluxes are produced by differences in drug concentrations in different regions. Thus, diffusion is one of the most significant process in all fields of pharmaceutical research either in vitro or in vivo. This is justified by the fact that everything is subject to thermal fluctuations, and drug molecules or particles immersed in aqueous environments are in continuous riotous motion. Therefore, understanding of these random motions is crucial for a sound interpretation of drug processes.

2.1 Random Walks and Regular Diffusion

Particles under the microscope exhibiting Brownian motion demonstrate clearly that they possess kinetic energy. We are also familiar with the diffusional spreading of molecules from the classical experiment in which a drop of dye is carefully placed in an aqueous solution. Fick's laws of diffusion describe the spatial and temporal variation of the dye molecules in the aqueous solution. However, before presenting Fick's differential equation, attention will be given to a proper answer for the fundamental question: How much do the molecules move on average during diffusional spreading?

The correct answer to the above question is a law of physics: "the mean square displacement is proportional to time." We can intuitively reach this conclusion with particles executing an imaginary one-dimensional random walk. A simple model is presented in Figure 2.1, ignoring the detailed structure of the liquid and temperature effects and assuming no interaction between particles. The particles are placed at $z = 0$ and start their random walk at $t = 0$ moving at a distance δ either to the right or to the left once every $t_\circ$ units of time; thus, the particles execute i steps in time $t = it_\circ$. Equal probabilities $(1/2)$ are assigned for each movement of the particles (either to the right or to the left). This means that the successive jumps of particles are statistically independent and therefore the walk is unbiased. We say that the particles are blind since they have no "memory" of their previous movement(s).

Fig. 2.1 A one-dimensional random walk of particles placed at $z = 0$ at $t = 0$. The particles occupy only the positions 0, $\pm\delta$, $\pm 2\delta$, $\pm 3\delta$, $\pm 4\delta$

The question arises: How far will a particle travel in a given time interval? The average distance a particle travels is given by mean square displacement evaluated as follows: The position of a particle along the z-axis after i steps z_i is

$$z_i = z_{i-1} \pm \delta, \tag{2.1}$$

where z_{i-1} is the position of the particle at the previous $(i-1)$-th step. Taking the square of (2.1) we get the square displacement

$$z_i^2 = z_{i-1}^2 \pm 2\delta z_{i-1} + \delta^2,$$

which if averaged for the total number of particles, provides their mean square displacement $\langle z_i^2 \rangle$:

$$\langle z_i^2 \rangle = \langle z_{i-1}^2 \rangle \pm 2\delta \langle z_{i-1} \rangle + \delta^2 = \langle z_{i-1}^2 \rangle + \delta^2. \tag{2.2}$$

The second term in the brackets vanishes since the plus sign corresponds to half of the particles and the minus sign to the other half. Given that $z_0 = 0$ and applying (2.2) for the successive steps $1, 2, \ldots, i$, we get

$$\langle z_1^2 \rangle = \delta^2, \langle z_2^2 \rangle = 2\delta^2, \ldots, \langle z_i^2 \rangle = i\delta^2. \tag{2.3}$$

Since as previously mentioned the number of steps is proportional to time ($i = t/t_\circ$), we can express the positioning of particles as a function of time t using (2.3):

$$\langle z^2(t) \rangle = \left(\delta^2 / 2t_\circ \right) t. \tag{2.4}$$

The use of 2 in the denominator of the previous equation will be explained in Section 2.4. The last expression shows that the mean square displacement of the particles is proportional to time, t:

$$\langle z^2(t) \rangle \propto t. \tag{2.5}$$

The same result is obtained if one considers a simple random walk in two dimensions, i.e., the walk is performed on a two-dimensional lattice. Here, the walker (particle) moves either vertically or horizontally at each time step ($t_\circ$ units of time) with equal probabilities. Two configurations for eight-time-step random walks are shown in Figure 2.2A, along with the trail of a random walk of 10,000 steps, Figure 2.2B. In the general case and assuming that the lattice spacing is δ, the position of the walker on the plane after i steps z_i is

$$z_i = \delta \sum_{j=1}^{i} u_j,$$

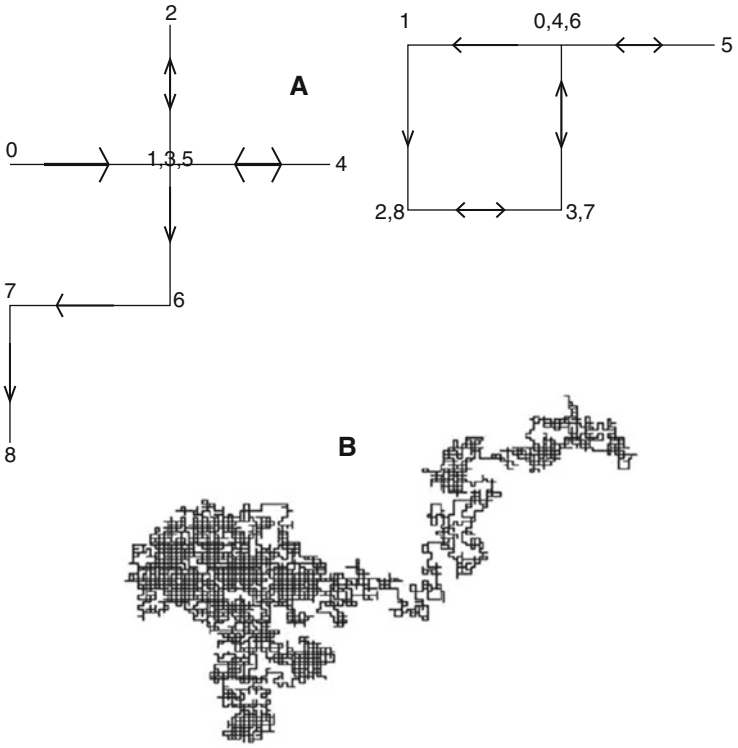


Fig. 2.2 (**A**) Two configurations of eight-step random walks in two dimensions. The numbers correspond to the successive eight steps and the arrows indicate the direction of movement. (**B**) A random walk of 10, 000 steps

where u_j is a (unit) vector pointing to a nearest-neighbor site; it represents the j-th step of the walk on the two-dimensional lattice. The mean displacement $\langle z_i \rangle$ of the walker can be obtained if z_i is averaged for the total number of walkers, $\langle z_i \rangle = 0$. This equation is obtained from the previous one since $\langle u_j \rangle = 0$. Moreover, the mean square displacement can be obtained from the previous equation if one takes into account that $\langle u_j u_j \rangle = 1$, and $\langle u_j u_k \rangle = 0$:

$$\begin{aligned} \langle z_i^2 \rangle &= \left\langle \left[\delta \sum_{j=1}^{i} u_j \right]^2 \right\rangle \\ &= \delta^2 \left\langle (u_1 + u_2 + \ldots + u_i)(u_1 + u_2 + \ldots + u_i) \right\rangle \\ &= \delta^2 \sum_{j=1}^{i} \langle u_j u_j \rangle + \delta^2 \sum_{\substack{j=1 \\ k \neq j}}^{i} \langle u_j u_k \rangle = i\delta^2. \end{aligned} \tag{2.6}$$

Substituting $i = t/t_\circ$ in the last equation, (2.4) is recovered using the factor $\frac{1}{2}$ for the derivation once again.

The theory for motion in three dimensions results in the same law if the same assumptions are applied and motions in the three directions are statistically independent. The important result for regular diffusion is that its time dependence is universal regardless of the dimension of the medium. This square root relation (2.5) has striking consequences for the distance covered by diffusing molecules. It takes *four* times as long to get *twice* as far while a particle can cover *half* the distance in a *quarter* of the time. Thus, transport by diffusion is very slow if there is far to go, but very rapid over very short distances. For example, the exchange and transport of solutes within cells and between cells and capillaries can be effectively maintained by diffusion due to the small size and close spacing of cells and capillaries in the body of mammals. On the contrary, the slowness of diffusion over large distances points to the necessity for a circulatory system to bring oxygen, for example, from the lungs to the brain or glucose from the liver to the muscles of the arms. To permit these exchanges, the bulk flow of blood carries a large number of solutes around the body in the vascular system by convection.

Equation (2.4) will help us to define and understand the meaning of the diffusion coefficient $\mathcal{D}$. This term corresponds to the proportionality constant of (2.4),

$$\mathcal{D} \triangleq \frac{\delta^2}{2t_\circ}, \tag{2.7}$$

has dimensions of area×time^{-1} and takes different values for different solutes in a given medium at a given temperature. Hence, the value of $\mathcal{D}$ is characteristic for a given solvent (or better, medium structure) at a given temperature of the diffusing tendency of the solute. For example, a small drug molecule in water at 25 °C has $\mathcal{D} \approx 10^{-5}\,\mathrm{cm}^2/\,\mathrm{s}$, while a protein molecule like insulin has $\mathcal{D} \approx 10^{-7}\,\mathrm{cm}^2/\,\mathrm{s}$. Using these values one can roughly calculate the time required for the drug and protein molecules to travel a distance of 1 mm; it takes $(0.1)^2/10^{-5} \approx 1000\,\mathrm{s} \approx 16.6\,\mathrm{min}$ for the drug and 1666.6 min for insulin. Hence, the value of $\mathcal{D}$ is heavily dependent on the size of the solute molecules. These numerical calculations are very useful in obtaining insight into the rapidity or slowness of a solute migration, e.g., drug release from controlled release formulations when regular diffusion is the operating mechanism.

2.2 Anomalous Diffusion

In the previous section we analyzed the random walk of molecules in Euclidean space and found that their mean square displacement is proportional to time (2.5). Interestingly, this important finding is not true when diffusion is studied in fractals and disordered media. The difference arises from the fact that the nearest-neighbor

sites visited by the walker are equivalent in spaces with integer dimensions but are not equivalent in fractals and disordered media. In these media the mean correlations between different steps $\langle u_j u_k \rangle$ are not equal to zero, in contrast to what happens in Euclidean space; cf. derivation of (2.6). In reality, the anisotropic structure of fractals and disordered media makes the value of each of the correlations $u_j u_k$ structurally and temporally dependent. In other words, the value of each pair $u_j u_k$ depends on where the walker is at the successive times j and k, and the Brownian path on a fractal may be a "fractal of a fractal" [9]. Since the correlations $u_j u_k$ do not average out, the final important result is $\langle u_j u_k \rangle \neq 0$, which is the underlying cause of anomalous diffusion. In reality, the mean square displacement does not increase linearly with time in anomalous diffusion and (2.5) is no longer exact.

To characterize the dynamic movement of particles on a fractal object, one needs two additional parameters: the spectral or fracton dimension d_s and the random-walk dimension d_w. Both terms are quite important when diffusion phenomena are studied in disordered systems. This is so since the path of a particle or a molecule undergoing Brownian motion is a random fractal. A typical example of a random fractal is the percolation cluster shown in Figure 1.5.

The definition of spectral dimension d_s refers to the probability $p(t)$ of a random walker returning to its origin after time t:

$$p(t) \propto t^{-d_s/2}. \tag{2.8}$$

According to (2.8), the value of d_s governs the decrease of the probability $p(t)$ with time. When diffusion is considered in Euclidean spaces the various dimensionality terms become identical: $d_t = d_s = d_f$. However, in fractal spaces the following inequalities hold: $d_t < d_s < d_f < d_e$, where d_e is the embedding dimension. For example, we found for the Sierpinski gasket (Figure 1.2A) $d_f = 1.5815$, while $d_s = 1.3652$ and the embedding dimension in this case is $d_e = 2$. The meaning of d_s can be understood if one considers a walker executing a random walk on a ramified system, like the Sierpinski gasket with $d_f = 1.5815$, Figure 1.2A. Due to the system's ramification, the walker has many alternatives of movement in the branched system, and therefore the probability of the walker being back at the origin is small. Hence, the value of d_s goes up in accord with (2.8) and is higher than one ($d_s > 1$), i.e., the topological dimension of a curve. In actual practice, the calculation of d_s is accomplished numerically. Analytical solutions for d_s are available when the recursion algorithm of the system is known, e.g., Sierpinski gasket.

Finally, a stochastic viewpoint may be associated with the relation (2.8) since the spectral dimension also characterizes the number $n(t)$ of distinct sites visited by the random walker up to time t:

$$n(t) \propto t^{d_s/2}. \tag{2.9}$$

The random-walk dimension d_w is useful whenever one has a specific interest in the fractal dimension of the trajectory of the random walk. The value of d_w is exclusively dependent on the values of d_f and d_s:

$$d_w = \min\left[2\frac{d_f}{d_s}, d_f\right].$$

The type of the random walk (recurrent or nonrecurrent) determines the minimum value of the two terms in the brackets of the previous equation. If the walker does not visit the same sites (nonrecurrent), then $d_w = 2d_f/d_s$. If the walk is of recurrent type, then the walker visits the same sites again and again and therefore the walker covers the available space (space-filling walk). Consequently, the meaning of d_w coincides with d_f ($d_w = d_f$). The mean square displacement in anomalous diffusion follows the pattern

$$\left\langle z^2(t)\right\rangle \propto t^{2/d_w}, \tag{2.10}$$

where d_w is the fractal dimension of the walk and its value is usually $d_w > 2$. The exponent d_w arises from the obstacles of the structure such as holes, bottlenecks, and dangling ends, i.e., the diffusional propagation is hindered by geometric heterogeneity. The previous equation is the fundamental relation linking the propagation of the diffusion front to the structure of the medium, and it recovers also the classical law of regular diffusion when $d_w = 2$.

In conclusion, the dynamic movement of particles on a fractal object may be described by functional characteristics such as the spectral dimension d_s and the random-walk dimension d_w. This anomalous movement of the molecules induces heterogeneous transport and heterogeneous reactions. Such phenomena present a challenge to several branches of science: chemical kinetics, surface and solid state physics, etc. Consequently, one may argue that all mechanisms involved in drug absorption, metabolism, enzymatic reactions, and cell microscopic reactions can be analyzed in the new heterogeneous context since these processes are taking place under topological constraints.

2.3 Fick's Laws of Diffusion

Apart from the above considerations of diffusion in terms of the distance traveled in time, the amount of substance transported per unit time is useful too. This approach brings us to the concept of the rate of diffusion. The two considerations are complementary to each other since the diffusion of molecules at the microscopic level results in the observed "flux" at the macroscopic level. Fick's laws of diffusion describe the flux of solutes undergoing classical diffusion.

The simplest system to consider is a solution of a solute with two regions of different concentrations c_l and c_r to the left and right, respectively, of a boundary separating the two regions, Figure 2.3. In reality, the rate of diffusion is the net flux, i.e., the difference between the two opposite unidirectional fluxes. There will be a net movement of solute molecules to the right if $c_l > c_r$ or to the left if $c_l < c_r$. When $c_l = c_r$, the unidirectional fluxes are equal and the net flux is zero. Since the

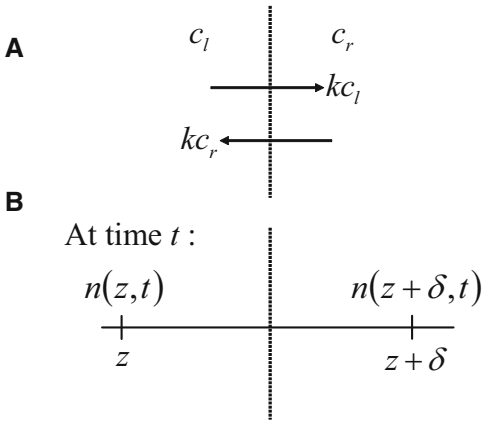


Fig. 2.3 A solute diffuses across a plane. (**A**) Solute diffusion from two regions of different concentrations c_l and c_r; the plane indicates the boundary of the regions. The transfer rate of material is proportional to concentrations c_l and c_r. (**B**) At a given time t there are $n(z,t)$ and $n(z+\delta,t)$ molecules at positions z and $z+\delta$, respectively

two fluxes across the boundary from left to right and vice versa are proportional to c_l and c_r, respectively, the net flux is proportional to the concentration difference across the boundary.

The derivation of Fick's first law of diffusion requires a reconsideration of Figure 2.3A in terms of the one-dimensional random walk as shown in Figure 2.3B. Let us suppose that at time t, there are $n(z,t)$ molecules at the left position z and $n(z+\delta,t)$ molecules at the right position $z+\delta$, Figure 2.3B. Since equal probabilities $(1/2)$ are assigned for the movement of the molecules (either to the right or to the left), half of the $n(z,t)$ and $n(z+\delta,t)$ molecules will cross the plane at the next instant of time $t+t_\circ$, moving in opposing directions. The net number of molecules crossing the plane to the right is $-\frac{1}{2}\left[n\left(z+\delta,t\right)-n\left(z,t\right)\right]$ and the corresponding net flux J of the diffusate is

$$J\left(z,t\right)=-\frac{1}{2\mathcal{A}t_\circ}\left[n\left(z+\delta,t\right)-n\left(z,t\right)\right],$$

where $\mathcal{A}$ is the area of the plane and $t_\circ$ is the time interval. Multiplying and dividing the right part by δ^2 and rearranging, we get

$$J\left(z,t\right)=-\frac{\delta^2}{2t_\circ}\frac{1}{\delta}\left[\frac{n\left(z+\delta,t\right)}{\mathcal{A}\delta}-\frac{n\left(z,t\right)}{\mathcal{A}\delta}\right].$$

The terms in the brackets express the concentration of molecules per unit volume $\mathcal{A}\delta$, i.e., $c(z+\delta,t)\equiv c_r\left(t\right)$ and $c(z,t)\equiv c_l\left(t\right)$ at positions $z+\delta$ and z, respectively, while the term $\delta^2/2t_\circ$ is the diffusion coefficient $\mathcal{D}$; the presence of 2 in the denominator explains its use in (2.4). We thus obtain

$$J\left(z,t\right)=-\mathcal{D}\frac{c\left(z+\delta,t\right)-c\left(z,t\right)}{\delta}.$$

Since the term in the brackets in the limit $\delta \to 0$ is the partial derivative of $c(z,t)$ with respect to z, one can write

$$J(z,t) = -\mathcal{D}\frac{\partial c(z,t)}{\partial z}. \tag{2.11}$$

The minus sign indicates that the flow occurs from the concentrated to the dilute region of the solution. Equation (2.11) is Fick's first law, which states that the net flux is proportional to the gradient of the concentration function (at z and t). Flux has dimensions of mass×area^{-1}×time^{-1}.

Since the flux J is the flow of material $\dot{q}(z,t)$ from the left to the right through the surface $\mathcal{A}$, (2.11) is rewritten as follows:

$$\dot{q}(z,t) = -\mathcal{D}\mathcal{A}\frac{\partial c(z,t)}{\partial z}. \tag{2.12}$$

From this relationship it is clear that the force acting to diffuse the material q through the surface is the concentration gradient $\partial c/\partial z$. This gradient may be approximated by differences

$$\frac{\partial c(z,t)}{\partial z} \approx \frac{\Delta c(z,t)}{\Delta z} = \frac{c(z+\delta,t) - c(z,t)}{\delta} = \frac{c_r(t) - c_l(t)}{\delta}, \tag{2.13}$$

and the previous expression becomes

$$\dot{q}(t) \triangleq R_{lr} = -\frac{\mathcal{D}\mathcal{A}}{\delta}\left[c_r(t) - c_l(t)\right], \tag{2.14}$$

where R_{lr} is the transfer rate of material. This equation usually takes one of two similar forms:

$$\dot{q}(t) = -CL_{lr}\left[c_r(t) - c_l(t)\right] \qquad \text{or} \qquad \dot{q}(t) = -P\mathcal{A}\left[c_r(t) - c_l(t)\right]. \tag{2.15}$$

The new introduced parameter $CL_{lr} \triangleq \mathcal{D}\mathcal{A}/\delta$ is called *clearance*, and it has dimensions of flow, volume×time^{-1}. The clearance has a bidirectional use and indicates the volume of the solution that is cleared from drug per unit of time because of the drug movement across the plane. For an isotropic membrane, structural and functional characteristics are identical at both sides of the membrane, $CL_{lr} = CL_{rl}$. In practice, the term "clearance" is rarely used except for the irreversible removal of a material from a compartment by unidirectional pathways of metabolism, storage, or excretion. The other new parameter $P \triangleq \mathcal{D}/\delta$ characterizes the diffusing ability of a given solute for a given membrane, and it is called *permeability*. Permeability has dimensions of length×time^{-1}.

We now write a general mass conservation equation stating that the rate of change of the amount of material in a region of space is equal to the rate of flow across the

boundary plus any that is created within the boundary. If the region is $z_1 < z < z_2$ and no material is created

$$\frac{\partial}{\partial t}\int_{z_1}^{z_2} \mathrm{d}q\,(z,t) = \frac{\partial}{\partial t}\int_{z_1}^{z_2} c\,(z,t)\,\mathrm{d}z = J\,(z_1,t) - J\,(z_2,t)\,.$$

Here, if we assume $\mathcal{D}$ constant in (2.11) and $z_2 = z_1 + \Delta z$, at the limit $\Delta z \to 0$, this relation leads to

$$\frac{\partial c\,(z,t)}{\partial t} = \mathcal{D}\frac{\partial^2 c\,(z,t)}{\partial z^2}. \tag{2.16}$$

This is the second Fick's law stating that the time rate of change in concentration (at z and t) is proportional to the curvature of the concentration function (at z and t). There is a clear link between the two laws (2.11) and (2.16).

In order to examine the relevance of the two laws, let us consider that the layer separating the two regions in Figure 2.3A is not thin but has an appreciable thickness δ, while z is the spatial coordinate along it. According to (2.11), if $\partial c/\partial z$ is constant, then the flux J is constant. This happens when c is a linear function of z. Consequently, $\partial^2 c/\partial z^2 = 0$ in (2.16) and this implies the steady-state condition $\partial c\,(z,t)\,/\partial t = 0$, where the concentration is stationary in time. Under these conditions, as many drug molecules diffuse in from the side of higher concentration as diffuse out to the side of lower concentration. This can be accomplished experimentally if the concentrations c_l and c_r in the two regions of Figure 2.3A are maintained constant. With boundary conditions $c(0,t) = c_l$ and $c(\delta,t) = c_r$, and initial condition $c(z,0) = 0$, the solution of (2.16) is given by [10]

$$\begin{aligned} c\,(z,t) = {} & c_l - (c_l - c_r)\,\frac{z}{\delta} \\ & - \frac{4c_l}{\pi}\sum_{i=1}^{\infty}\frac{1}{2i-1}\sin\left[(2i-1)\,\pi\,\frac{z}{\delta}\right]\exp\left[-\frac{(2i-1)^2\,\pi^2}{\delta^2}\mathcal{D}t\right] \\ & + \frac{2\,(c_l - c_r)}{\pi}\sum_{i=1}^{\infty}\frac{(-1)^{i+1}}{i}\sin\left(i\pi\,\frac{z}{\delta}\right)\exp\left(-\frac{i^2\pi^2}{\delta^2}\mathcal{D}t\right). \end{aligned} \tag{2.17}$$

By using the above relationship, Figure 2.4 simulates the distance–concentration profiles $c\,(z,t)$ at times $t = 15$ min, 1 and 5 h with $\mathcal{D} = 0.1$ cm^2/ h, $\delta = 1$ cm, $c_l = 10$ and $c_r = 2$ g/ l. Since there is no solute inside the layer initially ($c(z,0) = 0$), for early times (e.g., $t = 15$ min) the solute molecules undergo diffusion with two opposite directions, from the boundaries to the interior of the layer ($\partial c/\partial z < 0$ and $J\,(z,t) > 0$ for $0 \le z < z^{\bullet}$; $\partial c/\partial z > 0$ and $J\,(z,t) < 0$ for $z^{\bullet} \le z < 1$ cm with $z^{\bullet} \approx 0.6$ cm according to Figure 2.4). As time grows, the diffusion becomes unidirectional with $\partial c/\partial z < 0$ and $J\,(z,t) > 0$ because $c_l > c_r$. As time goes by

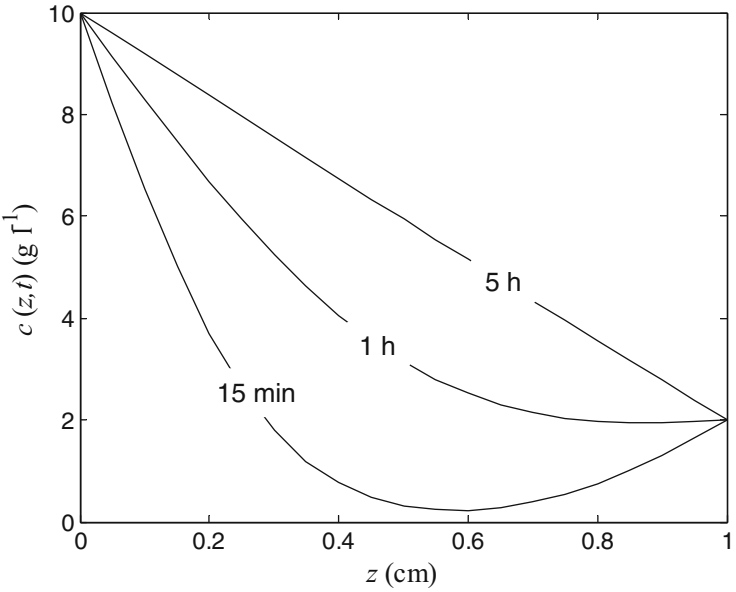


Fig. 2.4 Simulation of distance–concentration profiles $c\,(z,t)$ at times $t = 15$ min, 1 and 5 h with $\mathcal{D} = 0.1$ cm^2/ h, $\delta = 1$ cm, $c_l = 10$ and $c_r = 2$ g/ l

(e.g., $t = 5$ h), the steady state is reached, the solution of the partial differential equation 2.16 is $c\,(z,.) = c_l - (c_l - c_r)\,\frac{z}{\delta}$ and according to the definition 2.11 the net flux

$$J\,(.,.) = \frac{\mathcal{D}}{\delta}\,(c_l - c_r)$$

is constant.

If we postulate that molecules move independently, the concentration $c\,(z,t)$ at some point z is proportional to the probability density $p\,(z,t)$ of finding a molecule there. Thus, the diffusion partial differential equation (2.16) holds when probability densities are substituted for concentrations:

$$\frac{\partial p\,(z,t)}{\partial t} = \mathcal{D}\frac{\partial^2 p\,(z,t)}{\partial z^2}. \tag{2.18}$$

If a molecule is initially placed at $z = 0$, then the solution of the previous equation is

$$p\,(z,t) = (4\pi\mathcal{D}t)^{-1/2}\exp\left(-\frac{z^2}{4\mathcal{D}t}\right).$$

For $t \gg 1$ at any z, we obtain $p\,(z,t) \propto t^{-1/2}$. This behavior in a homogeneous medium corresponds to (2.8), giving the probability density in a fractal medium with spectral dimension d_s.

2.4 Classical Kinetics

Pharmacy, like biology and physiology, is wet and dynamic. Drug molecules immersed in the aqueous environment of intravascular, extravascular, and intracellular fluids participate in reactions, such as reversible binding to membrane or plasma proteins; biotransformation or transport processes, e.g., drug release from a sustained release formulation; drug uptake from the gastrointestinal membrane; and drug permeation through the blood–brain barrier. This classification is very rough since some of these processes are more complex. For example, drug release is basically a mass transport phenomenon but may involve reaction(s) too, e.g., polymer dissolution and/or polymer transition from the rubbery to the glassy state. However, irrespective of the detailed characteristics, the common and principal component of the underlying mechanism of numerous drug processes is diffusion. This is the case for the ubiquitous passive transport processes that rely on diffusion exclusively. The value of $\mathcal{D}$ depends on the nature of the environment of the diffusing species. If the environment changes from one point to another, the value of $\mathcal{D}$ may depend on position. Usually, we deal with systems in which the environment of the diffusing species is the same everywhere, so that $\mathcal{D}$ is a constant. The diffusion coefficient is constant for diffusion of dilute solute in a uniform solvent. This case takes in a large number of important situations, and if the dilute solute is chemically the same as the solvent but is isotopically tagged, then the diffusion is termed self-diffusion. In contrast, chemical reactions can be either reaction-limited or diffusion-limited. In the following sections we will discuss them separately.

2.4.1 Passive Transport Processes

There appear to be two main ways for solutes to pass through cell membranes, namely, transcellular and paracellular. The most important is the transcellular route, whereby compounds cross the cells by traversing the cell membrane following either passive diffusion or carrier-mediated transport. Undoubtedly, the transcellular passive diffusion is the basic mechanism of solute permeation through cell membranes. According to this mechanism the solute leaves the fluid bathing the membrane, dissolves in the substance of the membrane, diffuses across in solution, and then emerges into the intracellular fluid. Accordingly, the mathematical treatment of drug diffusion across a membrane can be based on (2.12), which is a very useful expression of Fick's first law of diffusion. This equation is used extensively in the pharmaceutical sciences. It describes the mass (number of molecules, or moles, or amount) transported per unit time, $\dot{q}$, across an area $\mathcal{A}$ with a concentration gradient $\partial c/\partial z$ at right angles to the area. According to this definition, the numerical value of the diffusion coefficient $\mathcal{D}$, expressed in mass units, corresponds to the amount of solute that diffuses per unit time across a unit area under the influence of a unit concentration gradient.

For a passive transport process, the concentration gradient across the membrane can be considered constant and therefore the gradient can be approximated by differences as in (2.13) to obtain

$$\dot{q}(t) = \frac{\mathcal{D}'\mathcal{A}}{\delta}\left[c_l(t) - c_r(t)\right],$$

where $\mathcal{D}'$ is a modified diffusion coefficient, for restricted diffusion inside the membrane. The value of $\mathcal{D}'$ is much smaller than the diffusion coefficient $\mathcal{D}$ in free solution. The minus sign is not used in the previous equation since the rate of transport corresponds to the solute transfer from the external to the internal site ($c_l > c_r$). Furthermore, if sink conditions prevail ($c_l \gg c_r$), the previous equation can be simplified to

$$\dot{q}(t) = CLc(t) = P\mathcal{A}c(t). \tag{2.19}$$

The last equation reveals that estimates for P can be obtained in an experimental setup if the permeation rate $\dot{q}(t)$ and the total membrane area $\mathcal{A}$ available for transport are measured and the drug concentration $c(t)$ in the donor compartment remains practically constant. What is implicit from all the above is that the diffusion coefficient $\mathcal{D}'$ is at the origin of the definition of the clearance CL and permeability P, and these parameters are incorporated into the global rate constant of the rate equations used in pharmacokinetics. For example, the first-order absorption rate constant k_a in the following equation is proportional to the diffusion coefficient $\mathcal{D}'$ of drug in the gastrointestinal membrane:

$$\dot{c}_b(t) = k_a c_{GI}(t),$$

where $c_b(t)$ and $c_{GI}(t)$ denote drug concentration (amount absorbed/volume of distribution) in blood and in the gastrointestinal lumen (amount dissolved in the gastrointestinal fluids/volume of gastrointestinal fluids), respectively. In other words, $\mathcal{D}'$ controls the rate of drug absorption from the gastrointestinal tract.

2.4.2 *Reaction Processes: Diffusion- or Reaction-Limited?*

Pharmacokinetics has been based on the concepts of classical chemical kinetics. However, the applicability of the rate equations used in chemical kinetics presupposes that the reactions are really reaction-limited. In other words, the typical time for the two chemical species to react when placed in close proximity (reaction time t_{reac}) is larger than the typical time needed for the two species to reach each other (diffusion time t_{diff}) in the reaction space. When the condition $t_{\text{reac}} > t_{\text{diff}}$ is met, then one can use the global concentrations of the reactant species in the medium to obtain the classical rate equations of chemical kinetics. This is so since the rate of

the reaction is proportional to the global concentrations of the reactant species (law of mass action). The inequality $t_{\text{reac}} > t_{\text{diff}}$ underlines the fact that the two reactant species have encountered each other more than one time previously in order to react effectively.

The opposite case, $t_{\text{reac}} < t_{\text{diff}}$, indicates that the two reactant species actually react upon their first encounter. The diffusion characteristics of the species control the rate of the reaction, and therefore these reactions are called *diffusion limited*. Consider, for example, a system consisting of species A and B with n_A and n_B molecules of A and B, respectively. The problem of the reaction rate between A and B is in essence reduced to the rate at which A and B molecules will encounter one another. The principal parameters governing the reaction rate are the diffusion coefficients $\mathcal{D}_A$ and $\mathcal{D}_B$ of the reactant species since they determine the diffusing tendency of the species. Focusing on B molecules, it can be proven that the rate of B molecules diffusing to an A molecule is proportional to the diffusion coefficient of B, the number of B molecules, and the distance between A and B, namely, $4\pi\mathcal{D}_B(\rho_A + \rho_B)n_B$, where $\rho_A + \rho_B$ is the distance between the centers of A and B molecules; accordingly, the total rate of A and B encounters is $4\pi\mathcal{D}_B(\rho_A + \rho_B)n_Bn_A$. In an analogous manner the total rate of A and B encounters, viewed in terms of the A molecules, is $4\pi\mathcal{D}_A(\rho_A + \rho_B)n_Bn_A$. The mean of these separate rates provides a reasonable expression for the rate per unit volume for A and B molecules separately:

$$\text{Rate of } A \text{ and } B \text{ encounters} = 2\pi(\mathcal{D}_A + \mathcal{D}_B)(\rho_A + \rho_B)n_An_B.$$

Although the previous equation signifies the importance of the diffusion characteristics of the reactant species, it cannot be used to describe adequately the rate of the reaction. The reason is that the concept of global concentrations for the n_A and n_B molecules is meaningless, since a unit volume cannot be conceived due to the local fluctuations of concentrations. Hence, the local concentrations of the reactants determine the rate of the reaction for diffusion-limited reactions. Accordingly, local density functions with different diffusion coefficients for the reactant species are used to describe the diffusion component of reaction–diffusion equations describing the kinetics of diffusion-limited reactions.

2.4.3 Carrier-Mediated Transport

The transport of some solutes across membranes does not resemble diffusion and suggests a temporary, specific interaction of the solute with some component (protein) of the membrane characterized as "carrier," e.g., the small-peptide carrier of the intestinal epithelium. The rate of transport increases in proportion to concentration only when this is small, and it attains a maximal rate that cannot be exceeded even with a large further increase in concentration. The kinetics of carrier-mediated transport is theoretically treated by considering carrier–solute complexes in the same manner as enzyme–substrate complexes following the principles of enzyme-

catalyzed reactions in Michaelis–Menten kinetics. In both biotransformation and carrier-mediated transport, unrestricted diffusion is considered for the reactant species. Due to the analogous formulation of the two processes, the equations describing the rates of biotransformation,

$$\dot{c}(t) = \frac{V_{\max} c(t)}{k_M + c(t)}, \tag{2.20}$$

and carrier-mediated transport,

$$\dot{c}(t) = \frac{R_{\max} c(t)}{k_M + c(t)}, \tag{2.21}$$

are similar. In these expressions, $c(t)$ is the solute (substrate) concentration, k_M is the Michaelis constant, $V_{\max}$ is the maximum biotransformation rate, and $R_{\max}$ is the maximum transport rate. Both equations indicate that the rate of biotransformation or carrier-mediated transport becomes independent of substrate (solute) concentration when this is large. In this case, the rate of biotransformation or carrier-mediated transport is said to exhibit *saturation kinetics*. The graphical representation of the previous equations is shown in Figure 2.5.

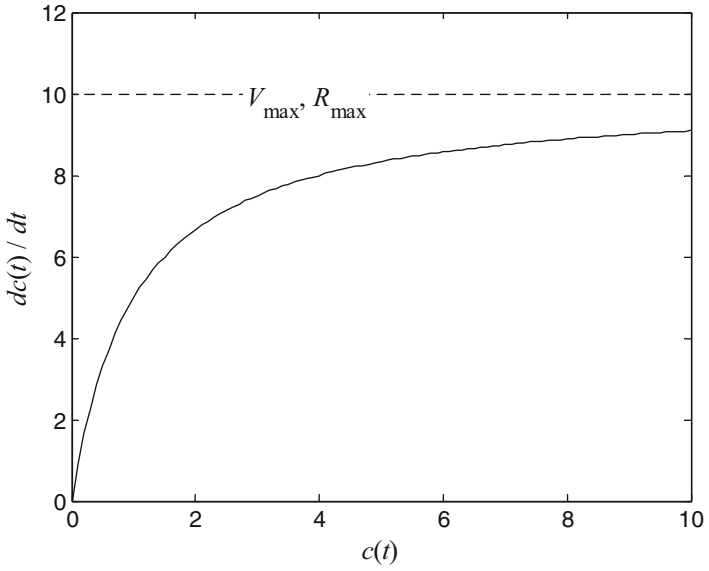


Fig. 2.5 The rate of biotranformation or carrier-mediated transport vs. solute concentration. The plateau value corresponds to $V_{\max}$ or $R_{\max}$. k_M and $V_{\max}$ were set to 1 and 10, respectively, with arbitrary units

2.5 Fractal-like Kinetics

The undisputable dogma of chemistry whether in chemical synthesis or classical chemical kinetics is to "stir well the system." The external stirring re-randomizes the positioning of the reactant species, and therefore the rate of the reaction follows the classical pattern imposed by the order of the reaction. However, many reactions and processes take place under dimensional or topological constraints that introduce spatial heterogeneity. A diffusion process under such conditions is highly influenced, drastically changing its properties. A general well-known result is that in such constrained spaces, diffusion is slowed down and diffusion follows an anomalous pattern. Obviously, the kinetics of the diffusion-limited reactions (processes) are then sensitive to the peculiarities of the diffusion process. In other words, the transport properties of the diffusing species or the reactants largely determine the kinetics of the diffusion-limited processes. Under these circumstances one can no longer rely on classical rate equations and a different approach is necessary. The drastic and unexpected consequences of nonclassical kinetics of diffusion-limited reactions are called *fractal-like kinetics*. An extensive review on the ubiquitous presence of fractals and fractal concepts in pharmaceutical sciences has been published recently [11]; the essentials for this "understirred" type of kinetics are delineated below.

2.5.1 Segregation of Reactants

Classical homogeneous kinetics assumes that the reactants are located in a three-dimensional vessel, and that during the reaction process the system is constantly stirred, thus causing the positions (locations) of the reactants to be constantly re-randomized as a function of time. However, there are important chemical reactions, which are called "heterogeneous," in which the reactants are spatially constrained by either walls or phase boundaries, e.g., liquid–solid boundaries. This is the case for in vivo drug dissolution as well as for many bioenzymatic and membrane reactions. Due to dimensional or topological constraints these heterogeneous reactions take place under understirred conditions. The most dramatic manifestation of such highly inefficient stirring is the spontaneous segregation of reactants in $A + B$ reactions [12–14]. This means that correlations begin to develop between the reactants' positions, which subsequently have a profound effect on the rate of a diffusion-controlled reaction. The build-up of such correlations is strongly dependent on the dimensionality, being more pronounced the further one goes below three-dimensional spaces. This is so because quantitatively the parameter values in the diffusion laws are very different in different dimensions. In addition, if the space where the reaction takes place is not smooth, but highly irregular, this has an added

effect on the building of such correlations. This happens if the space is a fractal structure characterized by its own dimensionality, which as discussed in Chapter 1 could be different from the integer 1, 2, or 3.

An important segregation effect is related to the violation of Wenzel's old law for heterogeneous reactions; this law states that the larger the interface, the higher the reaction rate [15]. Thus, the most classical way to speed up a heterogeneous process, e.g., drug dissolution, is to grind the material in order to increase the surface area. At the macroscopic level, this law has been verified in numerous physicochemical studies [16] as well as in in vitro drug dissolution studies and in vivo bioavailability studies using micro instead of macro drug particles. However, violation of Wenzel's law has been observed in simulation studies [17, 18] at the microscopic level. Simulations for the catalytic reaction $A + B \rightarrow AB \uparrow$, which takes place only on the rims of surfaces, indicate that the steady-state rate per unit surface area is not constant but rather depends on the size of the sample. In reality, lower reaction rates were observed for a connected catalyst compared to a disjointed one despite the fact that equal lengths for both designs were used. This is due to the lower segregation of the reactants on the rims of the disjointed catalyst, which results in a higher rate coefficient for the catalytic reaction. The clear message taken from these results is that shredding a sample not only increases the surface area but can also increase the reactivity per unit area. The latter observation violates Wenzel's law.

2.5.2 Time-Dependent Rate Coefficients

The spatial reactant correlations result in building a depletion zone around each reactant, which grows steadily with time. This means that in the close neighborhood of each reactant there is a void, a space that is empty of reactants. The net result is that the reactant distribution for the two-reactant case ($A + B \rightarrow C$) shows clear segregation of unlike species (A from B) and aggregation of like species (either A or B). Naturally, the diffusion-controlled reaction slows down, since as reactants get further apart, they must travel longer distances to find another reactant to react with (cf. equation 2.9). A curious effect now is that the rate constant k of the reaction is no longer "constant," but depends on the growth of this depletion zone and consequently is time-dependent:

$$k(t) = k_{\circ} t^{-\lambda} \qquad (t > t_{\circ}),$$

where $k(t)$ is the instantaneous rate coefficient since it depends on time t, and λ is the fractal kinetics exponent with $0 \leq \lambda < 1$. In fact, $k(t)$ crosses over from a constant regime at short times, $t < t_{\circ}$, to a power-law decrease at longer times, $t > t_{\circ}$. The switching time $t_{\circ}$ depends on the experimental conditions. This behavior is the hallmark of fractal kinetics [17].

Under homogeneous conditions (e.g., vigorous stirring), $\lambda = 0$ and therefore $k(t)$ is a constant giving back the classical kinetics result. The previous equation

has been applied to the study of various reactions in fractals as well as in many other nonclassical situations. For instance, theory, simulations, and experiments have shown that the value of λ for $A+A$ reactions is related to the spectral dimension d_s of the walker (species) as follows [9, 19]:

$$\lambda = 1 - \frac{d_s}{2}.$$

From this relationship, we obtain $\lambda = 1/3$ since the value of d_s is $\approx 4/3$ for $A + A$ reactions taking place in random fractals in all embedded Euclidean dimensions [9, 20]. It is also interesting to note that $\lambda = 1/2$ for an $A + B$ reaction in a square lattice for very long times [13]. Thus, it is now clear from theory, computer simulation, and experiment that elementary chemical kinetics are quite different when reactions are diffusion limited, dimensionally restricted, or occur on fractal surfaces [9, 12, 21–23].

We emphasize that the fractal-like kinetic characteristics are not observed only under "bing-bang" type conditions (also called batch) as discussed above but also under quasi-steady-state conditions (cf. Section 8.5.1). Consider, for example, the homodimeric reaction with two molecules of a single substrate reacting to form product ($A + A \rightarrow C$). Under homogeneous conditions the rate at quasi-steady state will be proportional to substrate concentration squared, $c^2(t)$, i.e., it is time-independent (by definition). However, the rate for the bimolecular $A + A$ diffusion-limited reaction under topological or dimensional constraints will be proportional to $c^{\gamma}(t)$. Surprisingly, the effective reaction order γ is higher than 2 and is related to the spectral dimension d_s and in turn to the fractal kinetics exponent λ [9]:

$$\gamma = 1 + \frac{2}{d_s} = 1 + (1 - \lambda)^{-1},$$

with $d_s \leq 2$. Typical values for the Sierpinski gasket and the percolation cluster are $\gamma = 2.46$ and $\gamma = 2.5$, respectively. If $d_s = 1$, so that diffusion is compact, then $\gamma = 3$ for the bimolecular $A + A$ reaction. In all these cases, the mechanism of diffusion is bimolecular. However, the increase in the effective reaction order arises from the distribution of the species, which as time goes by becomes "less random," i.e., it is actually more ordered.

Before we close this section some major, unique kinetic features and conclusions for diffusion-limited reactions that are confined to low dimensions or fractal dimensions or both can now be derived from our previous discussion. First, a reaction medium does not have to be a geometric fractal in order to exhibit fractal kinetics. Second, the fundamental linear proportionality $k \propto \mathcal{D}$ of classical kinetics between the rate constant k and the diffusion coefficient $\mathcal{D}$ does not hold in fractal kinetics simply because both parameters are time-dependent. Third, diffusion is compact in low dimensions and therefore fractal kinetics is also called *compact kinetics* [24, 25] since the particles (species) sweep the available volume compactly. For dimensions $d_s > 2$, the volume swept by the diffusing species is no longer

compact and species are constantly exploring mostly new territory. Finally, the initial conditions have no importance in classical kinetics due to the continuous re-randomization of species but they may be very important in fractal kinetics [17].

2.5.3 *Effective Rate Equations*

The dependence of kinetics on dimensionality is due to the physics of diffusion. This modifies the kinetic differential equations for diffusion-limited reactions, dimensionally restricted reactions, and reactions on fractal surfaces. All these chemical kinetic patterns may be described by power-law equations with time-invariant parameters like

$$\dot{c}(t) = -\kappa c^{\gamma}(t), \quad c(t_0) = c_0, \tag{2.22}$$

with $\gamma \geq 2$. Under these conditions, the traditional rate law for the $A + A$ reaction with concentration squared exhibits a characteristic reduction of the rate constant with time:

$$\dot{c}(t) = -k(t)\, c^2(t), \quad c(t_0) = c_0, \tag{2.23}$$

where $k(t) = k_\circ t^{-\lambda}$. Conversely, (2.23) is equivalent to a time-invariant rate law (2.22) with an increased kinetic order γ. New parameters λ and $k_\circ$ are given by

$$\lambda = (\gamma - 2)/(\gamma - 1) \text{ and } k_\circ = \kappa^{1/(\gamma-1)} (\gamma - 1)^{(2-\gamma)/(\gamma-1)}$$

with $0 \leq \lambda < 1$.

In traditional chemical kinetics $\lambda = 0$, the rate constant is time-invariant, and the effective kinetic order γ equals molecularity 2. As the reaction becomes increasingly diffusion-limited or dimensionally restricted, λ increases, the rate constant decreases more quickly with time, and the kinetic order in the time-invariant rate law increases beyond the molecularity of the reaction. When the reaction is confined to a one-dimensional channel, $\gamma = 3.0$, or it can be as large as 50 when isolated on finely dispersed clusters or islands [9, 22]. The kinetic order is no longer equivalent to the molecularity of the reaction. The increase in kinetic order results in behavior with a higher effective cooperativity. The kinetic orders in some cases reflect the fractal dimension of the physical surface on which the reaction occurs.

This anomaly stems from the nonrandomness of the reactant distributions in low dimensions. Although in a classical reaction system the distribution of the reactants stays uniformly random, in a fractal-like reaction system the distribution tends to become "less random." Similar changes take place in other reactions and other spaces. Such findings are well established today, and they have been observed

experimentally and theoretically. Also, results from Monte Carlo simulations (a powerful tool in this field) are in very good agreement with these findings.

The solution of the differential equations above is a power function of time, namely $c(t) = \beta t^{\alpha}$ with parameters β and α satisfying the initial condition $c(t_0) = c_0$. Usually β and α are estimated by curve fitting on experimental data, and the parameters of (2.22) and (2.23) are obtained by

$$\kappa = -\alpha\beta^{1/\alpha} \text{ and } \gamma = 1 - 1/\alpha$$

and

$$k_{\circ} = -\alpha/\beta \text{ and } \lambda = 1 + \alpha,$$

respectively. Since we have assumed $\gamma \geq 2$ or $0 \leq \lambda < 1$, the parameter α satisfies $-1 \leq \alpha < 0$.

2.5.4 *Enzyme-Catalyzed Reactions*

In the same vein and under dimensionally restricted conditions, the description of the Michaelis–Menten mechanism can be governed by power-law kinetics with kinetic orders with respect to substrate and enzyme given by noninteger powers. Under quasi-steady-state conditions, Savageau [26] defined a fractal Michaelis constant and introduced the fractal rate law. The behavior of this fractal rate law is decidedly different from the traditional Michaelis–Menten rate law:

- the effective k_M decreases as the concentration of enzyme increases, and
- the kinetic order of the overall reaction with respect to total enzyme is greater than unity.

These properties are likely to have an important influence on the behavior of intact biochemical systems, e.g., within the living cell, enzymes do not function in dilute homogeneous conditions isolated from one another. The postulates of the Michaelis–Menten formalism are violated in these processes and other formalisms must be considered for the analysis of kinetics in situ. The intracellular environment is very heterogeneous indeed. Many enzymes are now known to be localized within two-dimensional membranes or quasi-one-dimensional channels, and studies of enzyme organization in situ [27] have shown that essentially all enzymes are found in highly organized states. The mechanisms are more complex, but they are still composed of elementary steps governed by fractal kinetics.

Power-law formalism was used by Savageau [28] to examine the implications of fractal kinetics in a simple pathway of reversible reactions. Starting with elementary chemical kinetics, that author proceeded to characterize the equilibrium behavior of a simple bimolecular reaction, then derived a generalized set of conditions for microscopic reversibility, and finally developed the fractal kinetic rate law

for a reversible Michaelis–Menten mechanism. By means of this fractal kinetic framework, the results showed that the equilibrium ratio is a function of the amount of material in a closed system, and that the principle of microscopic reversibility has a more general manifestation that imposes new constraints on the set of fractal kinetic orders. So, Savageau concluded that fractal kinetics provide a novel means to achieve important features of pathway design.

2.5.5 *Importance of the Power-Law Expressions*

Power-law expressions are found at all hierarchical levels of organization from the molecular level of elementary chemical reactions to the organismal level of growth and allometric morphogenesis. This recurrence of the power law at different levels of organization is reminiscent of fractal phenomena. In the case of fractal phenomena, it has been shown that this self-similar property is intimately associated with the power-law expression [29]. The reverse is also true; if a power function of time describes the observed kinetic data or if a reaction rate higher than 2 is revealed, the reaction takes place in fractal physical support.

The power-law formalism is a mathematical language or representation with a structure consisting of ordinary nonlinear differential equations whose elements are products of power-law functions. Power-law formalism meets two of the most important criteria for judging the appropriateness of a kinetic representation for complex biological systems: the degree to which the formalism is systematically structured, which is related to the issue of mathematical tractability, and the degree to which actual systems in nature conform to the formalism, which is related to the issue of accuracy.

2.6 Fractional Diffusion Equations

Before closing this chapter we would like to mention briefly a novel consideration of diffusion based on the recently developed concepts of fractional kinetics [30]. From our previous discussion it is apparent that if $d_s \leq 2$, diffusion is recurrent. This means that diffusion follows an anomalous pattern described by (2.10); the mean squared displacement grows as $\langle z^2(t) \rangle \propto t^{\gamma}$ with the exponent $\gamma \neq 1$. To deal with this, a consistent generalization of the diffusion equation (2.18) could have a fractional order temporal derivative such as

$$\frac{\partial^{\gamma} p(z,t)}{\partial t^{\gamma}} = \mathcal{D}_{\gamma} \frac{\partial^2 p(z,t)}{\partial z^2},$$

where $\mathcal{D}_{\gamma}$ is the fractional diffusion coefficient and the fractional order γ depends on d_w, the fractal dimension of the walk. The previous fractional diffusion equation

generalizes Fick's second law, and therefore it allows scientists to describe complex systems with anomalous behavior in much the same way as simpler systems [30].

Also, in order to appreciate the extent of spatial heterogeneity, Berding [31] introduced a heterogeneity function for reaction–diffusion systems evolving to spatially inhomogeneous steady-state conditions. The same author discusses particular applications and compares specific reaction–diffusion mechanisms with regard to their potential for heterogeneity.

Chapter 3
Nonlinear Dynamics

A wonderful harmony arises from joining together the seemingly unconnected.

Heraclitus of Ephesus (544-483 BC)

Series of measurements from many physiological processes appear random. On the other hand, we are used to thinking that the determinants of variability cannot be known because of the multiplicity and interconnectivity of the factors affecting the phenomena. This idea relies on the classical view of randomness, which requires that a complex process have a large (perhaps infinite) number of degrees of freedom that are not directly observed but whose presence is manifested through fluctuations. However, over the last two decades, scientists from various fields of research have shown that randomness generated by deterministic dynamic processes exhibits spectra practically indistinguishable from spectra of pure random processes. This is referred to as *chaotic behavior*, a specific subtype of nonlinear dynamics, which is the science dealing with the analysis of dynamic systems [32, 33].

The paradox with the term "chaos" is the contradiction between its meaning in colloquial use and its mathematical sense. Routinely, we use the word chaos in everyday life as a synonym for randomness having catastrophic implications. In mathematics, however, "chaos" refers to irregular behavior of a process that appears to be random, but is not. Accordingly, this apparent random-looking behavior poses a fundamental dilemma regarding the origin of randomness in a set of irregular observations from a dynamic process: Is the system chaotic or not? In other words, does the irregular behavior of the observations arise from noise or chaos?

Figure 3.1 illustrates the difference between random and chaotic systems:

- Subplot (A) shows a series of uniformly distributed random numbers between 0 and 1.
- In (B), the plot was generated by the logistic map, a deterministic model of the form $y_{i+1} = 4y_i\,(1 - y_i)$.

It is impossible to distinguish the two models visually. The subplots C and D are the so-called *pseudophase* plots of the two sequences of plots A and B, respectively: each y_i is plotted against its consequent y_{i+1}. The random sequence (A) produces

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_3

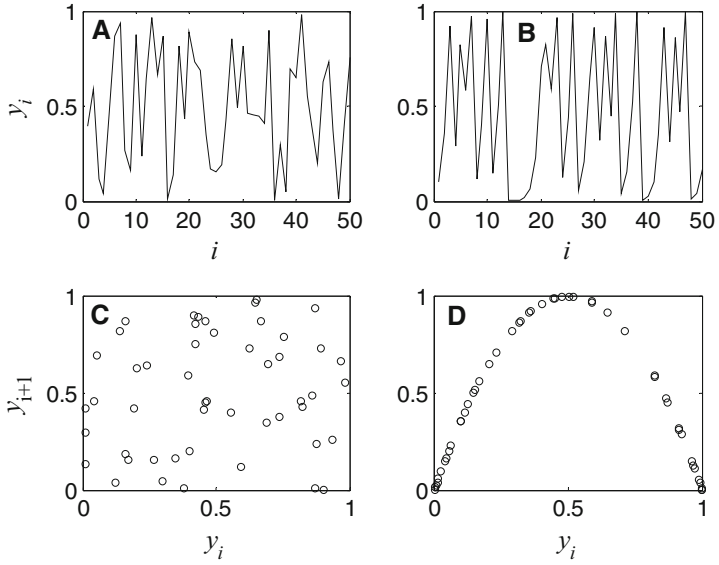


Fig. 3.1 The difference between random (**A**, **C**) and chaotic (**B**, **D**) processes pictured as a series of numbers (**A**, **B**) and as pseudophase plots (**C**, **D**)

scattered points (C) showing that there is no correlation between successive points. In contrast, the points of the deterministic sequence (B) lie in a well-formed line (D).

The key property in this complex, unpredictable, random-like behavior is nonlinearity. When a system (process, or model, or both) consists only of linear components, the response is proportional to its stimulus and the cumulative effect of two stimuli is equal to the summation of the individual effects of each stimulus. This is the superposition principle, which states that every linear system can be studied by breaking it down into its components (thus reducing complexity). In contrast, for nonlinear systems, the superposition principle does not hold; the overall behavior of the system is not at all the same as the summation of the individual behaviors of its components, making complex, unpredictable behavior a possibility. Nevertheless, not every nonlinear system is chaotic, which means that nonlinearity is a necessary but not a sufficient condition for chaos.

The basic ideas of chaos were introduced more than a hundred years ago; however, its significance and implications were realized relatively recently because chaos was studied in detail after the wide dissemination of computers in the 1970s. Although its study started from the fields of mathematics, astronomy, and physics, scientists from almost every field became interested in these ideas. The life sciences are good candidates for chaos due to the complexity of biological processes, although many consider the advanced mathematics and modeling techniques used a drawback. However, during the last 20 years the science of chaos has evolved into a truly interdisciplinary field of research that has changed the way scientists look at phenomena.

3.1 Dynamic Systems

A dynamic system is a deterministic system whose state is defined at any time by the values of several variables $\underline{y}(t)$, the so-called *states* of the system, and its evolution in time is determined by a set of rules. These rules, given a set of initial conditions $\underline{y}(0)$, determine the time evolution of the system in a unique way. This set of rules can be either

- differential equations of the form

$$\dot{\underline{y}}(t) = \underline{g}\left(\underline{y}, t, \theta\right),$$

  and the system is called a *flow*, or
- discrete equations in which every consequent generation of the variables $\underline{y}$ is given by an equation of the form

$$\underline{y}_{i+1} = \underline{g}\left(\underline{y}_i, \theta\right),$$

  where $\underline{y}_i$ stands for the i-th generation of the variable $\underline{y}$, and then the system is called a *map*.

In the above definitions, θ represents a set of parameters of the system, having constant values. These parameters are also called *control parameters*. The set of the system's variables forms a representation space called the *phase space* [33]. A point in the phase space represents a unique state of the dynamic system. Thus, the evolution of the system in time is represented by a curve in the phase space called *trajectory* or *orbit* for the flow or the map, respectively. The number of variables needed to describe the system's state, which is the number of initial conditions needed to determine a unique trajectory, is the dimension of the system. There are also dynamic systems that have infinite dimension. In these cases, the processes are usually described by differential equations with partial derivatives or time-delay differential equations, which can be considered as a set of infinite in number ordinary differential equations. The fundamental property of the phase space is that trajectories can never intersect themselves or each other. The phase space is a valuable tool in dynamic systems analysis since it is easier to analyze the properties of a dynamic system by determining topological properties of the phase space than by analyzing the time series of the values of the variables directly.

Stable limit sets in the phase space are of supreme importance in experimental and numerical settings because they are the only kind of limit set that can be observed naturally, that is, by simply letting the system run (cf. Appendix A).

3.2 Attractors

Dynamic systems are classified in two main categories: conservative and nonconservative systems. Conservative systems have the property of conserving the volume that is formed by an initial set of points in phase space as time goes by, although

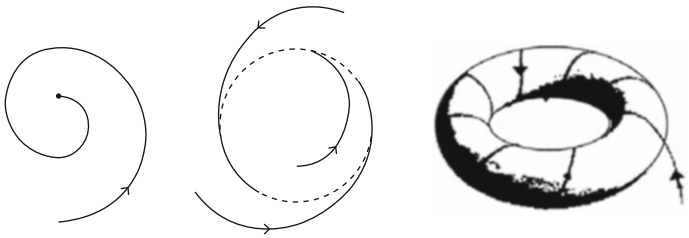


Fig. 3.2 A schematic representation of various types of attractors. Reprinted from [34] with permission from Springer

the shape of the volume may change. In other words, a volume in phase space resembles an incompressible liquid. On the other hand, nonconservative systems do not possess this property and an initial volume in phase space, apart from changing its shape, may also grow or shrink. In the latter case (when the volume shrinks) the system is called *dissipative*. Most processes in nature, including biological processes, are dissipative.

The trajectories of dissipative dynamic systems, in the long run, are confined in a subset of the phase space, which is called an *attractor* [33], i.e., the set of points in phase space where the trajectories converge. An attractor is usually an object of lower dimension than the entire phase space (a point, a circle, a torus, etc.). For example, a multidimensional phase space may have a point attractor (dimension 0), which means that the asymptotic behavior of the system is an equilibrium point, or a limit cycle (dimension 1), which corresponds to periodic behavior, i.e., an oscillation. Schematic representations for the point, the limit cycle, and the torus attractors are depicted in Figure 3.2. The point attractor is pictured on the left: regardless of the initial conditions, the system ends up in the same equilibrium point. In the middle, a limit cycle is shown: the system always ends up doing a specific oscillation. The torus attractor on the right is the two-dimensional equivalent of a circle. In fact, a circle can be called a one-torus, the two-dimensional torus can be called a two-torus, and there is also the three-torus and generally the m-torus. The trajectory on a two-torus is a two-dimensional oscillation with the ratio of the frequencies of the two oscillations being irrational. Because the trajectory never passes through the same point twice, in infinite time it fills the entire surface of the torus. This type of trajectory is called *quasi-periodic*. Being an attractor, the torus attracts all trajectories to fall on its surface.

Even the states of systems with infinite dimension, like systems described by partial differential equations, may lie on attractors of low dimension. The phase space of a system may also have more than one attractor. In this case the asymptotic behavior, i.e., the attractor where a trajectory ends up, depends on the initial conditions. Thus, each attractor is surrounded by an attraction basin, which is the part of the phase space where the trajectories from all initial conditions end up.

3.3 Bifurcation

A dynamic system may exhibit qualitatively different behavior for different values of its control parameters θ. Thus, a system that has a point attractor for some value of a parameter may oscillate (limit cycle) for some other value. The *critical point* where the behavior changes is called a *bifurcation point*, and the event a *bifurcation* [33]. More specifically, this kind of bifurcation, i.e., the transition from a point attractor to a limit cycle, is referred to as *Hopf bifurcation*.

Consider the one-dimensional map

$$y_{i+1} = g\left(y_i, \theta\right) = \theta y_i \left(1 - y_i\right). \tag{3.1}$$

This difference equation is called a *logistic map*, and represents a simple deterministic system, where given a y_i one can calculate the consequent point y_{i+1} and so on. We are interested in solutions $y_i \geq 0$ with $\theta > 0$. This model describes the dynamics of a single species population [33]. For this map, the fixed points y^* on the first iteration are solutions of

$$y_1^* = \theta y_1^* \left(1 - y_1^*\right),$$

namely

$$y_{1A}^* = 0 \; y_{1B}^* = \left(\theta - 1\right)/\theta,$$

with the corresponding characteristic multipliers (cf. Appendix A)

$$\xi_{1A} = \theta \; \xi_{1B} = 2 - \theta.$$

As θ increases from zero but with $0 < \theta < 1$, the only realistic fixed point that is nonnegative is y_{1A}^*, which is stable since $0 < \xi_{1A} < 1$. The first bifurcation of y_{1A}^* is observed for $\theta = 1$. When $1 < \theta < 3$, on the one hand, the fixed point y_{1A}^* becomes unstable since $\xi_{1A} > 1$, and on the other hand, the positive fixed point y_{1B}^* is stable since $-1 < \xi_{1B} < 1$. Although there are two steady states, for any initial condition different from $y = 0$, the system will end up after a few steps in y_{1B}^* (Figure 3.3A, fixed point of period 1 for $\theta = 2.7$). The second bifurcation comes at y_{1B}^* at $\theta = 3$ where $\xi_{1B} = -1$, and so locally, near y_{1B}^*, we have a periodic solution.

To see what is happening when θ passes through the bifurcation value $\theta = 3$, we examine the stability at the second iteration. The second iteration can be thought of as a first iteration in a model where the iterative time step is 2. The fixed points are solutions of

$$y_2^* = \theta^2 y_2^* \left(1 - y_2^*\right)\left[1 - \theta y_2^* \left(1 - y_2^*\right)\right].$$

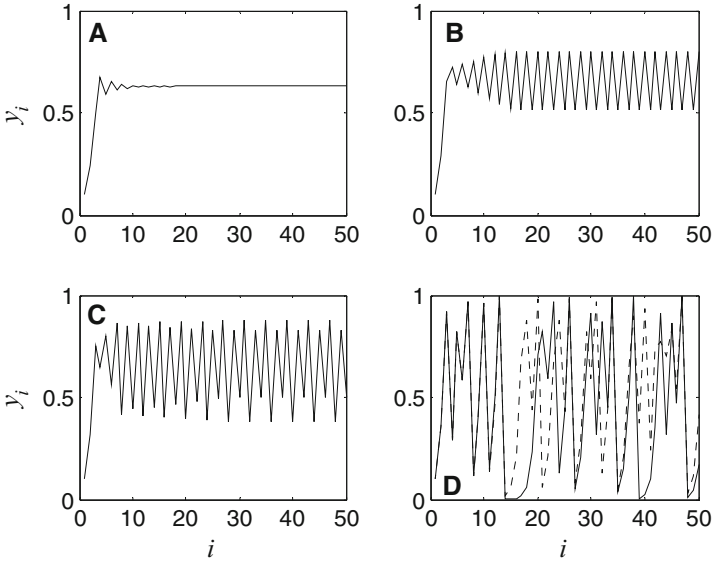


Fig. 3.3 The logistic map, for various values of the parameter θ. (**A**) $\theta = 2.7$, (**B**) $\theta = 3.2$, (**C**) $\theta = 3.5$, (**D**) two chaotic trajectories for $\theta = 4$ are coplotted. The initial condition for all *solid line* plots (**A** to **D**) is $y_0 = 0.1$

This equation leads to the following solutions:

$$y^*_{2A} = 0,\ y^*_{2B} = \tfrac{\theta-1}{\theta},\ y^*_{2C} = \tfrac{\theta+1-\sqrt{\theta^2-2\theta-3}}{2\theta},\ y^*_{2D} = \tfrac{\theta+1+\sqrt{\theta^2-2\theta-3}}{2\theta},$$

when $3 < \theta < 1 + \sqrt{6}$. The corresponding characteristic multipliers are

$$\xi_{2A} = \theta^2,\ \xi_{2B} = (2-\theta)^2,\ \xi_{2C} = \xi_{2D} = -\theta^2 + 2\theta + 4.$$

Hence, $\xi_{2A} > 1$, $\xi_{2B} > 1$, $-1 < \xi_{2C} < 1$, and $-1 < \xi_{2D} < 1$. Thus, the y^*_{2C} and y^*_{2D} of the second iteration are stable. What this means is that there is a stable equilibrium of the second iteration, i.e., if we start at y^*_{2C} or y^*_{2D}, for example, we come back to it after 2 iterations. What happens now is that for any initial condition, except $y = 0$ and $y = (\theta - 1)/\theta$, the system after a few steps will end up forming a never-ending succession of the two values of y^*_{2C} and y^*_{2D} (Figure 3.3B, fixed points of period 2 for $\theta = 3.2$).

As θ continues to increase ($1 + \sqrt{6} < \theta$), the characteristic multipliers ξ_{2C} and ξ_{2D} pass through $\xi = -1$, and so these 2-period solutions become unstable. At this stage, we look at the fourth iterate and we find, as might now be expected, that a 4-cycle periodic solution appears (Figure 3.3C, fixed point of period 4 for $\theta = 3.5$). The period doubles repeatedly and goes to infinity as one approaches a critical point θ_c at which instability sets in for all periodic solutions, e.g., for the model (3.1), $\theta_c \approx 3.5699456$. Above θ_c all fixed points are unstable and the system is chaotic.

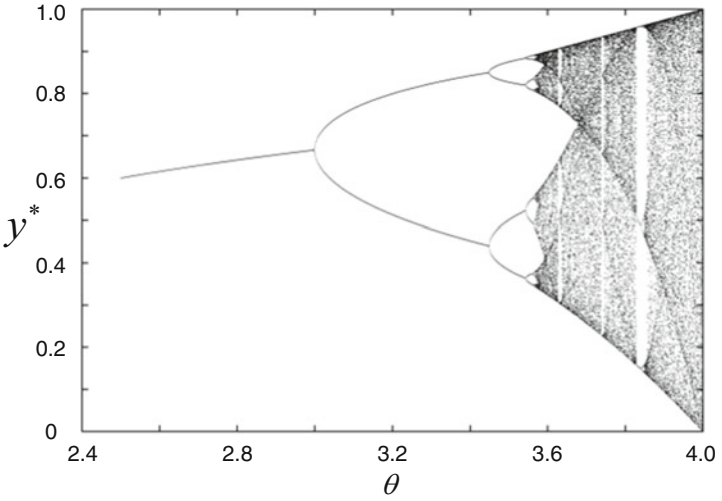


Fig. 3.4 The bifurcation diagram of the logistic map

The bifurcation situation is illustrated in Figure 3.4, where the stable fixed points y^* are plotted as a function of the parameter θ. These bifurcations are called *pitchfork bifurcations*, for obvious reasons from the picture they generate in Figure 3.4. For example, if $3 < \theta < 1 + \sqrt{6}$, then the periodic solution is between the two y^* that are the intersections of the vertical line through the θ value and the curve of equilibrium points. From Figure 3.4, we note that the difference between the values of θ at which two successive bifurcations take place decreases. It was actually found that the ratio of two successive intervals of θ between successive bifurcations is universally constant, namely $\delta = 4.66920161$, not only for this specific system, but for all systems of this kind, and it is referred to as the *Feigenbaum constant* [33]. Although we have concentrated here on the logistic map, this kind of behavior is typical of maps with dynamics like (3.1); that is, they all exhibit bifurcations to higher periodic solutions eventually leading to chaos.

So, apart from the regular behavior, which is either steady-state, periodic, or quasi-periodic behavior (trajectory on a torus, Figure 3.2), some dynamic systems exhibit chaotic behavior, i.e., trajectories follow complicated aperiodic patterns that resemble randomness. Necessary but not sufficient conditions in order for chaotic behavior to take place in a system described by differential equations are that it must have dimension at least 3, and it must contain nonlinear terms. However, a system of three nonlinear differential equations need not exhibit chaotic behavior. This kind of behavior may not take place at all, and when it does, it usually occurs only for a specific range of the system's control parameters θ.

3.4 Sensitivity to Initial Conditions

As pointed out, for $\theta > \theta_c$ there exist infinitely many unstable steady states of period 1, 2, 4, 8, ... and no stable steady states. This means that almost any initial condition leads to an aperiodic trajectory that looks random as in Figure 3.3D, but actually the behavior is chaotic. In this figure, two chaotic orbits for $\theta = 4$ are coplotted. Only the initial conditions of the two trajectories differ slightly. For the solid line the initial condition is $y = 0.1$, whereas for the dashed line it is $y = 0.10001$. Although the difference is extremely small, the effect is not at all negligible. The orbits follow an indistinguishable route only for the first 10 steps. Thereafter, they deviate dramatically. Thus, sensitivity to the initial conditions, together with its main consequence of long-term unpredictability, is exhibited.

Hence, the main characteristic of chaotic behavior is the sensitivity to initial conditions. This means that nearby trajectories, whose initial conditions are only slightly different, follow completely different evolutions in time. This property has the implication of unpredictability of the time evolution of the system in the long run due to our inability to know the initial conditions with infinite accuracy. The deviation of two initially neighboring trajectories increases exponentially with time, i.e., proportional to $\exp(\lambda t)$, where the exponent λ is called the *Lyapunov exponent* [33, 35]. Lyapunov exponents are a generalization of the eigenvalues at an equilibrium point and of characteristic multipliers. They depend on the initial conditions and they can be used to determine the stability of quasi-periodic and chaotic behavior as well as of equilibrium points and periodic solutions. For a flow, the Lyapunov exponents are equal to the real parts of the eigenvalues at the equilibrium point, and for a map, they are equal to the magnitudes of the characteristic multipliers at the fixed point. A dynamic system has the same number of Lyapunov exponents as its dimension. The Lyapunov exponents express the deviation of initially nearby trajectories in each "direction." So, a Lyapunov exponent may be negative for a stable "direction," which expresses the exponential approach of two nearby trajectories, and positive for exponential deviation, which expresses the divergence of two nearby trajectories. A system of high dimension may have Lyapunov exponents of all signs and is considered chaotic if at least one of them is positive, which states that at least in one "direction" there exists sensitivity to the initial conditions.

Because chaotic systems may have both negative and positive Lyapunov exponents, their asymptotic behavior can be limited in an attractor as well, where the negative exponents express the convergence to the attractor and the positive the exponential divergence (chaotic behavior) within the attractor. These chaotic attractors are not elementary topological entities with integer dimensions like a point, a circle, or a torus. Instead they have a fractal dimension, which defines an extremely complicated object of infinite detail, though confined in a finite space. This kind of attractor is called a *strange attractor* [33], and the integer dimension of the entire phase space in which the attractor lives is called the *embedding dimension*

of the attractor. The two concepts, exponential divergence of initially neighboring trajectories and confinement in a compact space, appear contradictory. However, the fractal structure of the strange attractor makes their coexistence feasible.

3.5 Reconstruction of the Phase Space

The concepts of nonlinear dynamics do not apply only to abstract mathematical models that are described by maps or flows. Useful results can be obtained from observations gathered from real processes as well. Real-life observations, like biological signals, are usually time series of measured quantities. Instead of studying a time series statistically, the idea is to consider it as if it came out of a dynamic system. Then, one tries to reconstruct its phase space (pseudophase space in the case of observed data, when the state variables are unknown) and see whether any structure is detectable, either visually or using certain mathematical and numerical tools [36–38]. The absence of any structure in phase space (e.g., a scatter of points) means that the system is random (Figure 3.1C). However, the presence of structure is evidence of the dynamic origin of the time series and the existence of an attractor (Figure 3.1D). The dimension of the attractor can give us information about the dynamic behavior of the whole system. If, for example, the dimension of the attractor is not an integer, it corresponds to a strange attractor and the system exhibits chaotic behavior. The embedding dimension of the attractor, which is actually the dimension of the reconstructed phase space and in the case of a strange attractor should be the next greater integer of the fractal dimension, gives the least number of independent variables, or states, needed to describe the system.

The phase space reconstruction of a time series is accomplished by the method of delays. An embedding dimension d_e is chosen, plus a time delay t°, and then the phase space is constructed using as variables $y(t)$, $y(t + t^\circ)$, $\ldots$, $y(t + (d_e - 1)\, t^\circ)$, for all t. It is evident that the choice of d_e and t° is crucial for the reconstruction. There are certain theorems and tests that help in the proper choice of these parameters, but experience and trial are also valuable tools. It must be mentioned though that due to the automated character of the algorithms, the danger of misleading results always exists. During the past years an overuse of these techniques was noticed and many of the results obtained by this rationale were either wrong or led to erroneous conclusions due to poor application of the techniques and algorithms.

Example 1. The Rössler Strange Attractor

Figure 3.5 illustrates the model of the *Rössler strange attractor* [33]. The set of nonlinear differential equations is

$$\begin{aligned} \dot{y}_1 &= -y_2 - y_3, & y_1(0) &= 3, \\ \dot{y}_2 &= y_1 + 0.2y_2, & y_2(0) &= 3, \\ \dot{y}_3 &= 0.4 + y_1 y_3 - 5.7y_3, & y_3(0) &= 0. \end{aligned}$$

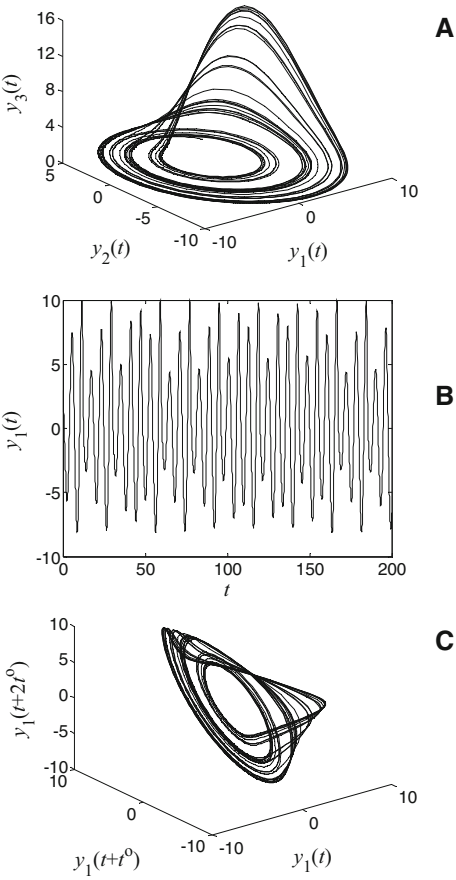


Fig. 3.5 The Rössler strange attractor. (**A**) The phase space. (**B**) The state variable $y_1\,(t)$. (**C**) Reconstruction in the pseudophase space

The single trajectory plotted in the three-dimensional phase space never passes through the same point a second time, yet it never leaves a compact volume, thus forming a fractal object of infinite detail (fractal dimension ≈ 2.07), Figure 3.5A. The state variable y_1 plotted in Figure 3.5B as a function of time exhibits obvious aperiodicity. In Figure 3.5C, the Rössler attractor is reconstructed in pseudophase space with the method of delays, making use only of the data from the y_1 variable, as if y_1 were an observable quantity and nothing more of the underlying dynamics were known. Of course, here the dimension of the system is also known and one does not have to try other values for the dimension. Every value of $y_1(t)$ is plotted against $y_1(t + t^\circ)$ and $y_1(t + 2t^\circ)$ with lag time $t^\circ = 1$. The reconstructed phase space is not identical to the original one; however, the main topology and features are depicted adequately. ■

3.6 Estimation and Control in Chaotic Systems

Parameter estimation is a key factor in modeling. We usually fit the established model to experimental data in order to estimate model parameters both for simulation and control. However, a task so common in a classical system is quite difficult in a chaotic one. The sensitivity of the system's behavior with respect to the initial conditions and the control parameters makes it very hard to assess the parameters using tools such as least squares fitting. Thus, efforts have been made to deal with this problem [39]. For nonlinear data analysis, a combination of statistical and mathematical tests on the data to discern inner relationships among the data points (determinism vs. randomness), periodicity, quasiperiodicity, and chaos is used. These tests are in fact nonparametric indices. They do not reveal functional relationships, but rather directly calculate process features from time series records. For example, the dimensionality of a time series, which results from the phase space reconstruction procedure, and the Lyapunov exponent are such nonparametric indices. Some others are also commonly used:

- *Correlation dimension*. The correlation dimension is calculated by measuring the Hausdorff dimension according to the Grassberger method [37, 40]. The dimension of the system relates to the fewest number of independent variables necessary to specify a point in the state space [41]. With random data, the dimension increases with an increase of the embedding space. In deterministic data sets the dimension levels off, even though the presence of random noise may yield a slow rise.
- *Singular value decomposition* and eigenvalues of the singular value matrix phase plots. By applying singular value decomposition to the embedded matrix one can improve the appearance of the trajectories in phase space by separating out the noise and the different frequencies from each other, which is important when one is working with experimental data [38, 42]. The eigenvalues give a strong indication of the dimension of the system. A random system shows no demarcation of values, whereas a deterministic system does, as the embedding dimension increases. Each column of data is equivalent to an independent variable; by plotting one column vector vs. another, one can construct the phase space and observe the flows with arrows indicating the direction [43].

The above indices contrast with those destined for linear data analysis:

- The *autocorrelation* (or correlation) function is obtained by multiplying each $y(t)$ by $y(t - t^\circ)$, where t° is a time delay, and summing the products over all points [44]. Examination of the sum plotted as a function of t° reveals the level of dependency of data points on their neighbors. The correlation time is the value of t° for which the value of the correlation function falls to $\exp(-1)$. When the correlation function falls abruptly to zero, that indicates that the data are without a deterministic component; a slow fall to zero is a sign of stochastic or deterministic behavior; when the data slowly drop to zero and show periodic

behavior, then the data are highly correlated and are either periodic or chaotic in nature [38, 44].
- Following a fast Fourier transform of the data, the *power spectrum* shows the power (the Fourier transform squared) as a function of frequency. Random and chaotic data sets fail to demonstrate a dominant frequency. Periodic or quasi-periodic data sets will show one or more dominant frequencies [38].

Chaotic systems are characterized by extreme sensitivity to tiny perturbations. This phenomenon is also known as the *butterfly effect*. This famous term was coined by Lorenz [45], who noticed that long-term prediction of the weather using his system of differential equations was impossible. Lorenz observed that tiny differences in the initial conditions start to grow at a greater and greater speed, until the predictions become nonsense. In an analogous manner, the flapping of a single butterfly's wing today will produce a tiny change in the state of the atmosphere, which in the long run will diverge from that which would otherwise exist in the unperturbed state.

The butterfly effect is often regarded as a troublesome property, and for many years it was generally believed that chaotic motions are neither predictable nor controllable. Von Neumann around 1950 first reported a differing view that small, carefully chosen, preplanned atmospheric disturbances could lead after some time to desired large-scale changes in the weather. Using this chaotic sensitivity, recent work demonstrates that the butterfly effect permits the use of tiny feedback perturbations to control trajectories in chaotic systems, a capability without counterpart in nonchaotic systems [46]. Indeed, it is possible to accomplish this only because the chaotic systems are characterized by exponential growth of small disturbances. This exponential growth implies that we can reach any accessible target extremely quickly, using only a small perturbation.

The relevant research fits broadly into two categories [47]. First, one may ask to select a desired behavior among an infinite variety of behaviors naturally present in chaotic systems, and then stabilize this behavior by applying only tiny changes to an accessible system parameter. Second, one can use the sensitivity of chaotic systems to direct trajectories rapidly to a desired state and steer the system to a general target in state space (not necessarily a periodic orbit). This means that chaotic systems can achieve great flexibility in their ultimate performance.

The presence of chaos may be a great advantage for control in a variety of situations. Typically, in a nonchaotic system, small controls can only change the system dynamics slightly. Short of applying large controls or greatly modifying the system, we are stuck with whatever system performance already exists. In a chaotic system, on the other hand, we are free to choose among a rich variety of dynamic behaviors. Thus, we anticipate that it may be advantageous to design chaos into systems, allowing such variety without requiring large controls or the design of separate systems for each desired behavior.

3.7 Physiological Systems

The application of nonlinear dynamics in physiological systems offers a new basis in the way certain pathological phenomena emerge. The main characteristic is that a pathological symptom is considered as a sudden qualitative change in the temporal pattern of an illness, such as when a bifurcation takes place. This change can be caused either by endogenous factors or by an exterior stimulus that changes one or more critical control parameters. According to this rationale, therapeutic strategies should aim to invert the progress of the disease and restore normal physiological conditions by interfering with the control parameters. This is in contrast to the classical approach, which focuses on eliminating the symptoms with a linear rationale that relates the therapeutic stimulus to the effect through a proportionality. This is a general concept also referred to as *dynamical disease*, a term introduced by Mackey and Glass [32, 48–50] (cf. also Section 13.1.3). It is widely appreciated that chaotic behavior dominates physiological systems. Moreover, periodic or other nonchaotic states are considered pathological, whereas chaotic behavior is considered to be the normal, healthy state. The reason for this should be associated with a fundamental advantage of nonlinear over classical systems. Indeed, small variations of the control parameters may offer finer, more rapid, and more energy-efficient controllability of the system compared to linear systems [51]. This may explain why nature prefers chaos to regularity, and of course the latter is a good enough reason for applied biological sciences such as biopharmaceutics, pharmacokinetics, and pharmacodynamics to adopt this rationale to a greater extent.

Part II
Modeling in Biopharmaceutics

Most drugs on the market today are taken orally. Provided a drug is well absorbed, this route of administration results in effective therapy with a minimum of inconvenience to the patient. Therefore, much effort in pharmaceutical research has been devoted to developing oral dosage forms that can deliver the drug to the systemic circulation in a timely and efficient manner. *Biopharmaceutics* refers to the study of the effect of formulation characteristics, the physicochemical properties of the drug, and the physiology of drug absorption. The main object of this discipline is to close the gap between the in vitro data collected in the laboratory and the in vivo performance of the pharmaceutical formulation. Since drug molecules can penetrate the epithelial barrier of the gastrointestinal tract only if they are in solution, drug dissolution and release are the most important biopharmaceutical processes for oral drug delivery. In parallel, the physiological conditions prevailing in the gastrointestinal tract play a significant role and should be taken into account when drug absorption is studied.

Most research on gastrointestinal absorption is based on the concept of homogeneity, that is, the description of average behavior. Some of the most often used paradigms are those borrowed from chemical engineering literature to model hydrodynamics, permeability, and absorption. For example, in the field of dissolution testing, a well-stirred (homogeneous) dissolution medium is used to mimic the in vivo conditions [52]. Calculations associated with the effective intestinal permeability or the unstirred water layer thickness in permeability studies assume that the hydrodynamics of the solution in the intestinal segment obey the well-stirred model [53]. The tank and tube models, often used for the analysis of drug dissolution and uptake in the gastrointestinal tract [54–56], are accompanied with the assumptions of perfect mixing and homogeneous flow, respectively.

One can argue, however, that the assumptions of homogeneity and well-stirred media are not only obvious, but that they are also in fact contrary to the evidence, given the anatomical and physiological complexity of the gastrointestinal tract. In vivo drug dissolution, release, and uptake are heterogeneous processes since they take place at interfaces of different phases, e.g., liquid–solid and liquid–membrane boundaries, while diffusion, which is the principal mechanism of all

processes, operates under topological constraints. In addition, all processes occur in heterogeneous environments, e.g., variable stirring conditions in the lumen.

Therefore, a proper analysis of drug dissolution, release, and uptake should take into account the heterogeneous character of these processes. In this second part of the book, we consider these important biopharmaceutical processes, placing emphasis on their heterogeneous features.

Chapter 4
Drug Release

An equation relating the rate of release of solid drugs suspended in ointment bases into perfect sinks is derived. . . . The amount of drug released . . . is proportional to the square root of time.

Takeru Higuchi
School of Pharmacy, University of Wisconsin, Madison
Journal of Pharmaceutical Sciences 50:874–875 (1961)

The term "release" encompasses several processes that contribute to the transfer of drug from the dosage form to the bathing solution (e.g., gastrointestinal fluids, dissolution medium). The objective of this chapter is to present the spectrum of mathematical models that have been developed to describe drug release from controlled-release dosage forms. These devices are designed to deliver the drug at a rate that is governed more by the dosage form and less by drug properties and conditions prevailing in the surrounding environment. The release mechanism is an important factor in determining whether both of these objectives can be achieved. Depending on the release mechanism, the controlled-release systems can be classified into

1. diffusion-controlled,
2. chemically controlled, and
3. swelling-controlled.

By far, diffusion is the principal release mechanism, since apart from the diffusion-controlled systems, diffusion takes place at varying degrees in both chemically and swelling-controlled systems. The mathematical modeling of release from diffusion-controlled systems relies on the fundamental Fick's law (2.11), (2.16) with either concentration-independent or concentration-dependent diffusion coefficients. Depending on the formulation characteristics of the device, various types of diffusion can be conceived, i.e., diffusion through an inert matrix, a hydrogel, or a membrane. For chemically controlled systems, the rate of drug release is controlled by

- the degradation and in some cases the dissolution of the polymer in erodible systems or

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_4

- the rate of the hydrolytic or enzymatic cleavage of the drug–polymer chemical bond in pendant chain systems.

For swelling-controlled systems the swelling of the polymer matrix after the inward flux of the liquid bathing the system induces the diffusion of drug molecules toward the bathing solution.

In the following sections of this chapter we present the mathematical models used to describe drug release from hydroxypropyl methylcellulose (HPMC) controlled-release dosage forms. HPMC is the most widely used hydrophilic polymer for oral drug delivery systems. Since HPMC exhibits high swellability, drug release from HPMC-based systems is the result of different simultaneously operating phenomena. In addition, different types of HPMC are commercially available and therefore a universal pattern of drug release from HPMC-based systems cannot be pointed out. Accordingly, a wide spectrum of models has been used to describe drug release kinetics from HPMC-based matrix tablets. The sequential presentation below of the mathematical models presented attempts to provide hints to their interrelationships, along with their time evolution, and avoids a strict classification, e.g., empirical vs. mechanistically based models. The last part of the chapter is devoted to the rapidly emerging applications of Monte Carlo simulation in drug release studies. Finally, a brief mention of applications of nonlinear dynamics to drug release phenomena is made at the end of the chapter.

4.1 The Higuchi Model

In 1961 Higuchi [57] analyzed the kinetics of drug release from an ointment assuming that the drug is homogeneously dispersed in the planar matrix and the medium into which drug is released acts as a perfect sink, Figure 4.1. Under these pseudo-steady-state conditions, Higuchi derived (4.1) for the cumulative amount $q(t)$ of drug released at time t:

$$q(t) = \mathcal{A}\sqrt{\mathcal{D}(2c_0 - c_s)c_s t}, \quad c_0 > c_s, \tag{4.1}$$

where $\mathcal{A}$ is the surface area of the ointment exposed to the absorbing surface, $\mathcal{D}$ is the diffusion coefficient of drug in the matrix medium, and c_0 and c_s are the initial drug concentration and the solubility of the drug in the matrix, respectively. Although a planar matrix system was postulated in the original analysis [57], modified forms of (4.1) were published [58–60] for different geometries and matrix characteristics, e.g., granular matrices.

Equation (4.1) is frequently written in simplified form:

$$\frac{q(t)}{q_\infty} = k\sqrt{t}, \tag{4.2}$$

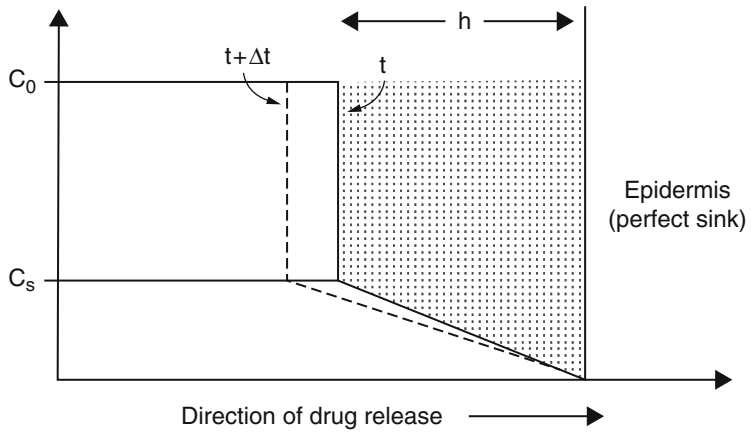


Fig. 4.1 The spatial concentration profile of drug (*solid line*) existing in the ointment containing the suspended drug in contact with a perfect sink according to Higuchi's assumptions. The broken line indicates the temporal evolution of the profile, i.e., a snapshot after a time interval Δt. For the distance h above the exposed area, the concentration gradient $(c_0 - c_s)$ is considered constant assuming that c_0 is much higher than c_s

where q_∞ is the cumulative amount of drug released at infinite time and k is a composite constant with dimension time$^{-1/2}$ related to the drug diffusional properties in the matrix as well as the design characteristics of the system. For a detailed discussion of the assumptions of the Higuchi derivation in relation to a valid application of (4.2) to real data, the reader can refer to the review of Siepmann and Peppas [61].

Equation (4.2) reveals that the fraction of drug released is linearly related to the square root of time. However, (4.2) cannot be applied throughout the release process since the assumptions used for its derivation are not obviously valid for the entire release course. Additional theoretical evidence for the time limitations in the applicability of (4.2) has been obtained [62] from an exact solution of Fick's second law of diffusion for thin films of thickness δ under perfect sink conditions, uniform initial drug concentration with $c_0 > c_s$, and assuming constant diffusion coefficient of drug $\mathcal{D}$ in the polymeric film. In fact, the short-time approximation of the exact solution is

$$\frac{q(t)}{q_\infty} = 4\sqrt{\frac{\mathcal{D}t}{\pi\delta^2}} = k'\sqrt{t}, \tag{4.3}$$

where $k' = 4\sqrt{\mathcal{D}/\pi\delta^2}$. Again, the proportionality between the fraction of drug released and the square root of time is justified (4.3). These observations have led to a rule of thumb, which states that the use of (4.2) for the analysis of release data is recommended only for the first 60% of the release curve $(q(t)/q_\infty \leq 0.60)$. This arbitrary recommendation does not rely on strict theoretical and experimental

findings and is based only on the fact that completely different physical conditions have been postulated for the derivation of the equivalent (4.2) and (4.3), while the underlying mechanism in both situations is classical diffusion. In this context, a linear plot of the cumulative amount of drug released $q(t)$ or the fraction of drug released $q(t)/q_\infty$ (utilizing data up to 60% of the release curve) vs. the square root of time is routinely used in the literature as an indicator for diffusion-controlled drug release from a plethora of delivery systems.

In 2011, an issue of the International Journal of Pharmaceutics (Vol. 418, No. 1, pp. 1–148, 10 October 2011) entitled "Mathematical modeling of drug delivery systems: Fifty years after Takeru Higuchi's models" was published commemorating the 50-th anniversary of the Higuchi's publication [57].

4.2 Systems with Different Geometries

One of the first physicochemical studies [63] dealing with diffusion in glassy polymers published in 1968 can be considered as the initiator of the realization that two apparently independent mechanisms of transport, a Fickian diffusion and a Case II transport, contribute in most cases to the overall drug release. Fick's law governs the former mechanism, while the latter reflects the influence of polymer relaxation on the molecules' movement in the matrix [64]. The first studies on this topic [65, 66] were focused on the analysis of Fickian and non-Fickian diffusion as well as the coupling of relaxation and diffusion in glassy polymers. The models used to describe drug release from different geometries are quoted below:

1. Fickian diffusional release from a thin polymer film. Equation (4.3) gives the short-time approximation of the fractional drug released from a thin film of thickness δ.
2. Case II release from a thin polymer film. The fractional drug release $q(t)/q_\infty$ follows zero-order kinetics [67, 68] according to

$$\frac{q(t)}{q_\infty} = \frac{2k_0}{c_0\delta}t, \tag{4.4}$$

   where k_0 is the Case II relaxation constant and c_0 is the drug concentration, which is considered uniform.
3. Case II radial release from a cylinder. The following equation describes the fractional drug released, $q(t)/q_\infty$, when Case II drug transport with radial release from a cylinder of radius ρ is considered [68]:

$$\frac{q(t)}{q_\infty} = \frac{2k_0}{c_0\rho}t - \left(\frac{k_0}{c_0\rho}t\right)^2. \tag{4.5}$$

4. Case II one-dimensional radial release from a sphere. For a sphere of radius ρ with Case II one-dimensional radial release, the fractional drug released, $q(t)/q_\infty$, is given [68] by

$$\frac{q(t)}{q_\infty} = \frac{3k_0}{c_0\rho}t - 3\left(\frac{k_0}{c_0\rho}t\right)^2 + \left(\frac{k_0}{c_0\rho}t\right)^3. \tag{4.6}$$

5. Case II radial and axial release from a cylinder. We quote below a detailed analysis of Case II radial and axial release from a cylinder [69] since (4.4) and (4.5) are special cases of the general equation derived in this section.

The analysis of Case II drug transport with axial and radial release from the cylinder depicted in Figure 4.2 is based on two assumptions:

- a boundary is formed between the glassy and rubbery phases of the polymer, and
- the movement of this boundary takes place under constant velocity.

First, the release surface is determined. A cylinder of height $2L$ that is allowed to release from all sides can be treated as a cylinder of height L that can release from the round side and the top only, Figure 4.2. This second case is easier to analyze and is also implied in [68] for the release of drug from a thin film of thickness $L'/2$. If the big cylinder of Figure 4.2 is cut in half across the horizontal line, two equal cylinders, each of height L, are formed. If drug release from the two newly formed areas (top and bottom) of the two small cylinders is not considered, the two cylinders of height L' exhibit the same release behavior as the big cylinder, i.e., $q(t)_{2L} = 2q(t)_L$ and $q_{\infty,2L} = 2q_{\infty,L}$; consequently,

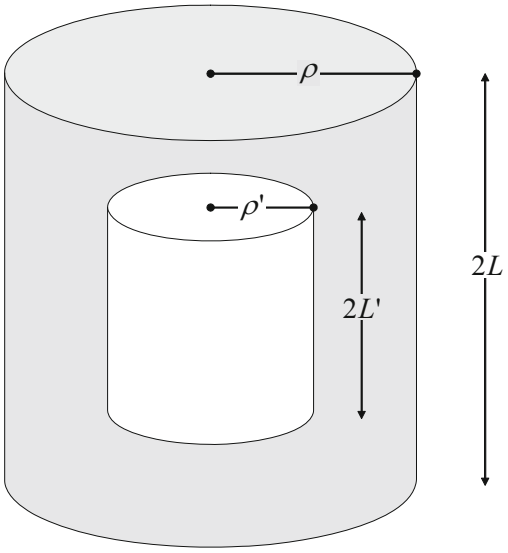


Fig. 4.2 Case II drug transport with axial and radial release from a cylinder of height $2L$ and radius ρ at $t = 0$. Drug release takes place from all sides of the big cylinder. The drug mass is contained in the gray region. After time t the height of the cylinder is reduced to $2L'$ and its radius to ρ' (*small cylinder*)

$$\frac{q(t)_{2L}}{q_{\infty,2L}} = \frac{q(t)_L}{q_{\infty,L}}.$$

This proportionality demonstrates that the analysis of the release results can describe both of the following cases: either a cylinder of height L that releases from the round and top surfaces or a cylinder of height $2L$ that releases from all sides, Figure 4.2.

At zero time, the height and radius of the cylinder are L and ρ, respectively, Figure 4.2. After time t the height of the cylinder decreases to L' and its radius to ρ' assuming Case II drug transport for both axial and radial release. The decrease rate of radius ρ' and height L' of the cylinder can be written

$$\dot{\rho}' = \dot{L}' = -\frac{k_0}{c_0}, \tag{4.7}$$

where k_0 is the Case II relaxation constant and c_0 is the drug concentration (considered uniform). The assumed value of the penetration layer speed is implied from the analysis of the cases studied in [67, 68], which are simpler than the present case. Initial conditions for the above equations are simply $\rho'(0) = \rho$ and $L'(0) = L$.

After integration of (4.7), we obtain the following equations as well as the time for which each one is operating:

$$\begin{aligned} \rho' &= \rho - (k_0/c_0)\, t, \quad t \le (c_0/k_0)\, \rho, \\ L' &= L - (k_0/c_0)\, t, \quad t \le (c_0/k_0)\, L. \end{aligned} \tag{4.8}$$

This means that the smaller dimension of the cylinder (ρ or L) determines the duration of the phenomenon.

The amount of drug released at any time t is given by the following mass balance equation:

$$q(t) = c_0 \pi \left(\rho^2 L - \rho'^2 L'\right). \tag{4.9}$$

Substituting (4.8) into (4.9), the following expression for mass $q(t)$ as a function of time t is obtained:

$$q(t) = c_0 \pi \left[\rho^2 L - \left(\rho - \frac{k_0}{c_0} t\right)^2 \left(L - \frac{k_0}{c_0} t\right)\right].$$

And for the mass released at infinite time, we can write

$$q_\infty = c_0 \pi \rho^2 L.$$

From the previous equations, the fraction released $q(t)/q_\infty$ as a function of time t is obtained:

$$\frac{q(t)}{q_\infty} = \left(\frac{2k_0}{c_0\rho} + \frac{k_0}{c_0 L}\right)t - \left(\frac{k_0^2}{c_0^2\rho^2} + \frac{2k_0^2}{c_0^2\rho L}\right)t^2 + \frac{k_0^3}{c_0^3\rho^2 L}t^3. \tag{4.10}$$

This equation describes the entire fractional release curve for Case II drug transport with axial and radial release from a cylinder. Again, (4.10) indicates that the smaller dimension of the cylinder (ρ or L) determines the total duration of the phenomenon. When $\rho \gg L$, (4.10) can be approximated by

$$\frac{q(t)}{q_\infty} = \frac{k_0}{c_0 L}t,$$

which is identical to (4.4) with the difference of a factor of 2 due to the fact that the height of the cylinder is $2L$. When $\rho \ll L$, (4.10) can be approximated by

$$\frac{q(t)}{q_\infty} = \frac{2k_0}{c_0\rho}t - \left(\frac{k_0}{c_0\rho}t\right)^2,$$

which is also identical to (4.5). These results demonstrate that the previously obtained (4.4) and (4.5) are special cases of the general solution (4.10).

4.3 The Power-Law Model

Peppas and coworkers [66, 70] introduced a semiempirical equation (the so-called power law) to describe drug release from polymeric devices in a generalized way:

$$\frac{q(t)}{q_\infty} = kt^\lambda, \tag{4.11}$$

where k is a constant reflecting the structural and geometric characteristics of the delivery system expressed in dimensions of time$^{-\lambda}$, and λ is a release exponent the value of which is related to the underlying mechanism(s) of drug release. Equation (4.11) enjoys a wide applicability in the analysis of drug release studies and the elucidation of the underlying release mechanisms. Apart from its simplicity, the extensive use of (4.11) is mainly due to the following characteristics:

- Both Higuchi equations (4.1) and (4.3), which describe Fickian diffusional release from a thin polymer film, are special cases of (4.11) for $\lambda = 0.5$; also, (4.4) is a special case of (4.11) for $\lambda = 1$.
- It can describe adequately the first 60% of the release curve when (4.5) and (4.6) govern the release kinetics [68, 69].
- The value of the exponent λ obtained from the fitting of (4.11) to the first 60% of the experimental release data, from polymeric-controlled delivery systems of different geometries, is indicative of the release mechanism, Table 4.1.

Table 4.1 Values of the exponent λ in (4.11) and the corresponding release mechanisms from polymeric-controlled delivery systems of various geometries [65].

Exponent λ			**Release**
Thin film	Cylinder	Sphere	**mechanism**
0.5	0.45	0.43	Fickian diffusion
$0.5 < \lambda < 1.0$	$0.45 < \lambda < 0.89$	$0.43 < \lambda < 0.85$	Anomalous transport
1.0	0.89	0.85	Case II transport

From the values of λ listed in Table 4.1, only the two extreme values 0.5 and 1.0 for thin films (or slabs) have a physical meaning. When $\lambda = 0.5$, pure Fickian diffusion operates and results in diffusion-controlled drug release. It should be recalled here that the derivation of the relevant (4.3) relies on short-time approximations and therefore the Fickian release is not maintained throughout the release process. When $\lambda = 1.0$, zero-order kinetics (Case II transport) are justified in accord with (4.4). Finally, the intermediate values of λ (cf. the inequalities in Table 4.1) indicate a combination of Fickian diffusion and Case II transport, which is usually called *anomalous transport*.

It is interesting to note that even the more realistic model adhering to the Case II radial and axial drug release from a cylinder (4.10) can be described by the power-law equation. In this case, pure Case II drug transport and release is approximated (Table 4.1) by the following equation:

$$\frac{q(t)}{q_\infty} \approx kt^{0.89}. \tag{4.12}$$

A typical example of comparison between (4.10) and (4.12) when $\rho < L$ is shown in Figure 4.3. One should note the resemblance, along the first 60% of the curves, to the kinetic profiles derived from these equations.

4.3.1 Higuchi Model vs. Power-Law Model

Drug release data are frequently plotted as percent (or fractional) drug released vs. $t^{1/2}$. This type of plot is usually accompanied by linear regression analysis using $q(t)/q_\infty$ as dependent and $t^{1/2}$ as independent variable. This routinely applied procedure can lead to misinterpretations regarding the diffusional mechanism, as is shown below using simulation studies [71].

Simulated data were generated from (4.11) using values for λ and k ranging from 0.4 to 0.65 and from 0.05 to 0.5, respectively. The range of λ values is the neighborhood of the Higuchi exponent 0.5, which is the theoretical value for a diffusion-controlled release process. Moreover, values of λ in the range $0.4 - 0.65$ are frequently quoted in the literature for the discernment of drug release mechanisms (pure diffusion, anomalous transport, and combination) from HPMC matrix devices of different geometries [67, 68]. The values assigned to k are similar

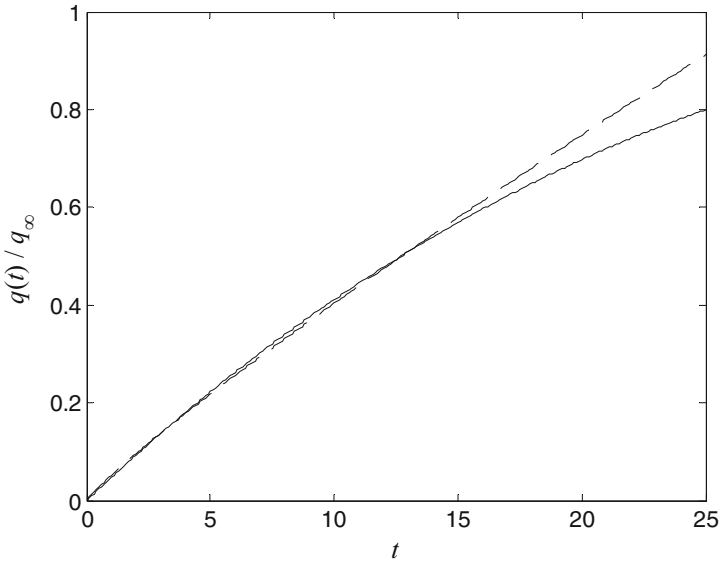


Fig. 4.3 Fractional drug release $q\,(t)\,/q_\infty$ vs. time (arbitrary units) for Case II transport with axial and radial release from a cylinder. Comparison of the solutions presented by (4.10) with $k_0 = 0.01$, $c_0 = 0.5$, $\rho = 1$, $L = 2.5$ (*dashed line*) and (4.12) with $k = 0.052$ (*solid line*)

to the estimates obtained when (4.3) is fitted to drug release data, whereas k has dimension of time$^{-1/2}$. The constraint $q\,(t)\,/q_\infty \leq 1$ was used for each set generated from (4.11). The duration of the simulated release experiment was arbitrarily set equal to 8 ($t \leq 8$). Therefore, the number of the simulated data generated from (4.11) varied according to the specific value assigned to k using in all cases a constant time step, 0.01. The pairs of data $(q\,(t)\,/q_\infty, t)$ generated from (4.11) were further analyzed using linear regression analysis in accord with (4.3).

Table 4.2 shows the results of linear regression analysis ($q\,(t)\,/q_\infty$ vs. $t^{1/2}$) for the data generated from (4.11). As expected, the theoretically correct sets of data ($\lambda = 0.5$) exhibited ideal behavior (intercept$= 0$, $R^2 = 1$). Judging from the determination coefficient R^2 values in conjunction with the number of data points utilized in regression, all other sets of data with $\lambda \neq 0.5$ are also described nicely if one does not apply a more rigorous analysis, e.g., plot of residuals. It is also worthy of mention that the positive intercepts were very close to zero and only in two cases ($k = 0.4, \lambda = 0.4$; $k = 0.5, \lambda = 0.4$) were they found to be in the range $0.10-0.11$. In parallel, any negative intercepts were very close to the origin of the axes.

These observations indicate that almost the entire set of data listed in Table 4.2 and generated from (4.11) can be misinterpreted as obeying (4.3). Under real experimental conditions the discernment of kinetics is even more difficult when linear regression of $q\,(t)\,/q_\infty$ vs. $t^{1/2}$ is applied. This is so if one takes into account

Table 4.2 Results of linear regression $q(t)/q_\infty$ vs. $t^{1/2}$ for data generated from (4.11). (a) Estimates not statistically significant different from zero were obtained. (b) Number of data points utilized in regression.

k	λ	Intercept	Slope	R^2	N^b
0.05	0.40	0.01287	0.03668	0.9970	800
	0.45	0.006719	0.04305	0.9993	800
	0.50	0^a	0.05	1	800
	0.55	−0.00576	0.05760	0.9994	800
	0.60	−0.01545	0.06571	0.9976	800
	0.65	−0.02436	0.07501	0.9950	800
0.30	0.40	0.0772	0.02201	0.9970	800
	0.45	0.04031	0.2583	0.9993	800
	0.50	0^a	0.3	1	800
	0.55	−0.04418	0.3456	0.9994	800
	0.60	−0.08866	0.3925	0.9976	743
	0.65	−0.1258	0.4349	0.9949	637
0.40	0.40	0.1030	0.2935	0.9970	800
	0.45	0.05270	0.3451	0.9993	766
	0.50	0^a	0.4	1	625
	0.55	−0.04676	0.4513	0.9994	529
	0.60	−0.08829	0.4987	0.9976	460
	0.65	−0.1253	0.5422	0.9948	4409
0.50	0.40	0.1117	0.3800	0.9969	565
	0.45	0.05243	0.4424	0.9993	466
	0.50	0^a	0.5	1	400
	0.55	−0.04649	0.5525	0.9993	352
	0.60	−0.0878	0.6002	0.9975	317
	0.65	−0.1245	0.6432	0.9947	290

- the usually small number of experimental data points available,
- the constraint for the percentage of drug released, $q(t)/q_\infty \leq 0.60$,
- the experimental error of data points,
- the high variability or lack of data points at the early stages of the experiment, and
- the possible presence of a delay in time.

Therefore, it is advisable to fit (4.11) directly to experimental data using nonlinear regression. Conclusions concerning the release mechanisms can be based on the estimates for λ and the regression line statistics [71].

4.4 Recent Mechanistic Models

Although the empirical and semiempirical models described above provide adequate information for the drug release mechanism(s), better insight into the release process can be gained from mechanistic models. These models have the advantage of being

more accurate and predictive. However, mechanistic models are more physically realistic and therefore mathematically more complex since they describe all concurrent physicochemical processes, e.g., diffusion, dissolution, swelling. Additionally, they require the use of time- and/or position-, direction-dependent diffusivities. This mathematical complexity is the main disadvantage of the mechanistic models since explicit analytical solutions of the partial differential equations cannot be derived. In this case, one has to rely on numerical solutions and less frequently on implicit analytical solutions.

Although the emphasis of this section will be on the most recent mechanistic approaches, the work of Fu et al. [72] published in 1976 should be mentioned since it deals with the fundamental release problem of a drug homogeneously distributed in a cylinder. In reality, Fu et al. [72] solved Fick's second law equation assuming constant cylindrical geometry and no interaction between drug molecules. These characteristics imply a constant diffusion coefficient in all three dimensions throughout the release process. Their basic result in the form of an analytical solution is

$$\frac{q(t)}{q_\infty} = 1 - \frac{8}{h^2\rho^2}\left[\sum_{i=1}^{\infty}\alpha_i^{-2}\exp\left(-\mathcal{D}\alpha_i^2 t\right)\right]\left[\sum_{j=1}^{\infty}\beta_j^{-2}\exp\left(-\mathcal{D}\beta_j^2 t\right)\right],$$

where $\beta_j = (2j+1)\,\pi/(2h)$, α_i are the roots of the equation $J_0(\rho\alpha) = 0$, and J_0 is the zero-order Bessel function. Here, h denotes the half-length, ρ the radius of the cylinder, and i and j are integers. Note that for small t the series is very slowly converging. Even keeping 100 terms of the above series is still not a good enough approximation of $q(t)/q_\infty$, for $t \approx 0$. For long times all terms with high values of α and β decay rapidly and only the term with the lowest value survives. The series reduces to a simple exponential after some time.

Gao et al. [73, 74] developed a mathematical model to describe the effect of formulation composition on the drug release rate for HPMC-based tablets. An effective drug diffusion coefficient $\mathcal{D}'$ was found to control the rate of release as derived from a steady-state approximation of Fick's law in one dimension:

$$\frac{q(t)}{q_\infty} = \frac{\mathcal{A}}{V}\sqrt{\frac{\mathcal{D}'t}{\pi}},$$

where $\mathcal{A}$ is the surface area and V the volume available for release, while $\mathcal{D}'$ corresponds to the quotient $\mathcal{D}/\tau$, where $\mathcal{D}$ is the classical drug diffusion coefficient in the release medium and τ is the tortuosity of the diffusing matrix.

In a series of papers Narasimhan and Peppas [75–77] developed models that take into account the dissolution of the polymer carrier. According to the theory, the polymer chain, at the surface of the system, disentangles (above a critical water concentration) and diffuses into the release medium. The kinetics of the polymer mass loss is controlled by the dissolution rate constant of the polymer and the decreasing with time surface area of the device. Symmetry planes in axial and radial

direction, placed at the center of the matrix, for the water and drug concentration profiles allow the development of an elegant mathematical analysis. Fick's second law of diffusion for cylindrical geometry is used to model both water and drug diffusion. Since both the composition and the dimensions of the device change with time while the diffusion coefficients for both species are considered to be dependent on the water content, the complex partial differential equations obtained are solved numerically. The model has been used successfully to describe the effect of the initial theophylline loading of HPMC-based tablets on the resulting drug release rate.

Recently, a very sophisticated mechanistic model called the *sequential layer* model was presented [78–83]; the model considers inhomogeneous polymer swelling, drug dissolution, polymer dissolution, and water and drug diffusion with nonconstant diffusivities and moving boundary conditions. The raptation theory was used for the description of polymer dissolution, while water and drug diffusion were described using Fick's second law of diffusion. An exponential dependence of the diffusion coefficients on the water content was taken into account. Moving boundaries were considered since the polymer swells, the drug and the polymer dissolve, thereby making the interface matrix/release medium not stationary. The model was applied successfully in the elucidation of the swelling and drug release behavior from HPMC matrices using chlorpheniramine maleate, propranolol HCl, acetaminophen, theophylline, and diclofenac as model drugs.

4.5 Monte Carlo Simulations

In a Monte Carlo simulation we attempt to follow the time evolution of a model that does not proceed in some rigorously predefined fashion, e.g., Newton's equations of motion. Monte Carlo simulations are appropriate for models whose underlying mechanism(s) are of a stochastic nature and their time evolution can be mimicked with a sequence of random numbers, which is generated during the simulation. The repetitive Monte Carlo simulations of the model with different sequences of random numbers yield results that agree within statistical error but are not identical. The goal is to understand the stochastic component of the physical process making use of the perfect control of "experimental" conditions in the computer-simulation experiment, examining every aspect of the system's configuration in detail. Since the mass transport phenomena, e.g., drug diffusion and the chemical processes, e.g., polymer degradation encountered in drug release studies, are random processes, Monte Carlo simulations are used to elucidate the release mechanisms. In the next section we demonstrate the validity of the Higuchi law using Monte Carlo simulations and in the following two sections we focus on the use of Monte Carlo simulations for the description of drug release mechanisms based on Fickian diffusion from Euclidean or fractal spaces. Finally, the last portion of this section deals with Monte Carlo simulations of drug release from bioerodible microparticles.

4.5.1 *Verification of the Higuchi Law*

The presuppositions for the application of the Higuchi law (4.2) have been discussed in Section 4.1. However, it is routinely quoted in the literature without a rigorous proof that only the first 60% of the release curve data should be utilized for a valid application of (4.2). Recently, this constraint has been verified for the Higuchi model using Monte Carlo computer simulations [84] (cf. Appendix B).

To mimic the conditions of the Higuchi model, a one-dimensional matrix of 200 sites has been constructed, Figure 4.4. Each site is labeled with the number of particles it currently hosts. Initially all sites have 10 particles, i.e., the total number n_0 of particles monitored is 2000. Drug molecules move inside the matrix by the mechanism of Fickian diffusion and cannot move to a site unless this site is empty. Thus, the system is expected to behave as if its "concentration" was much higher than its "solubility," which is the basic assumption made in the theoretical derivation of the Higuchi equation. The matrix can leak only from the site at its edge in full analogy with Figure 4.1. The diffusive escape process is simulated by selecting a particle at random and moving it to a randomly selected nearest-neighbor site. If the new site is an empty site, then the move is allowed and the particle is moved to this new site. If the new site is already occupied, the move is rejected. A particle is removed from the lattice as soon as it migrates to the leak site. After each particle move, time is incremented by arbitrary time units, the Monte Carlo microSteps (MCS), during which the movement takes place. One MCS is the smallest time unit in which an event can take place. The increment is chosen to be $1/n\,(t)$, where $n\,(t)$ is the number of particles remaining in the system. This is a typical approach in Monte Carlo simulations. The number of particles that are present inside the cylinder as a function of time is monitored until the cylinder is completely empty of particles. Figure 4.5 shows the simulation results for the first 60% of the release data; the slope of the line is 0.51 very close to the value 0.50 expected by the Higuchi equation.

The simulation results presented in Figure 4.5 provide an indirect proof of the valid use of the first 60% of the release data in line with (4.2). Needless to say, the Monte Carlo simulations in Figure 4.5 do not apply to the diffusion problem associated with the derivation of (4.3).

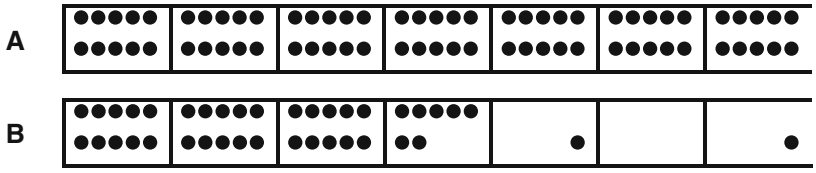


Fig. 4.4 Schematic of a system used to study diffusion under the Higuchi assumptions. (**A**) Initial configuration of the system, (**B**) evolution after time t. Particles are allowed to leak only from the right side of the system. Reprinted from [84] with permission from Springer

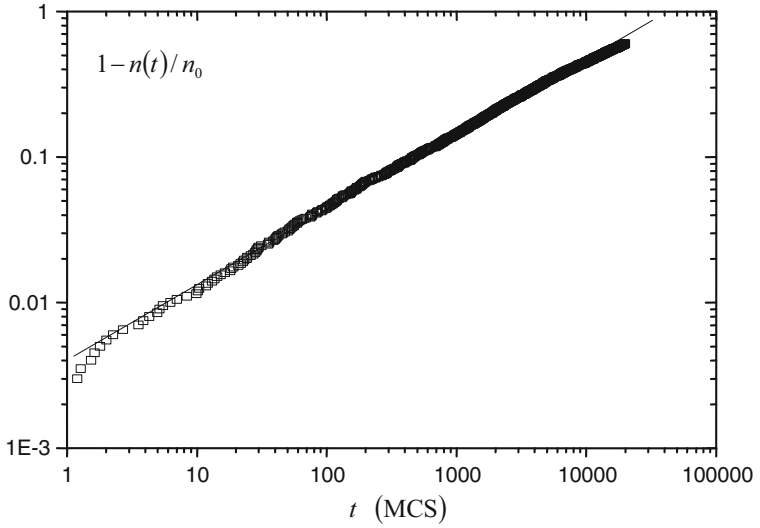


Fig. 4.5 Log-log plot of $1 - n(t)/n_0$ vs. time. Simulation results are indicated as points using the first 60% of the release data. The slope of the fitted line is 0.51 and corresponds to the exponent of the Higuchi equation. The theoretical prediction is 0.50

4.5.2 Drug Release from Homogeneous Cylinders

The general problem that we will focus on in this section is the escape of drug molecules[1] from a cylindrical vessel. Initially, theoretical aspects are presented demonstrating that the Weibull function can describe drug release kinetics from cylinders, assuming that the drug molecules move inside the matrix by a Fickian diffusion mechanism. Subsequently, Monte Carlo simulations will be used to substantiate the theoretical result and provide a link between the Weibull model and the physical kinetics of the release process [84].

4.5.2.1 Theoretical Aspects

A simple approximate solution is sought for the release problem, which can be used to describe release even when interacting particles are present. The particles are assumed to move inside the vessel in a random way. The particle escape rate is expected to be proportional to the number $n(t)$ of particles that exist in the vessel at time t. The rate will also depend on another factor, which will show how "freely" the particles are moving inside the vessel, how easily they can find the exits, how many of these exits there are, etc. This factor is denoted by g. Hence, a differential equation for the escape rate can be written

[1]The terms "drug molecule" and "particle" will be used in this section interchangeably.

$$\dot{n}(t) = -agn(t),$$

where a is a proportionality constant and the negative sign means that $n(t)$ decreases with time. If the factor g is kept constant, it may be included in a and in this case the solution of the previous equation is

$$n(t) = n_0 \exp(-at)$$

using the initial condition $n(0) = n_0$. The last equation is similar to the asymptotic result derived by Fu et al. [72] for pure Fickian diffusion inside a cylinder for long times (cf. Section 4.4).

It stands to reason to assume that the factor g should be a function of time since as time elapses a large number of drug molecules leave the vessel and the rest can move more freely. Thus, in general one can write that $g = g(t)$ and the previous differential equation becomes

$$\dot{n}(t) = -ag(t)n(t). \tag{4.13}$$

A plausible assumption is to consider that $g(t)$ has the form $g(t) \propto t^{-\mu}$.

We are interested in supplying a short-time approximation for the solution of the previous equation. There are two ways to calculate this solution. The direct way is to make a Taylor expansion of the solution. The second, more physical way, is to realize that for short initial time intervals the release rate $\dot{n}(t)$ will be independent of $n(t)$. Thus, the differential equation (4.13) can be approximated by $\dot{n}(t) = -ag(t)$. Both ways lead to the same result.

- For $\mu = 1/2$, (4.13) leads to $n(t) \propto \sqrt{t}$ (as a short-time approximation) exactly as predicted by the Higuchi law.
- For $\mu = 0$ we obtain, again as a short-time approximation, the result $n(t) = n_0 - at$ corresponding to *ballistic* exit (zero-order kinetics).

The above imply that choosing $g(t) = t^{-\mu}$ is quite reasonable. In this case (4.13) will be

$$\dot{n}(t) = -at^{-\mu}n(t).$$

Solving this equation we obtain

$$n(t) = n_0 \exp\left(-at^b\right), \tag{4.14}$$

where $b = 1 - \mu$.

The above reasoning shows that the stretched exponential function (4.14), or Weibull function as it is known, may be considered as an approximate solution of the diffusion equation with a variable diffusion coefficient due to the presence of particle interactions. Of course, it can be used to model release results even when no interaction is present (since this is just a limiting case of particles that are weakly interacting).

It is clear that it cannot be proven that the Weibull function is the best choice of approximating the release results. There are infinitely many choices of the form $g(t)$ and some of them may be better than the Weibull equation. This reasoning merely indicates that the Weibull form will probably be a good choice. The simulation results below show that it is indeed a good choice. The above reasoning is quite important since it provides a physical model that justifies the use of the Weibull function in order to fit experimental release data.

4.5.2.2 Simulations

A brief outline of the Monte Carlo techniques used for the problem of drug release from cylinders is described in Appendix B. The results obtained for cylinders of different dimensions are shown in Figure 4.6. In all cases it is possible to achieve a quite accurate fitting of the simulation results for $n(t)$ using the Weibull function [84]. It turns out that the exponent b takes values in the range 0.69 to 0.75. Figure 4.7 shows that the fitting is very accurate especially at the beginning, and it remains quite good until all of the drug molecules are released. The number of particles that have escaped from the matrix is equal to

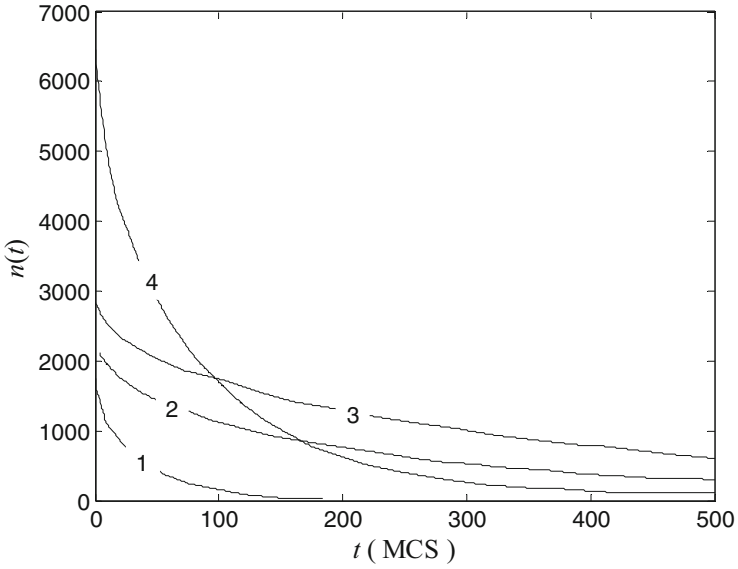


Fig. 4.6 Number of particles inside a cylinder as a function of time. (**1**) Cylinder with height of 31 sites and diameter 16 sites. Number of drug molecules $n_0 = 1750$. (**2**) Cylinder with height 7 sites and diameter 31 sites. Number of drug molecules $n_0 = 2146$. (**3**) Cylinder with height 5 sites and diameter 41 sites. Number of drug molecules $n_0 = 2843$. (**4**) Cylinder with height 51 sites and diameter 21 sites. Number of drug molecules $n_0 = 6452$

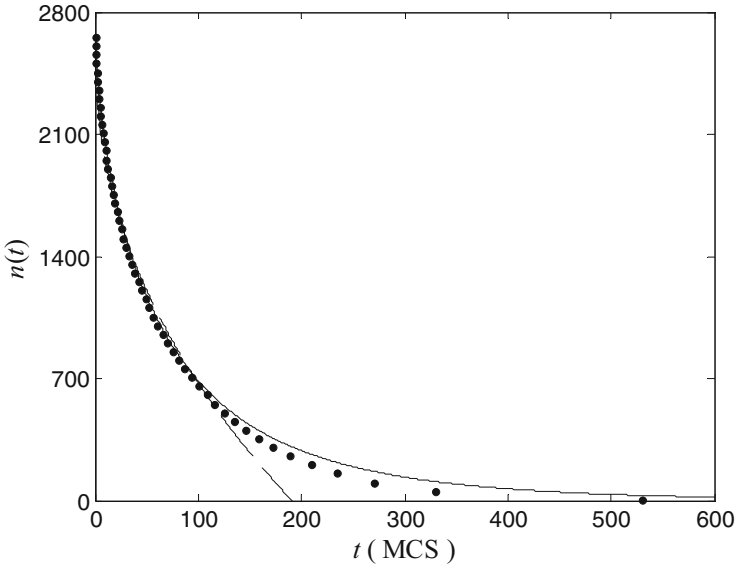


Fig. 4.7 Number of particles inside a cylinder as a function of time with initial number of drug molecules $n_0 = 2657$. Simulation for cylinder with height 21 sites and diameter 21 sites (*dotted line*). Plot of curve $n(t) = 2657 \exp\left(-0.049t^{0.72}\right)$, Weibull model fitting (*solid line*). Plot of curve $n(t) = 2657\left(1 - 0.094t^{0.45}\right)$, power-law fitting (*dashed line*)

$$\tilde{n}(t) = n_0 - n(t) = n_0\left[1 - \exp\left(-at^b\right)\right], \tag{4.15}$$

where a and b are parameters that have to be experimentally determined.

Ritger and Peppas [67, 68] have shown that the power law (4.11) describes accurately the first 60% of the release data. It is easy to show that the two models (4.11) and (4.14) coincide for small values of t. Note that $\tilde{n}(t)/n_0$ is directly linked to $q(t)/q_\infty$. From the Taylor expansion of $\exp(-\chi)$, we can say that for small values of χ we have $\exp(-\chi) \approx 1 - \chi$. From (4.15), setting $\chi = at^b$, one gets

$$\tilde{n}(t)/n_0 = at^b$$

for small values of at^b, which has the same form as the power-law model. For this approximation to hold, the quantity at^b has to be small. This does not mean that t itself must be small. As long as a is small, t may take larger values and the approximation will still be valid.

A comparison of the simulation results and fittings with the Weibull and the power-law model is presented in Figure 4.7. Obviously, the Weibull model describes quite well all release data, while the power law diverges after some time. Of course both models can describe equally well experimental data for the first 60% of the release curve.

4.5.2.3 The Physical Connection Between a, b and the System Geometry

The parameters a and b are somehow connected to the geometry and size of the matrix that contains the particles. This connection was investigated by performing release simulations for several cylinder sizes and for several initial drug concentrations [84]. The Weibull function was fitted to the simulated data to obtain estimates for a and b. If one denotes by N_{leak} the number of leak sites and by N_{tot} the total number of sites, in the continuum limit the ratio $N_{\text{leak}}/N_{\text{tot}}$ is proportional to the leak surface of the system. Plots of a vs. $N_{\text{leak}}/N_{\text{tot}}$ (not shown) were found to be linear and independent of the initial drug concentration; this implies that a is proportional to the specific leak surface, i.e., the surface to volume ratio. The slopes of the straight lines were found to be in the range $0.26 - 0.30$ [84]. The value of the slope can be related to the mathematical model presented in the theoretical section since the number of particles escaping at time dt was assumed to be proportional to $an\,(t)$; thus, the simulation results can be summarized as $an\,(t) = 0.28\,(N_{\text{leak}}/N_{\text{tot}})\,n\,(t)$. Assuming a uniform distribution of particles, $N_{\text{leak}}/N_{\text{tot}}$ is the probability that a particle is at a site that is just one step from the exit. Accordingly, $(N_{\text{leak}}/N_{\text{tot}})\,n\,(t)$ is the mean number of particles that are able to escape at a given instant of time. Since there are 6 neighboring sites in the three-dimensional space, the probability for a particle to make the escaping step is $1/6$ (≈ 0.17). It is quite close to the 0.28 value of the simulation. The difference is due to the fact that after some time, the distribution of particles is no longer uniform. There are more empty cells near the exits than inside, so the mean number of particles that are able to escape at a given instant is rather less than $(N_{\text{leak}}/N_{\text{tot}})\,n\,(t)$. This explains the higher value of the slope.

The plot of b values obtained from release simulations for several cylinder sizes and initial drug concentrations vs. $N_{\text{leak}}/N_{\text{tot}}$ (not shown) was also linear [84] with a slope practically independent of the initial concentration, $b = 0.65 + 0.4\,(N_{\text{leak}}/N_{\text{tot}})$. There are two terms contributing to b; one depends on $N_{\text{leak}}/N_{\text{tot}}$ and the other does not. Actually b is expected to be proportional to the specific surface, since a high specific surface means that there are a lot of exits, so finding an exit is easier. The constant term depends on the ability of the particles to move inside the matrix, the interaction between the particles, etc. The linear relationship yields the value of $b = 0.69$ when the exits cover the entire surface of the cylinder ($N_{\text{leak}} = N_{\text{tot}}$).

4.5.3 Release from Fractal Matrices

Apart from the classical mechanisms of release, e.g., Fickian diffusion from a homogeneous release device (cf. Sections 4.5.1 and 4.5.2) or Case II release there are also other possibilities. For example, the gastrointestinal fluids can penetrate the release device as it is immersed in the gastrointestinal tract fluids, creating areas of high diffusivity. Thus, the drug molecules can escape from the release device

through diffusion from these high diffusivity "channels." Now, the dominant release mechanism is diffusion, but in a complex disordered medium. The same is true when the polymer inside the release device is assuming a configuration resembling a disordered medium. This is a model proposed for HPMC matrices [85]. Several diffusion properties have to be modified when we move from Euclidean space to fractal and disordered media.

4.5.3.1 The Pioneering Work of Bunde et al.

The problem of the release rate from devices with fractal geometry was first studied by Bunde et al. [86]. This study was based on a percolation fractal-cluster at the critical point, assuming cyclic boundary conditions, embedded on a two-dimensional square lattice. The concentration of open sites is known to be approximately $p = 0.593$ (cf. Section 1.7). The fractal dimension of the percolation fractal is known to be $91/48$. The simulation starts with a known initial drug concentration $c_0 = 0.5$ and with randomly distributed drug molecules inside the fractal matrix. The drug molecules move inside the fractal matrix by the mechanism of diffusion. Excluded volume interactions between the particles, meaning that two molecules cannot occupy the same site at the same time, were also assumed. The matrix can leak from the intersection of the percolation fractal with the boundaries of the square box where it is embedded. Bunde et al. [86] specifically reported that the release rate of drug in a fractal medium follows a power law and justified their finding as follows: "the nature of drug release drastically depends on the dimension of the matrix and is different depending on whether the matrix is a normal Euclidean space or a fractal material such as a polymer, corresponding to the fact that the basic laws of physics are quite different in a fractal environment."

4.5.3.2 Can the Power Law Describe the "Entire" Release Curve?

Based on the findings of Bunde et al. [86], one can also conceive that the entire, classical % release vs. time curves from devices of fractal geometry should also follow a power law with (a different) characteristic exponent. Although the power law has been extensively used for the description of the initial 60% of the release data, it has also been shown that the power law can describe the entire drug release profile of several experimental data [71]. Typical examples of fittings of (4.11) to experimental data of drug release from HPMC matrices along with the estimates obtained for k and λ are shown in Figures 4.8, 4.9, and 4.10 [71]. In all cases, the entire release profile was analyzed and the fitting results were very good. All these experimental results were explained [71] on the basis of the Bunde et al. [86] findings. However, it will be shown below that the conclusion that the release rate follows a power law is accurate only for infinite problems. For problems in which the finite size is inherent, as happens to be the case in drug release studies, a power law is valid only in the initial stages of the release process.

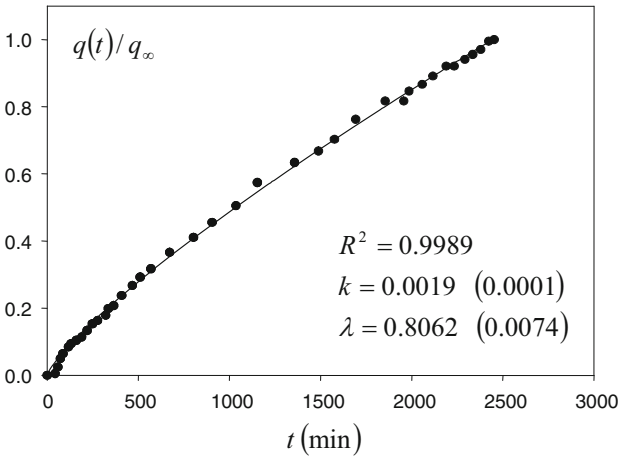


Fig. 4.8 Fitting of (4.11) to the entire set of fluoresceine release data from HPMC matrices [87]

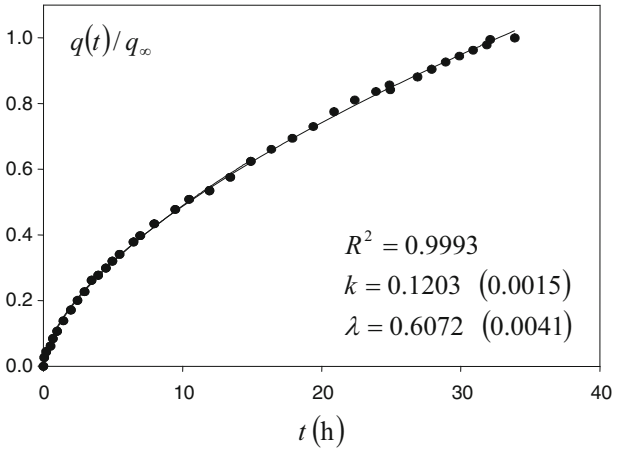


Fig. 4.9 Fitting of (4.11) to the entire set of buflomedil pyridoxal release data from HPMC matrices [88]

4.5.3.3 The Weibull Function Describes Drug Release from Fractal Matrices

Kosmidis et al. [89] reexamined the random release of particles from fractal polymer matrices using the percolation cluster at the critical point, Figure 4.11, following the same procedure as proposed by Bunde et al. [86]. The intent of the study was to derive the details of the release problem, which can be used to describe release when particles escape not from the entire boundary but just from a portion of the boundary of the release device under different interactions between the particles that are present.

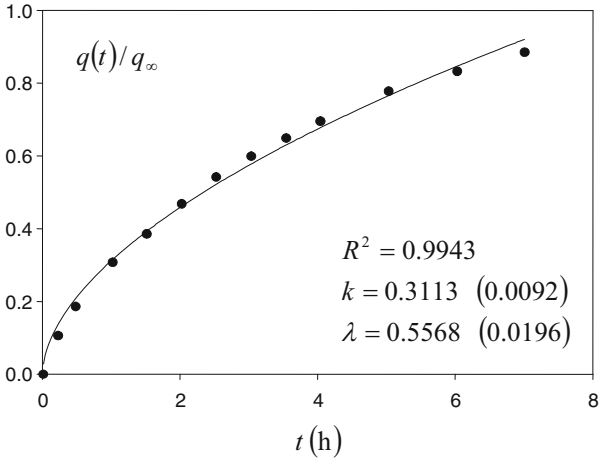


Fig. 4.10 A typical example of fitting (4.11) to chlorpheniramine maleate release data from HPMC K15M matrix tablets (tablet height 4 mm; tablet radius ratio 1 : 1) [83]

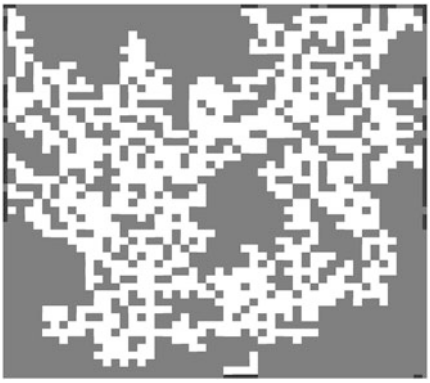


Fig. 4.11 A percolation fractal embedded on a 2-dimensional square lattice of size 50×50. Cyclic boundary conditions were used. We observe, especially on the boundaries, that there are some small isolated clusters, but these are not isolated since they are actually part of the largest cluster because of the cyclic boundary conditions. Exits (release sites) are marked in dark gray, while all lighter gray areas are blocked areas. Reprinted from [89] with permission from American Institute of Physics

The release problem can be seen as a study of the kinetic reaction $A + B \rightarrow B$ where the A particles are mobile, the B particles are static, and the scheme describes the well-known trapping problem [90]. For the case of a Euclidean matrix the entire boundary (i.e., the periphery) is made of the trap sites, while for the present case of a fractal matrix only the portions of the boundary that are part of the fractal cluster constitute the trap sites, Figure 4.11. The difference between the release problem and the general trapping problem is that in release, the traps are not randomly distributed

inside the medium but are located only at the medium boundaries. This difference has an important impact in real problems for two reasons:

- Segregation is known to play an important role in diffusion in disordered media (cf. Section 2.5.1). In the release problem the traps are *segregated* from the beginning, so one expects to observe important effects related to this segregation.
- The problem is inherently a finite-size problem. Results that otherwise would be considered as finite-size effects and should be neglected are in this case essential. At the limit of infinite volume there will be no release at all. Bunde et al. [86] found a power law also for the case of trapping in a model with a trap in the middle of the system, i.e., a classical trapping problem. In such a case, which is different from the model examined here, it is meaningful to talk about finite-size effects. In contrast, release from the surface of an infinite medium is impossible.

The fractal kinetics treatment of the release problem goes as follows [89]. The number of particles present in the system (vessel) at time t is $n(t)$. Thus, the particle escape rate will be proportional to the fraction g of particles that are able to reach an exit in a time interval dt, i.e., the number of particles that are sufficiently close to an exit. Initially, all molecules are homogeneously distributed over the percolation cluster. Later, due to the fractal geometry of the release system segregation effects will become important [17]. Accordingly, g will be a function of time, so that $g(t)$ will be used to describe the effects of segregation (generation of depletion zones), which is known to play an important role when the medium is disordered instead of homogeneous [17].

We thus expect a differential equation of the form of (4.13) to hold, where a is a proportionality constant, $g(t)\,n(t)$ denotes the number of particles that are able to reach an exit in a time interval dt, and the negative sign denotes that $n(t)$ decreases with time. This is a kinetic equation for an $A + B \rightarrow B$ reaction. The constant trap concentration $[B]$ has been absorbed in the proportionality constant a. The basic assumption of fractal kinetics [17] is that $g(t)$ has the form $g(t) \propto t^{-\mu}$. In this case, the solution is supplied by (4.14).

The form of this equation is a stretched exponential. In cases in which a system can be considered as infinite (for example, release from percolation fractals from an arbitrary site located at the middle of the volume) then the number of particles $n(t)$ inside the system is practically unchanged. Treating $n(t)$ as constant and letting $g(t) \propto t^{-\mu}$ in the right-hand side of (4.13) will lead to a power law for the quantity $\dot{n}(t)$. Since most physical problems belong to this class it is widely believed that the release rate from fractal matrices follows a power law. In the case of release from the periphery and if we want to study the system until all particles have escaped, as is often the case for practical applications, then (4.14) is of practical importance.

The above reasoning shows that the stretched exponential function (4.14), or Weibull function as it is known, may be considered as an approximate solution of the release problem. The advantage of this choice is that it is general enough for the description of drug release from vessels of various shapes, in the presence or

absence of different interactions, by adjusting the values of the parameters a and b. Monte Carlo simulation methods were used to calculate the values of the parameters a and (mainly) the exponent b [89].

4.5.3.4 Simulations

The drug molecules move inside the fractal matrix by the mechanism of diffusion, assuming excluded volume interactions between the particles. The matrix can leak at the intersection of the percolation fractal with the boundaries of the square box where it is embedded, Figure 4.11.

The diffusion process is simulated by selecting a particle at random and moving it to a randomly selected nearest-neighbor site. If the new site is an empty site, then the move is allowed and the particle is moved to this new site. If the new site is already occupied, the move is rejected since excluded volume interactions are assumed. A particle is removed from the lattice as soon as it migrates to a site lying within the leak area. After each particle move, time is incremented. As previously, the increment is chosen to be $1/n\,(t)$, where $n\,(t)$ is the number of particles remaining in the system. This is a typical approach in Monte Carlo simulations, and it is necessary because the number of particles continuously decreases, and thus, the time unit is MCS characterizing the system is the mean time required for all $n\,(t)$ particles present to move one step. The number of particles that are present inside the matrix as a function of time until a fixed number of particles (50 particles) remain in the matrix is monitored. The results are averaged using different initial random configurations over 100 realizations. The release rate $\dot{q}\,(t)$ is calculated by counting the number of particles that diffuse into the leak area in the time interval between t and $t + 1$.

Figure 4.12 shows simulation results (line) for the release of particles from a fractal matrix with initial concentration $c_0 = 0.50$, on a lattice of size 50×50. The simulation stops when more than 90% of the particles have been released from the matrix. This takes about $20,000$ MCS. In the same figure the data by Bunde et al. [86] (symbols), which cover the range $50 - 2,000$ MCS, are included. Because of the limited range examined in that study, the conclusion was reached that the release rate $\dot{q}\,(t)$ is described by a power law, with an exponent value between 0.65 and 0.75 [86]. With the extended range examined, Figure 4.12, this conclusion is not true, since in longer times $\dot{q}\,(t)$ deviates strongly from linearity, as a result of the finiteness of the problem.

In Figure 4.13, $n\,(t)\,/n_0$ is plotted as a function of time for different lattice sizes. The data were fitted with a Weibull function (4.14), where the parameter a ranges from 0.05 to 0.01 and the exponent b from 0.35 to 0.39. It has been shown [84] that (4.14) also holds in the case of release from Euclidean matrices. In that case the value of the exponent b was found to be $b \approx 0.70$.

These results reveal that the same law describes release from both fractal and Euclidean matrices. The release rate is given by the time derivative of (4.14). For early stages of the release, calculating the derivative of (4.14) and performing a Taylor series expansion of the exponential will result in a power law for the release

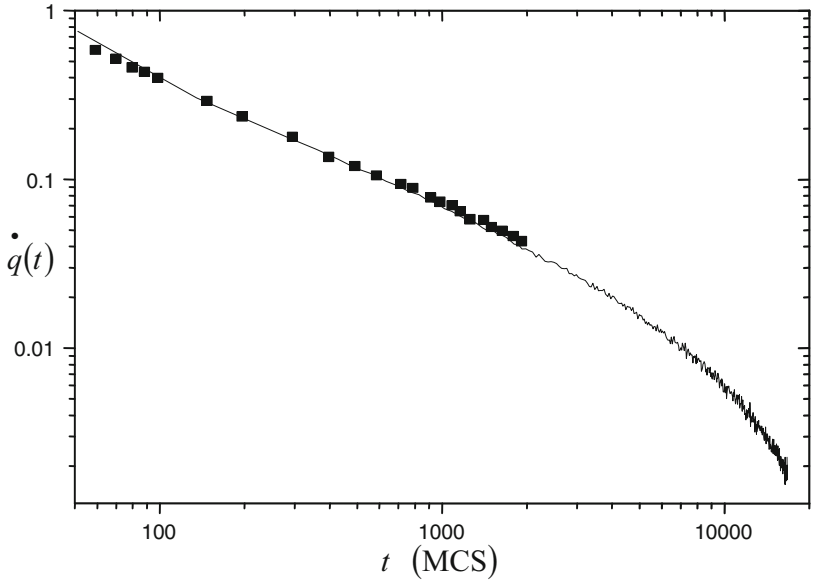


Fig. 4.12 Plot of the release rate $\dot{q}(t)$ vs. time. The lattice size is 50×50 and the initial concentration of particles is $c_0 = 0.50$. Points are the results given in [86], while the line is the result of the simulation in [89]

rate, just as Bunde et al. [86] have observed. If we oversimplify the release problem by treating it as a classical kinetics problem, we would expect a pure exponential function[2] instead of a stretched exponential (Weibull) function. The stretched exponential arises due to the segregation of the particles because of the fractal geometry of the environment. Concerning the release from Euclidean matrices [84], it has been demonstrated that the stretched exponential functional form arises due to the creation of a concentration gradient near the releasing boundaries. Note that although the functional form describing the release is the same in Euclidean and fractal matrices, the value of the exponent b is different, reflecting the slowing down of the diffusion process in a disordered medium. However, these results apparently point to a universal release law given by the Weibull function. The above considerations substantiate the use of the Weibull function as a more general form for drug release studies.

4.6 Discernment of Drug Release Kinetics

In the two previous sections the Weibull function was shown to be successful in describing the entire release profile assuming Fickian diffusion of drug from fractal as well as from Euclidean matrices. Since specific values were found for

[2]The classical kinetics solution is obtained by solving (4.13) in case of $g(t) = 1$.

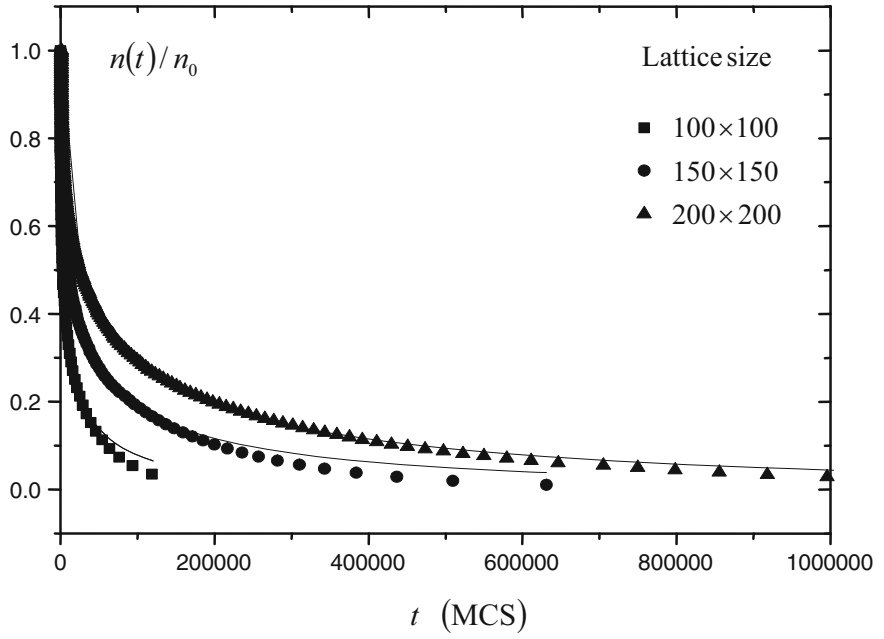


Fig. 4.13 Plot of the number of particles (normalized) remaining in the percolation fractal as a function of time t for lattice sizes 100 × 100, 150 × 150, and 200 × 200. $n(t)$ is the number of particles that remain in the lattice at time t and n_0 is the initial number of particles. Simulation results are represented by points. The *solid lines* represent the results of nonlinear fitting with a Weibull function

the exponent b for each particular case, a methodology based on the fitting results of the Weibull function (4.14) to the entire set of experimental %-release-time data can be formulated for the differentiation of the release kinetics[91]. Basically, successful fittings with estimates for b higher than one (sigmoid curves) rule out the Fickian diffusion of drug from fractal or Euclidean spaces and indicate a complex release mechanism. In contrast, successful fittings with estimates for b lower than one can be interpreted in line with the results of the Monte Carlo simulations of Sections 4.5.2 and 4.5.3. The exponent b of the Weibull function using the entire set of data was associated with the mechanisms of diffusional release as follows:

- $b < 0.35$: Not found in simulation studies [84, 89]. May occur in highly disordered spaces much different from the percolation cluster.
- $b \approx 0.35 - 0.39$: Diffusion in fractal substrate morphologically similar to the percolation cluster [89].
- $0.39 < b < 0.69$: Diffusion in fractal or disordered substrate different from the percolation cluster. These values were not observed in Monte Carlo simulation results [84, 89]. It is, however, plausible to assume this possibility since there has to be a crossover from fractal to Euclidian dimension.

- $b \approx 0.69 - 0.75$: Diffusion in normal Euclidean space [84].
- $0.75 < b < 1$: Diffusion in normal Euclidean substrate with contribution of another release mechanism. In this case, the power law can describe the entire set of data of a combined release mechanism (cf. below).
- $b = 1$: First-order release obeying Fick's first law of diffusion; the rate constant a controls the release kinetics, and the dimensionless solubility–dose ratio determines the final fraction of dose dissolved [92].
- $b > 1$: Sigmoid curve indicative of complex release mechanism. The rate of release increases up to the inflection point and thereafter declines.

When Fickian diffusion in normal Euclidean space is justified, further verification can be obtained from the analysis of 60% of the release data using the power law in accord with the values of the exponent quoted in Table 4.1. Special attention is given below for the values of b in the range $0.75 - 1.0$, which indicate a combined release mechanism. Simulated pseudodata were used to substantiate this argument assuming that the release obeys exclusively Fickian diffusion up to time $t = 90$ (arbitrary units), while for $t > 90$ a Case II transport starts to operate too; this scenario can be modeled using

$$\frac{q(t)}{q_\infty} = 1 - \exp\left(-0.05t^{0.70}\right) + \begin{cases} 0 & \text{for } t \le 90, \\ 0.004\,(t-90)^{0.89} & \text{for } t > 90. \end{cases} \tag{4.16}$$

Also, the following equation was used to simulate concurrent release mechanisms of Fickian diffusion and Case II transport throughout the release process:

$$\frac{q(t)}{q_\infty} = 1 - \exp\left(-0.05t^{0.70}\right) + 0.004t^{0.89}. \tag{4.17}$$

Pseudodata generated from (4.16) and (4.17) are plotted in Figure 4.14 along with the fitted functions

$$y(t) = 0.0652t^{0.5351} \text{ and } y(t) = 0.0787t^{0.5440}.$$

The nice fittings of the previous functions to the release data generated from (4.16) and (4.17), respectively, verify the argument that the power law can describe the entire set of release data following combined release mechanisms. In this context, the experimental data reported in Figures 4.8 to 4.10 and the nice fittings of the power-law equation to the entire set of these data can be reinterpreted as a combined release mechanism, i.e., Fickian diffusion and a Case II transport.

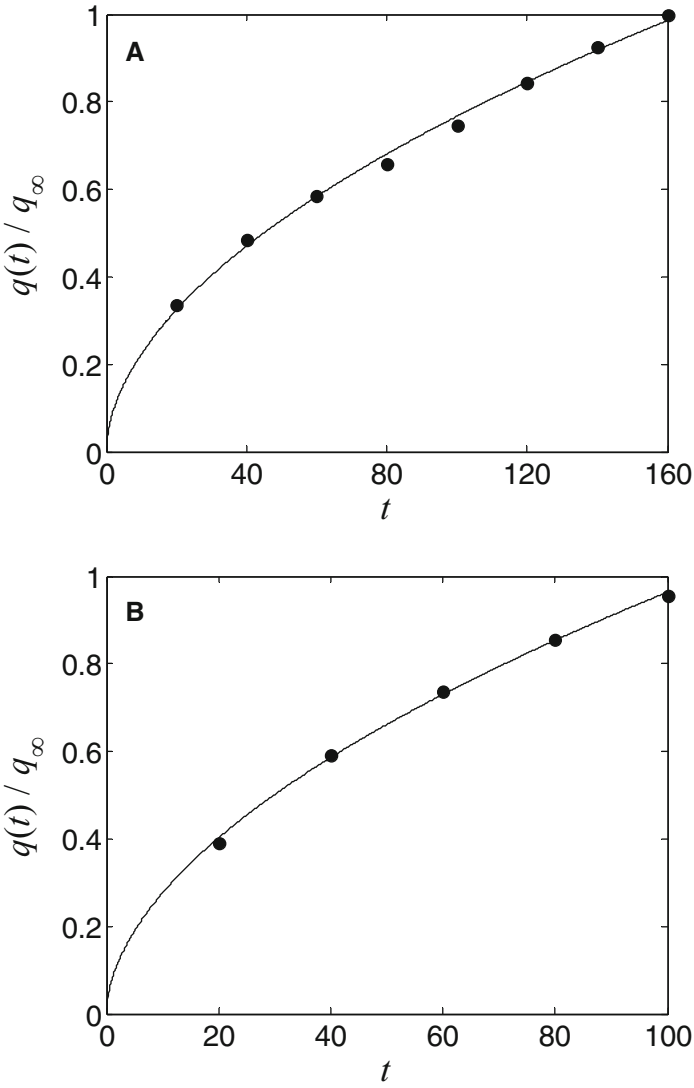


Fig. 4.14 (**A**) Points are simulation data produced using (4.16). The *solid line* is the fitting of the power law (4.11) to data. Best-fitting parameters are $k = 0.0652$ for the proportionality constant and $\lambda = 0.5351$ for the exponent. (**B**) Points are simulation data produced using (4.17). The *solid line* is the fitting of the power law (4.11) to data. Best-fitting parameters are $k = 0.0787$ for the proportionality constant and $\lambda = 0.5440$ for the exponent. Time is expressed in arbitrary units

4.7 Release from Bioerodible Microparticles

In bioerodible drug delivery systems various physicochemical processes take place upon contact of the device with the release medium. Apart from the classical physical mass transport phenomena (water imbibition into the system, drug dissolution, diffusion of the drug, creation of water-filled pores) chemical reactions (polymer degradation, breakdown of the polymeric structure once the system becomes unstable upon erosion) occur during drug release.

The mathematical model developed by Siepmann et al. [93] utilizes Monte Carlo techniques to simulate both the degradation of the ester bonds of the polymer poly-lactic-co-glycolic acid (PLGA) and the polymer's erosion (cleavage of the polymer chains throughout the PLGA matrix). Both phenomena are considered random, and the lifetime of the pixel representing the polymer's degradation is calculated as a function of a random variable obeying a Poisson distribution. The modeling of the physical processes (dissolution and diffusion) takes into account the increase of porosity of the matrix with time because of the polymer's erosion. This information is derived from the Monte Carlo simulations of the polymer's degradation–erosion and allows the calculation of the time- and position-dependent axial and radial diffusivities of the drug. Further, the diffusional mass transport processes are described using Fick's second law with spatially and temporally dependent diffusion coefficients. The numerical solution of the partial differential equation describing the kinetics of the three successive phases of drug release (initial burst, zero-order- and second rapid release) was found to be in agreement with the experimental release data of 5-fluorouracil loaded PLGA microparticles, Figure 4.15 [93]. This model has been further used to investigate the effect of the size of the biodegradable microparticles on the release rate of 5-fluorouracil [94].

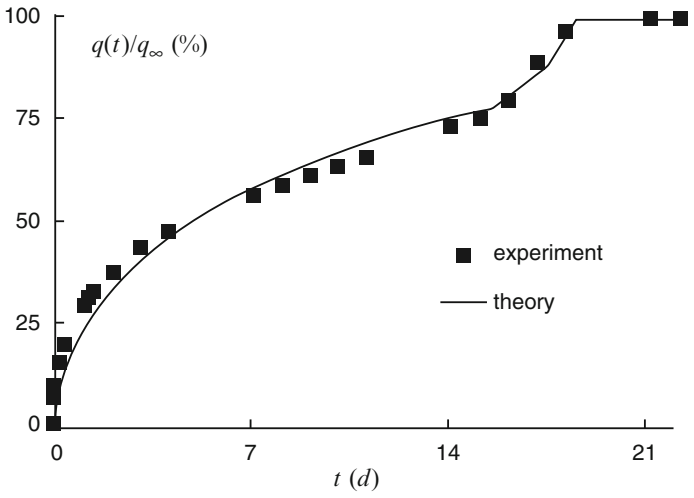


Fig. 4.15 Triphasic drug release kinetics from PLGA-based microparticles in phosphate buffer pH 7.4: experimental data (*symbols*) and fitted theory (*curve*). Reprinted from [93] with permission from Springer

4.8 Dynamic Aspects in Drug Release

Although the development of controlled drug delivery systems is usually based on the simple notion "a constant delivery is optimal," there are well-known exceptions. For example, drug administration in a periodic, pulsed manner is desirable for endogenous compounds, e.g., hormones [95]. The most classical example is the administration of insulin to diabetic patients in order to maintain blood glucose levels at an approximately constant level [96]. In reality, the pancreas behaves as a feedback controller, which changes its output with time in response to food intake or changes in metabolic activity. Hence, the delivery system should not simply maintain insulin levels within an acceptable physiological range to counterbalance the failure of the patient's pancreas to secrete sufficient insulin, but it should also mimic the normal pancreas's feedback controlling function. In other words, the delivery system should secrete insulin according to the (bio)sensed glucose levels in an automatic, periodic manner. These two steps, sensing and delivery, are the basic features of all self-regulated delivery systems regardless the variable, e.g., glucose, temperature, pressure, that is monitored to control the delivery of a pharmacological agent [97].

Since all these systems behave like autonomous oscillators fueled either directly or indirectly by the variable monitored, the factors involved in the production of pulsatile oscillations have been studied thoroughly. One of the most studied means for driving the periodic delivery of drugs is the utilization of chemical pH oscillators [98, 100, 101]. It was demonstrated that periodic drug delivery could be achieved as a result of the effect of pH on the permeability of acidic or basic drugs through lipophilic membranes. The model system of Giannos et al. [100] comprises a thin ethylene vinyl-acetate copolymer membrane separating a sink from an iodate-thiosulfate-sulfite pH oscillator compartment into which drugs like nicotine or benzoic acid are introduced. In the work of Misra and Siegel [98, 101] a model system consisting of the bromate-sulfite-marble pH oscillator in a continuously stirred tank reactor is used, along with acidic drugs of varying concentration. Figure 4.16 provides a schematic for the periodic flux of a drug through the membrane according to the pH oscillations. In one of the studies, Misra and Siegel [98] provided evidence that low concentrations of acidic drugs can attenuate and ultimately quench chemical pH oscillators by a simple buffering mechanism. In the second study, Misra and Siegel [101] demonstrated that multiple, periodic pulses of drug flux across the membrane can be achieved when the concentration of the drug is sufficiently low.

Another approach for periodically modulated drug release is based on an enzyme–hydrogel system, which, due to negative chemomechanical feedback instability, swells and de-swells regularly in the presence of a constant glucose level [102]. The enzyme glucose oxidase catalyses the conversion of glucose to gluconate and hydrogen ions; the latter affect the permeability of the poly(N-isopropylacrylamide-co-methacrylic acid) hydrogel membrane to glucose since the hydrogel swells with increasing pH and de-swells with decreasing pH, Figure 4.17. This system has been studied extensively from a dynamic point of view [99, 103].

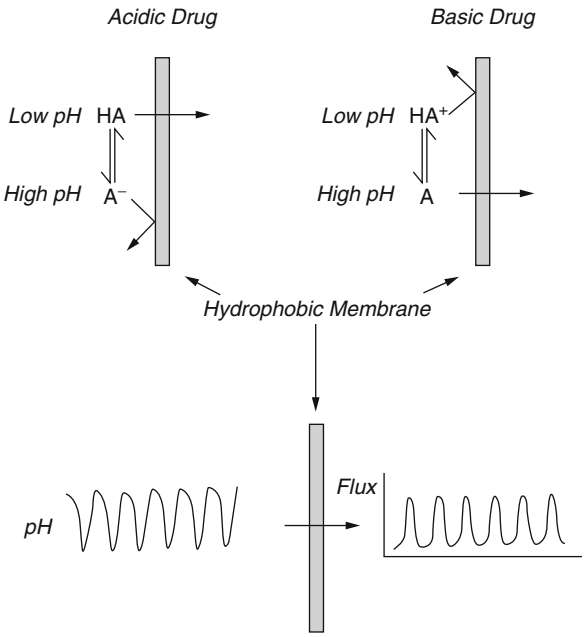


Fig. 4.16 Illustration of conversion of pH oscillations to oscillations in drug flux across a lipophilic membrane. Reprinted from [98] with permission from Wiley–Liss Inc., a subsidiary of John Wiley and Sons, Inc

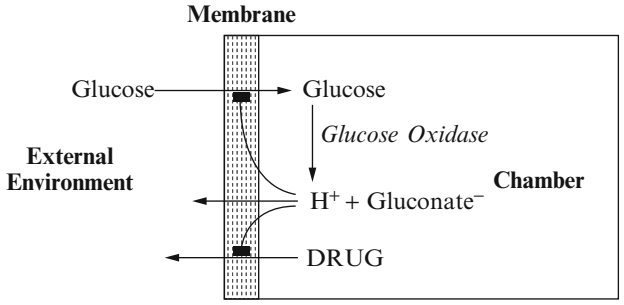


Fig. 4.17 Schematic of pulsating drug delivery device based on feedback inhibition of glucose transport to glucose oxidase through a hydrogel membrane. Changes in permeability to glucose are accompanied by modulation of drug permeability. Reprinted from [99] with permission from American Institute of Physics

It was found that the model allows, depending on system parameters and external substrate concentration, two separate single steady states, double steady state, and permanently alternating oscillatory behavior.

Chapter 5
Drug Dissolution

The rate at which a solid substance dissolves in its own solution is proportional to the difference between the concentration of that solution and the concentration of the saturated solution.

Arthur A. Noyes and Willis R. Whitney
Massachusetts Institute of Technology, Boston
Journal of the American Chemical Society 19:930–934 (1897)

The basic step in drug dissolution is the reaction of the solid drug with the fluid and/or the components of the dissolution medium. This reaction takes place at the solid–liquid interface and therefore dissolution kinetics are dependent on three factors, namely the flow rate of the dissolution medium toward the solid–liquid interface, the reaction rate at the interface, and the molecular diffusion of the dissolved drug molecules from the interface toward the bulk solution, Figure 5.1. As we stated in Section 2.4.2, a process (dissolution in our case) can be either diffusion or reaction-limited depending on which is the slower step. The relative importance of interfacial reaction and molecular diffusion (steps 2 and 3 in Figure 5.1, respectively) can vary depending on the hydrodynamic conditions prevailing in the microenvironment of the solid. This is so since both elementary steps 2 and 3 in Figure 5.1 are heavily dependent on the agitation conditions. For example, diffusion phenomena become negligible when externally applied intense agitation in in vitro dissolution systems gives rise to forced convection. Besides, the reactions at the interface (step 2) and drug diffusion (step 3) in Figure 5.1 are dependent on the composition of the dissolution medium. Again, the relative importance can vary according to the drug properties and the specific composition of the medium. It is conceivable that our limited knowledge of the hydrodynamics under in vivo conditions and the complex and position- and time-dependent composition of the gastrointestinal fluids complicates the study of dissolution phenomena in particular when one attempts to develop in vitro–in vivo correlations.

Early studies in this field of research formulated two main models for the interpretation of the dissolution mechanism: the diffusion layer model and the interfacial barrier model. Both models assume that there is a stagnant liquid layer in contact with the solid, Figure 5.2. According to the diffusion layer model (Figure 5.2A), the step that limits the rate at which the dissolution process occurs

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_5

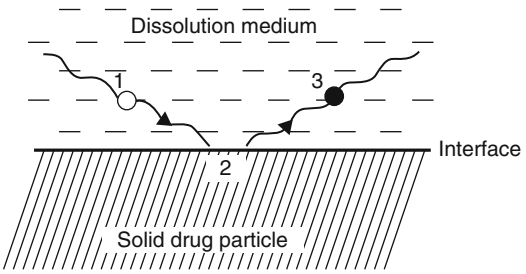


Fig. 5.1 The basic steps in the drug dissolution mechanism. (**1**) The molecules (∘) of solvent and/or the components of the dissolution medium are moving toward the interface; (**2**) adsorption–reaction takes place at the liquid–solid interface; (**3**) the dissolved drug molecules (•) move toward the bulk solution

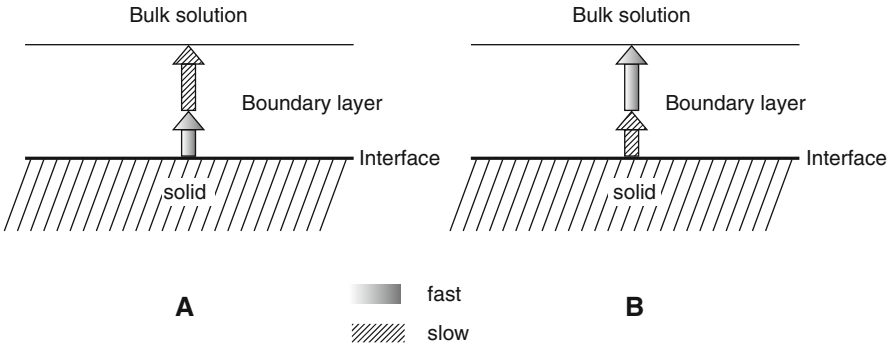


Fig. 5.2 Schematic representation of the dissolution mechanisms according to: (**A**) the diffusion layer model and (**B**) the interfacial barrier model

is the rate of diffusion of the dissolved drug molecules through the stagnant liquid layer rather than the reaction at the solid–liquid interface. For the interfacial barrier model (Figure 5.2B), the rate-limiting step of the dissolution process is the initial transfer of drug from the solid phase to the solution, i.e., the reaction at the solid–liquid interface.

Although the diffusion layer model is the most commonly used, various alterations have been proposed. The current views of the diffusion layer model are based on the so-called *effective diffusion boundary layer*, the structure of which is heavily dependent on the hydrodynamic conditions. In this context, Levich [104] developed the convection–diffusion theory and showed that the transfer of the solid to the solution is controlled by a combination of liquid flow and diffusion. In other words, both diffusion and convection contribute to the transfer of drug from the solid surface into the bulk solution. It should be emphasized that this observation applies even under moderate conditions of stirring.

5.1 The Diffusion Layer Model

Noyes and Whitney published [105] in 1897 the first quantitative study of a dissolution process. Using water as a dissolution medium, they rotated cylinders of benzoic acid and lead chloride and analyzed the resulting solutions at various time points. They found that the rate $\dot{c}(t)$ of change of concentration $c(t)$ of dissolved species was proportional to the difference between the saturation solubility c_s of the species and the concentration existing at any time t. Using k as a proportionality constant, this can be expressed as

$$\dot{c}(t) = k[c_s - c(t)] \; c(0) = 0. \tag{5.1}$$

Although it was not stated in the original article of Noyes and Whitney, it should be pointed out that the validity of the previous equation relies on the assumption that the amount used, q_0, is greater than or equal to the amount required to saturate the dissolution medium, q_s. Later on, (5.1) was modified [104, 106] and expressed in terms of the dissolved amount of drug $q(t)$ at time t while the effective surface area $\mathcal{A}$ of the solid was taken into account:

$$\dot{q}(t) = \frac{\mathcal{D}\mathcal{A}}{\delta}\left[c_s - \frac{q(t)}{V}\right] \; q(0) = 0, \tag{5.2}$$

where $\mathcal{D}$ is the diffusion coefficient of the substance, δ is the effective diffusion boundary layer thickness adjacent to the dissolving surface, and V is the volume of the dissolution medium. In this case, the first-order rate constant k (dimension of time^{-1}) appearing in (5.1) and governing the dissolution process is

$$k = \frac{\mathcal{D}\mathcal{A}}{\delta V}. \tag{5.3}$$

The integrated form of (5.2) gives the cumulative mass dissolved at time t:

$$q(t) = c_s V[1 - \exp(-kt)]. \tag{5.4}$$

The limit $t \to \infty$ defines the total drug amount, $q_s = c_s V$, that could be eventually dissolved in the volume V assuming that the amount used q_0 is greater than q_s. Thus, we can define the accumulated fraction of the drug in solution at time t as the ratio $q(t)/q_s$. Equation (5.4) expressed in terms of concentration ($c(t) = q(t)/V$) leads to the most useful form for practical purposes:

$$c(t) = c_s[1 - \exp(-kt)]. \tag{5.5}$$

Equation (5.5) is the classical equation quoted in textbooks indicating the exponential increase of concentration $c(t)$ approaching asymptotically the saturation solubility c_s.

Also, (5.1) indicates that initially ($t \to 0$) when $c\,(t)$ is small ($c\,(t) \leq 0.15c_s$) in comparison to c_s:

$$\dot{c}\,(t)\Big|_{t\to 0} = kc_s.$$

If this applies then we consider that sink conditions exist. Under sink conditions the concentration $c\,(t)$ increases linearly with time,

$$c\,(t) = kc_s t \; t \to 0, \tag{5.6}$$

and the dissolution rate is proportional to saturation solubility:

$$\dot{q}\,(t)\Big|_{t\to 0} = Vkc_s.$$

5.1.1 Alternative Classical Dissolution Relationships

The aforementioned analysis demonstrates that these classical concepts are in full agreement with Fick's first law of diffusion and the equivalent expressions in Sections 2.3 and 2.4. However, there are obvious deficiencies of the classical description of dissolution since the validity of (5.3) presupposes that all terms in this equation remain constant throughout the dissolution process. For example, the drug surface area $\mathcal{A}$ of powders and immediate release formulations is decreasing as dissolution proceeds. In fact, a dramatic reduction of the surface area takes place whenever the dose is not used in large excess, i.e., the drug mass divided by product of the volume of the dissolution medium and the drug's solubility is less than 10. This problem has been realized over the years and equations that take into account the diminution of the surface area have been published. For example, Hixson and Crowell [107] developed the following equation, which is usually called the *cube-root law*, assuming that dissolution occurs from spherical particles with a mono-disperse size distribution under sink conditions:

$$q_0^{1/3} - [q\,(t)]^{1/3} = k_{1/3}t, \tag{5.7}$$

where q_0 and $q\,(t)$ are the initial drug amount and the drug amount at time t after the beginning of the process, respectively, and $k_{1/3}$ is a composite cube-root rate constant. Alternatively, when sink conditions do not apply, the following equation (usually called the law of $2/3$) can be used:

$$[q\,(t)]^{-2/3} - q_0^{-2/3} = k_{2/3}t, \tag{5.8}$$

where $k_{2/3}$ is a composite rate constant for the law of $2/3$.

Although these approaches demonstrate the important role of the drug material's surface and its morphology on dictating the dissolution profile, they still suffer from limitations regarding the shape and size distribution of particles as well as the assumptions on the constancy of the diffusion layer thickness δ and the drug's diffusivity $\mathcal{D}$ throughout the process implied in (5.5), (5.6), (5.7), and (5.8). In reality, the parameters δ and $\mathcal{D}$ cannot be considered constant during the entire course of the dissolution process when poly-disperse powders are used and/or an initial phase of poor deaggregation of granules or poor wetting of formulation is encountered. In addition, the diffusion layer thickness appears to depend on particle size. For all aforementioned reasons, (5.5), (5.6), (5.7), and (5.8) have been proven adequate in modeling dissolution data only when the presuppositions of constancy of terms in (5.3) are fulfilled.

5.1.2 *Fractal Considerations in Drug Dissolution*

Drug particles are classically represented as ideal smooth spheres when dissolution phenomena are considered. The surface area of a spherical smooth object is a multiple of the scale, e.g., cm^2, and has a topological dimension $d_t = 2$. If one knows the radius ρ, the surface area of the sphere is $4\pi\rho^2$. However, many studies indicate that the surfaces of most materials are fractal [108]. The measured surface areas of irregular and rough surfaces increase with decreasing scale according to the specific surface structure. These surfaces have fractal dimensions d_f lying between the topological and the embedding dimensions: $2 < d_f < 3$.

Since the surface area of solids in dissolution studies is of primary importance, the roughness of the drug particles has been the subject of many studies. For example, Li and Park [109] used atomic force microscopy to determine the fractal properties of pharmaceutical particles. Moreover, analysis of the surface ruggedness of drugs, granular solids, and excipients using fractal geometry principles has been applied extensively [110–113]. Most of these studies underline the importance of surface ruggedness on dissolution. It is also interesting to note that considerations of the surface roughness are not restricted to the macroscopic level. The same concepts can also be applied to microscopic levels. A typical example is the importance of the surface roughness of proteins in binding phenomena [114].

Farin and Avnir [115] were the first to use fractal geometry to determine effects of surface morphology on drug dissolution. This was accomplished by the use of the concept of fractal reaction dimension d_r [116], which is basically the effective fractal dimension of the solid particle toward a reaction (dissolution in this case). Thus, (5.7) and (5.8) were modified [115] to include surface roughness effects on the dissolution rate of drugs for the entire time course of dissolution (5.9) and under sink conditions (5.10):

$$[q(t)]^{-\alpha} - q_0^{-\alpha} = \alpha k_{1/3}^* t, \tag{5.9}$$

$$q_0^{1-\alpha} - [q(t)]^{1-\alpha} = q_s(1-\alpha)\, k_{1/3}^* t, \tag{5.10}$$

where $\alpha = d_r/3$ and q_s is the drug amount that could be dissolved in the volume of the dissolution medium and $k^*_{1/3}$ is the dissolution rate constant of the modified cube-root. Although the previous equations describe quantitatively the dissolution of solids with fractal surfaces, their application presupposes that the value of d_r is known.

According to the classical scaling laws, an estimate of d_r can be obtained from the slope of a log–log plot of the initial rate of dissolution $\dot{q}(t)\Big|_{t\to 0}$ vs. the radius ρ of the various particle sizes. This kind of calculation relies on the fundamental proportionality

$$\dot{q}(t)\Big|_{t\to 0} \propto \mathcal{A} \propto \rho^{d_r-3},$$

where $\mathcal{A}$ is the effective surface area; the slope of $\log \dot{q}(t)\Big|_{t\to 0}$ vs. $\log \rho$ corresponds to $d_r - 3$, in agreement with the relationship for measurements regarding areas in Section 1.4.2. However, this approach for the calculation of d_r requires the execution of a number of experiments with a variety of particles of well-defined size and shape characteristics, which can also exhibit different d_r values.

For the aforementioned reasons, a simpler method requiring only a dissolution run with particles of a given size has been proposed for the estimation of d_r [117]. As can be seen from (5.9) and (5.10), on plotting the values of the left-hand side against time t, one can obtain the value of $k^*_{1/3}$ from the slope of the straight line. In practice, this involves choosing a starting value for d_r, e.g., 2, and, using an iterative method, searching for the linearity demanded by the previous equations for the experimental data pairs $(q(t), t)$. When this has been found, one knows values both for $k^*_{1/3}$ and d_r.

5.1.3 *On the Use of the Weibull Function in Dissolution*

In 1951, Weibull [118] described a more general function that can be applied to all common types of dissolution curves. This function was introduced in the pharmaceutical field by Langenbucher in 1972 [119] to describe the accumulated fraction of the drug in solution at time t, and it has the following form:[1]

$$\frac{q(t)}{q_\infty} = 1 - \exp\left[-(\lambda t)^\mu\right], \tag{5.11}$$

[1] In the pharmaceutical literature the exponential in the Weibull function is written as $\exp(-\lambda t^\mu)$ and therefore λ has dimension time$^{-\mu}$. In the version used herein (equation 5.11), the dimension of λ is time^{-1}.

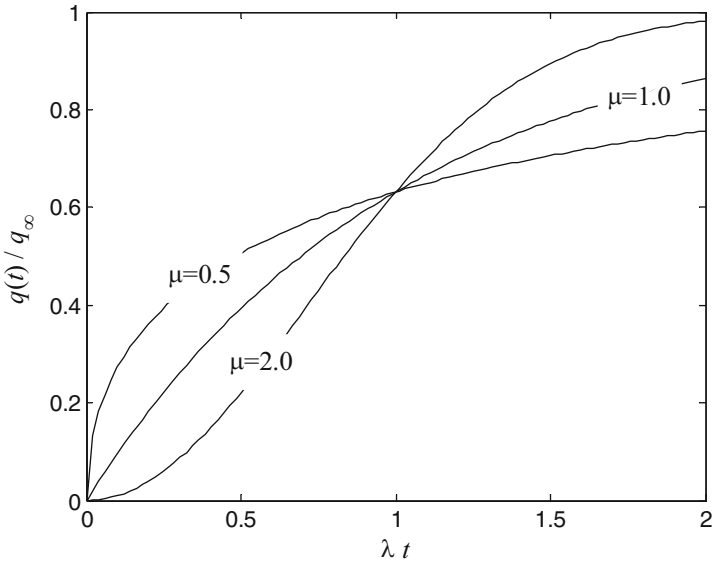


Fig. 5.3 Accumulated fraction of drug dissolved, $q(t)/q_\infty$ as a function of λt according to the Weibull distribution function (5.11)

where q_∞ is the total mass that can be eventually dissolved and λ, μ are constants. The *scale parameter* λ defines the time scale of the process, while the *shape parameter* μ characterizes the shape of the curve, which can be exponential ($\mu = 1$), *S*-shaped ($\mu > 1$), or exponential with a steeper initial slope ($\mu < 1$), Figure 5.3.

It is also worthy of mention that a gamma distribution function proposed by Djordjevic [120] for modeling in vitro dissolution profiles implies a relevant type of time dependency for the amount of drug dissolved.

The successful use of the Weibull function in modeling the dissolution profiles raises a plausible query: What is the rationale of its success? The answer will be sought in the relevance of the Weibull distribution to the kinetics prevailing during the dissolution process.

The basic theory of chemical kinetics originates in the work of Smoluchowski [121] at the turn of the twentieth century. He showed that for homogeneous reactions in three-dimensional systems the rate constant is proportional to the diffusion coefficient. In dissolution studies this proportionality is expressed with $k \propto \mathcal{D}$, where k is the intrinsic dissolution rate constant. In addition, both $\mathcal{D}$ and k are time-independent in well-stirred, homogeneous systems. However, that is not true for lower dimensions and disordered systems in chemical kinetics. Similarly, homogeneous conditions may not prevail during the entire course of the dissolution process in the effective diffusion boundary layer adjacent to the dissolving surface. It is very difficult to conceive that the geometric and hydrodynamic characteristics of this layer are maintained constant during the entire course of drug dissolution.

Accordingly, the drug's diffusional properties change with time and the validity of use of a classical rate constant k in (5.1) is questionable. It stands to reason that an instantaneous yet time-dependent rate coefficient $k(t)$ governing dissolution under inhomogeneous conditions can be written as

$$k(t) = k_\circ \left(\frac{t}{t_\circ}\right)^{-\gamma} \qquad \text{with} \qquad t \neq 0, \tag{5.12}$$

where $k_\circ$ is a rate constant not dependent on time, $t_\circ$ is a time scale parameter, and γ is a pure number. In a simpler form ($t_\circ = 1$), the previous relation is used in chemical kinetics to describe phenomena that take place under dimensional constraints or understirred conditions [17]. It is used here to describe the time dependency of the dissolution rate "constant" that originates from the change of the parameters involved in (5.3) during the dissolution process, i.e., the reduction of the effective surface area $\mathcal{A}$ and/or the inhomogeneous hydrodynamic conditions affecting δ and subsequently $\mathcal{D}$.

Using (5.12) to replace k in (5.1), also changing the concentration variables to amounts $Vc(t) = q(t)$, $V\mathrm{d}c(t) = \mathrm{d}q(t)$, and using, instead of $c_s V = q_s$, for generality purposes $c_\infty V = q_\infty$ (which applies to both $q_\infty = q_s$ and $q_\infty = q_0$), we obtain

$$\dot{q}(t) = k_\circ \left(\frac{t}{t_\circ}\right)^{-\gamma} [q_\infty - q(t)], \qquad q(t_0) = 0,$$

and after integration,

$$\frac{q(t)}{q_\infty} = 1 - \exp\left\{-\frac{k_\circ t_\circ}{1-\gamma}\left[\left(\frac{t}{t_\circ}\right)^{1-\gamma} - \left(\frac{t_0}{t_\circ}\right)^{1-\gamma}\right]\right\}.$$

Taking the limit as t_0 approaches zero, for $\gamma < 1$ we get the following equation:

$$\frac{q(t)}{q_\infty} = 1 - \exp\left[-\frac{k_\circ t_\circ}{1-\gamma}\left(\frac{t}{t_\circ}\right)^{1-\gamma}\right]. \tag{5.13}$$

This equation is identical to the Weibull equation (5.11) for

$$\lambda = \frac{1}{t_\circ}\left(\frac{k_\circ t_\circ}{1-\gamma}\right)^{1/(1-\gamma)} \qquad \text{and} \qquad \mu = 1 - \gamma.$$

Furthermore, (5.13) collapses to the "homogeneous" (5.4) when $\gamma = 0$. These observations reveal that the parameter μ of (5.11) can be interpreted in terms of the heterogeneity of the process. For example, an S-shaped dissolution curve with $\mu > 1$ in (5.11) for an immediate release formulation can now be interpreted as a heterogeneous dissolution process (with $\gamma < 0$ in equation 5.13), whose rate

increases with time during the upwards, concave initial limb of the curve and decreases after the point of inflection. This kind of behavior can be associated with an initial poor deaggregation or poor wetting.

Most importantly, it was shown that the structure of the Weibull function captures the time-dependent character of the rate coefficient governing the dissolution process. These considerations agree with Elkoski's [122] analysis of the Weibull function and provide an indirect, physically based interpretation [123] for its superiority over other approaches for the analysis of dissolution data. In other words, drug dissolution is a typical example of a heterogeneous process since, as dissolution proceeds, homogeneous conditions cannot be maintained in the critical region of the microenvironment of drug particles. Thus, drug dissolution exhibits fractal-like kinetics like other heterogeneous processes (e.g., adsorption, catalysis) since it takes place at the boundary of different phases (solid–liquid) under topological constraints.

5.1.4 Stochastic Considerations

The dissolution process can be interpreted stochastically since the profile of the accumulated fraction of amount dissolved from a solid dosage form gives the probability of the residence times of drug molecules in the dissolution medium. In fact, the accumulated fraction of the drug in solution, $q(t)/q_\infty$, has a statistical sense since it represents the cumulative distribution function of the random variable *dissolution time* T, which is the time up to dissolution for an individual drug fraction from the dosage form. Hence, $q(t)/q_\infty$ can be defined statistically as the probability that a molecule will leave the formulation prior to t, i.e., that the particular dissolution time T is smaller than t:

$$q(t)/q_\infty = \Pr[\text{leave the formulation prior to } t] = \Pr[T < t].$$

Conversely,

$$1 - q(t)/q_\infty = \Pr[\text{survive in the formulation to } t] = \Pr[T \geq t].$$

Since $q(t)/q_\infty$ is a distribution function, it can be characterized by its statistical moments. The first moment is defined as the *mean dissolution time* (*MDT*) and corresponds to the expectation of the time up to dissolution for an individual drug fraction from the dosage form:

$$MDT = E[T] = \int_0^\infty t \frac{\mathrm{d}q(t)}{q_\infty} = \frac{ABC}{q_\infty}, \tag{5.14}$$

where q_∞ is the asymptote of the dissolved amount of drug and ABC is the area between the cumulative dissolution curve and the horizontal line that corresponds to q_∞, Figure 5.4.

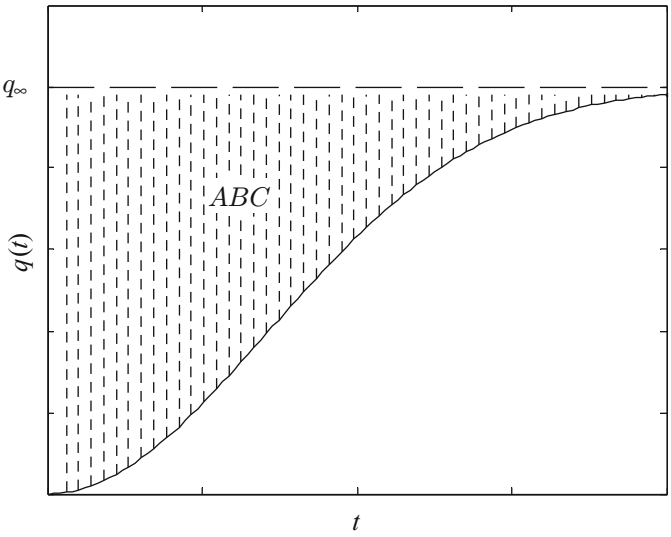


Fig. 5.4 The cumulative dissolution profile $q(t)$ as a function of time. The symbols are defined in the text

Since the fundamental rate equation of the diffusion layer model has the typical form of a first-order rate process (5.1), using (5.4) and (5.14), the *MDT* is found equal to the reciprocal of the rate constant k:

$$MDT = \frac{1}{k}. \tag{5.15}$$

As a matter of fact, all dissolution studies, which invariably rely on (5.1) and do not make dose considerations, utilize (5.15) for the calculation of the *MDT*. However, the previous equation applies only when the entire available amount of drug (dose) q_0 is dissolved. Otherwise, the mean dissolution time of the dose is not defined, i.e., *MDT* is infinite.

In fact, it will be shown below that *MDT* is dependent on the *solubility–dose ratio* if one takes into account the dose q_0 actually utilized [92]. Also, it will be shown that the widely used (5.15) applies only to a special limiting case. Multiplying both parts of (5.1) by V/q_0 (volume of the dissolution medium/actual dose), one gets the same equation in terms of the fraction of the actual dose of drug dissolved, $\varphi(t) \triangleq q(t)/q_0$:

$$\dot{\varphi}(t) = k\left[\tfrac{1}{\theta} - \varphi(t)\right], \varphi(0) = 0, \tag{5.16}$$

where θ is the solubility–dose ratio

$$\theta \triangleq \frac{q_0}{c_s V} = \frac{q_0}{q_s} \tag{5.17}$$

expressed as a dimensionless quantity. Equation (5.16) has two solutions:

- When $\theta \leq 1$ ($q_0 \leq q_s$), which means that the entire dose is eventually dissolved:

$$\varphi(t) = \begin{cases} \frac{1}{\theta}\left[1 - \exp(-kt)\right] & \text{for } t < t_\circ, \\ 1 & \text{for } t \geq t_\circ, \end{cases}$$

where $t_\circ = -\frac{\ln(1-\theta)}{k}$ is the time at which dissolution terminates ($\varphi(t_\circ) = 1$). Similarly to (5.14), the *MDT* is

$$MDT = \int_0^{t_\circ} t \mathrm{d}\varphi(t) = \frac{\theta + (1-\theta)\ln(1-\theta)}{k\theta}. \tag{5.18}$$

This equation reveals that the *MDT* depends on both k and θ. Figure 5.5 shows a plot of *MDT* as a function of θ for three different values of the rate constant k. Note that (5.15) is obtained from (5.18) for $\theta = 1$ (the actual dose is equal to the amount needed to saturate the volume of the dissolution medium). In other words, the classically used (5.15) is a special case of the general equation (5.18).

- When $\theta > 1$ ($q_0 > q_s$), which means that only a portion of the dose is dissolved and the drug reaches the saturation level $1/\theta$:

$$\varphi(t) = \frac{1}{\theta}\left[1 - \exp(-kt)\right].$$

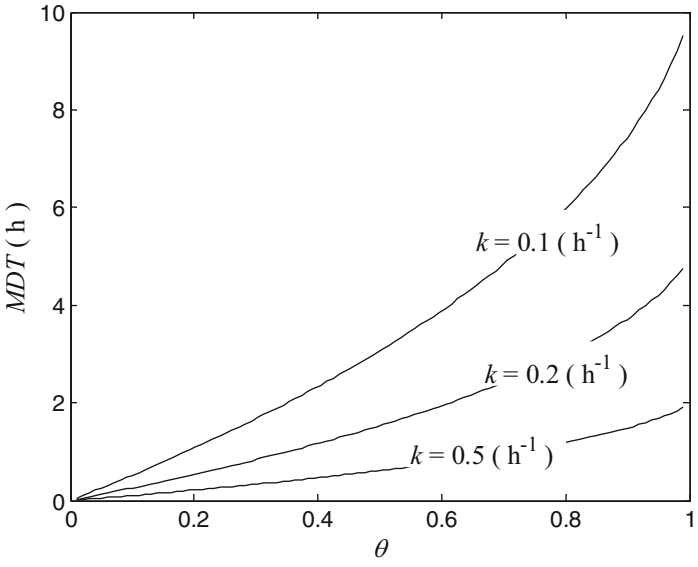


Fig. 5.5 Plot of *MDT* vs. θ using (5.18) for different values of k

The *MDT* is infinite because the entire dose is not dissolved. Therefore, the term mean saturation time, MDT_s, [124] has been suggested as more appropriate when we refer only to the actually dissolved portion of dose, in order to get a meaningful time scale for the portion of the dissolved drug dose:

$$MDT_s = \int_0^\infty t \frac{\mathrm{d}\varphi\,(t)}{1/\theta} = \frac{1}{k}, \tag{5.19}$$

which is independent of θ.

This analysis demonstrates that when $\theta \leq 1$, solubility–dose considerations should be taken into account in accord with (5.18) for the calculation of *MDT*; the *MDT* is infinite when $\theta > 1$. Equation (5.15) can be used to obtain an estimate for *MDT* only in the special case $\theta = 1$. Finally, (5.19) describes the MDT_s of the fraction of dose dissolved when $\theta > 1$.

5.2 The Interfacial Barrier Model

In the interfacial barrier model of dissolution it is assumed that the reaction at the solid–liquid interface is not rapid due to the high free energy of activation requirement and therefore the reaction becomes the rate-limiting step for the dissolution process (Figure 5.1), thus, drug dissolution is considered as a reaction-limited process for the interfacial barrier model. Although the diffusion layer model enjoys widespread acceptance since it provides a rather simplistic interpretation of dissolution with a well-defined mathematical description, the interfacial barrier model is not widely used because of the lack of a physically based mathematical description.

In recent years three novel models [124–126] have appeared that were proposed to describe the heterogeneous features of drug dissolution. They are considered here as continuous (in well-stirred media) or discrete (in understirred media) reaction-limited dissolution models. Their derivation and relevance are discussed below.

5.2.1 Continuous Reaction-Limited Dissolution Models

Lansky and Weiss [124] proposed a novel model by considering the reaction of the undissolved solute with the free solvent yielding the dissolved drug complexed with solvent:

$$[\text{undissolved drug}] + [\text{free solvent}] \rightarrow [\text{dissolved drug complexed with solvent}]\,.$$

Further, global concentrations as a function of time for the reactant species of the above reaction were considered, assuming that the solvent is not in excess and applying classical chemical kinetics. The following equation was found to describe the rate of drug dissolution in terms of the fraction of drug dissolved:

$$\dot{\varphi}(t) = k^* [1 - \varphi(t)] [1 - \theta\varphi(t)], \varphi(0) = 0, \tag{5.20}$$

where $\varphi(t)$ denotes the fraction of drug dissolved up to time t, and θ is the dimensionless solubility–dose ratio (5.17); k^* is a fractional (or relative) dissolution rate constant with dimensions time^{-1}. The fractional dissolution rate is a decreasing function of the fraction of dissolved amount $\varphi(t)$, as has also been observed for the diffusion layer model (5.16). However, (5.20) reveals a form of second-order dependency of the reaction rate on the dissolved amount $\varphi(t)$. In reality, a classical second-order dependency is observed for $\theta = 1$. These are unique features, which are not encountered in models dealing with diffusion-limited dissolution. All the above characteristics indicate that (5.20) describes the continuous–homogeneous character of the reaction of the solid with the solvent or the component(s) of the dissolution medium, i.e., a reaction-limited dissolution process in accord with the interfacial barrier model.

The solution of (5.20) for $\theta \neq 1$ yields the monotonic function

$$\varphi(t) = \frac{\exp[k^*(1-\theta)t] - 1}{\exp[k^*(1-\theta)t] - \theta}, \tag{5.21}$$

and for $\theta = 1$,

$$\varphi(t) = \frac{k^* t}{k^* t + 1},$$

with the same asymptotes as found above for the diffusion layer model, i.e., $\varphi(\infty) = 1$ for $\theta \leq 1$ and $\varphi(\infty) = 1/\theta$ for $\theta > 1$. It is interesting to note that both MDT and MDT_s for the model of the previous equation depend on the solubility–dose ratio θ when $\theta \neq 1$. Thus, the MDT for $\theta < 1$ is

$$MDT = -\frac{1}{k^*\theta} \ln(1 - \theta), \tag{5.22}$$

while the MDT_s for $\theta > 1$ is

$$MDT_s = \frac{1}{k^*} \ln\left(\frac{\theta}{\theta - 1}\right). \tag{5.23}$$

For $\theta = 1$ the MDT is infinite. It should be noted that the MDT for the diffusion layer model depends also on θ for $\theta < 1$ while the MDT_s is equal to $1/k$ when $\theta \geq 1$, (5.18) and (5.19). However, this dependency is different in the two models, cf. (5.18), (5.19), and (5.22), (5.23).

In 2008 Dokoumetzidis et al. [126] published a complete analysis of a continuous reaction-limited model of dissolution based on a bidirectional chemical reaction of the undissolved drug species with the free solvent molecules yielding the dissolved species of drug complex with solvent. This bidirectional reaction governed by the k_{+1} and k_{-1} rate constants can be considered in either sink conditions, where it corresponds to the unidirectional case and the entire drug amount is dissolved, or reaching chemical equilibrium, which corresponds to saturation of the solution. The model equation derived for the drug concentration $y(t)$ in mass per volume units is

$$\dot{y}(t) = k_{+1} (w_0)^b \left[\frac{q_0}{V} - y(t)\right]^a - k_{-1} y(t) \tag{5.24}$$

where q_0 is the initial quantity (dose) in mass units, w_0 is the initial concentration of the free solvent species, a and b are exponents dependent on the stoichiometry of the reaction and/or the geometry (surface) of the solid particles, and V is the volume of the dissolution medium.

Simpler, special cases of the above equation can be considered. For example, when $a = 1$ the undissolved species have equal probability to dissolve, implying that they are in a form of a well-mixed dispersion. Upon integration of the above equation for $a = 1$ one ends up

$$\varphi(t) = \frac{1}{q_{ss}} \left\{1 - \exp\left[-\left(k_{+1} (w_0)^b + k_{-1}\right) t\right]\right\}$$

where $\varphi(t)$ is the fraction of dose dissolved and

$$q_{ss} = \frac{k_{+1} (w_0)^b + k_{-1}}{k_{+1} (w_0)^b}$$

is a dimensionless constant. Other simpler cases are also considered in [126]. The models derived were fitted successfully to dissolution experimental data sets. On the contrary, functions based on the diffusion layer model fitted to experimental data, failed to reveal the governing role of saturation solubility in the dissolution process. One of the most important results of this study is that the rate of dissolution of a reaction-limited approach is driven by the concentration of the undissolved species and solubility is considered to be the concentration when the reaction equilibrium is reached.

5.2.2 A Discrete Reaction-Limited Dissolution Model

Dokoumetzidis and Macheras [125] developed a population growth model for describing drug dissolution under heterogeneous conditions. In inhomogeneous media, Fick's laws of diffusion are not valid, while global concentrations cannot

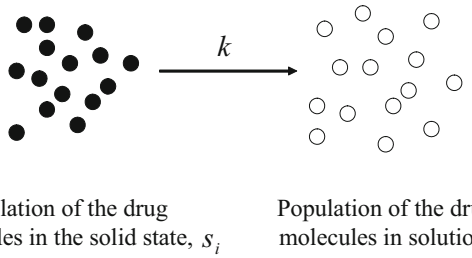


Fig. 5.6 A discrete, reaction-limited dissolution process interpreted with the population growth model of dissolution

be used in the dissolution rate equation. In order to face the problem of complexity and circumvent describing the system completely, the reaction of the solid with the solvent or the component(s) of the dissolution medium was described as the "birth" of the population of the dissolved drug molecules from the corresponding population of solid drug particles, Figure 5.6. In this context, only instants of the system's behavior are considered and what happens in the meanwhile is ignored. The jump from one instant to the next is done by a logical rule, which is not a physical law, but an expression that gives realistic results based on logical assumptions. The variable of interest (mass dissolved) is not considered as a continuous function of time, but is a function of a discrete time index specifying successive "generations."

Defining s_i and y_i as the populations of the drug molecules in the solid state and in solution in the i-th generation ($i = 0, 1, 2, \ldots$), respectively, the following finite difference equation describes the change of y_i between generations i and $i + 1$:

$$y_{i+1} = y_i + ks_i = y_i + k\left(q_0 - y_i\right), y_0 = 0,$$

where k is a proportionality constant that controls the reaction of the solid particles with the solvent or the components of the dissolution medium, and q_0 is the population of the drug molecules in the solid state corresponding to dose (Figure 5.6). The growth of y_i is not unlimited since the solubility of drug in the medium restricts the growth of y_i. Thus, the rate of dissolution decreases as the population of the undissolved drug molecules decreases as reaction proceeds. For each one of the drug particles of the undissolved population, the solubility q_s (expressed in terms of the amount needed to saturate the medium in the neighborhood of the particle) is used as an upper "local" limit for the population growth of the dissolved drug molecules. Accordingly, the growth rate is a function of the population level and can be assumed to decrease with increasing population in a linear manner:

$$k \to k\left(y_i\right) = k\left(1 - \frac{y_i}{q_s}\right),$$

where q_s is the saturation level of the population, i.e., the number of drug molecules corresponding to saturation solubility. Thus, the previous recursion relation is replaced with the nonlinear discrete equation:

$$y_{i+1} = y_i + k\,(q_0 - y_i)\left(1 - \tfrac{y_i}{q_s}\right),\ y_0 = 0.$$

This equation can be normalized in terms of dose by dividing both sides by q_0 and written more conveniently using $y_i/q_0 = \varphi_i$, $y_{i+1}/q_0 = \varphi_{i+1}$, and $\theta = q_0/q_s$:

$$\varphi_{i+1} = \varphi_i + k\,(1 - \varphi_i)\,(1 - \theta\varphi_i)\,,\ \varphi_0 = 0, \tag{5.25}$$

where φ_i and φ_{i+1} are the dissolved fractions of drug dose at generations i and $i+1$, respectively. The previous discrete equation, if written as

$$\varphi_{i+1} - \varphi_i = k\,(1 - \varphi_i)\,(1 - \theta\varphi_i)\,,\ \varphi_0 = 0, \tag{5.26}$$

becomes equivalent to its continuous analogue (5.20). As expected, (5.26) has the two classical fixed point, $\varphi_A^* = 1$ when $\theta \leq 1$ and $\varphi_B^* = 1/\theta$ when $\theta > 1$, Figure 5.7. All discrete features of (5.26) are in full analogy with the fractional dissolution rate differential equation (5.20), and it is for this reason that the two approaches are considered counterparts [124].

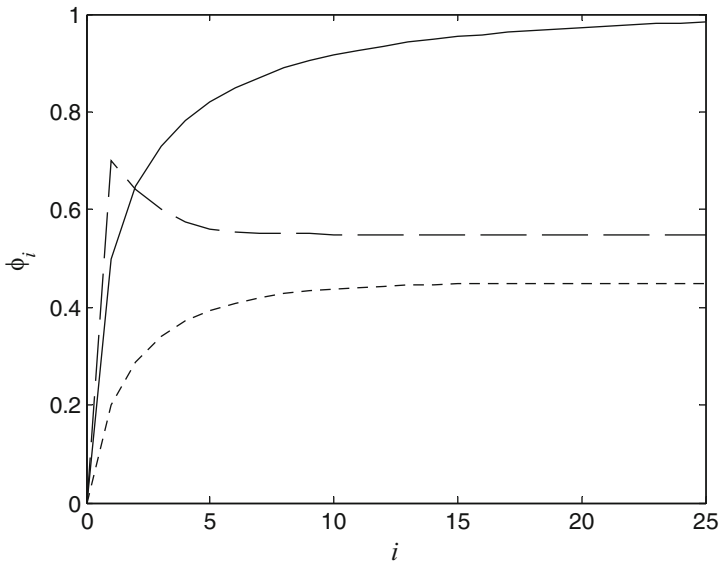


Fig. 5.7 Plot of the dissolved fraction φ_i as a function of generations i using (5.25) with $k = 0.5$, $\theta = 0.83$ (*solid line*); $k = 0.7$, $\theta = 1.82$ (*dashed line*); $k = 0.2$, $\theta = 2.22$ (*dotted line*)

Since difference equations exhibit dynamic behavior [127, 128], the stability of the fixed points of (5.25) is explored according to the methodology presented in Appendix A. The absolute value of the derivative of the right-hand side of (5.25) is compared with unity for each fixed point. There are the following cases:

- If $\theta < 1$, the derivative is equal to $1 - k\,(1 - \theta)$ and the condition for stability of the fixed point $\varphi_A^* = 1$ is

$$0 < k < \frac{2}{1-\theta}.$$

- If $\theta > 1$, the derivative is equal to $1 - k\,(\theta - 1)$ and the condition for stability of the fixed point $\varphi_B^* = 1/\theta$ is

$$0 < k < \frac{2}{\theta - 1}.$$

- If $\theta = 1$, the derivative is equal to unity and therefore the fixed point $\varphi_A^* = 1$ is neither stable or unstable.

Because of the discrete nature of (5.26), the first step always gives $\varphi_1 = k$; hence, k is always lower than 1, i.e., the theoretical top boundary of φ_i. Comparing the second step $\varphi_2 = k + k\,(1 - k)\,(1 - \theta k)$ with the first one $\varphi_1 = k$, one can obtain the conditions $k > 1/\theta$ and $\theta > 1$, which ensure that the first step is higher than the following steps (Figure 5.7B). The usual behavior encountered in dissolution studies, i.e., a monotonic exponential increase of φ_i reaching asymptotically 1, or the saturation level $1/\theta$, is observed when $\theta \leq 1$ (Figure 5.7A) or when $k < 1/\theta$ for $\theta > 1$ (Figure 5.7C), respectively.

As previously pointed out, when one uses (5.25) for $\theta > 1$ and values of k in the range $1/\theta < k < 2/\,(\theta - 1)$, the first step is higher than the plateau value followed by a progressive decline to the plateau (Figure 5.8A, B). For $1/\theta$ and k values close enough, the descending part of the dissolution curve is smooth, concave either upward (Figure 5.8B) or initially downward and then upward (Figure 5.8A); this decline can also take the form of a fading oscillation when k is close to $2/\,(\theta - 1)$ (Figure 5.8C, D). When k exceeds $2/\,|\theta - 1|$, the fixed points become unstable, bifurcating to a double-period stable fixed point. So we have both the unstable main point and the generated double-period stable point. This mechanism is called bifurcation and is common to dynamic systems (cf. Chapter 3).

Equation (5.26) can be used to estimate the proportionality constant k and θ from experimental data by plotting the fraction dissolved (φ_i) as a function of the generations i. Prior to plotting, the sampling times are transformed to generations defining arbitrarily a constant sampling interval as a "time unit." By doing so, an initial estimate for k can be obtained by reading the value of φ_i corresponding to the first datum point. When $\theta > 1$ an initial estimate for θ can be obtained from the highest value of the dissolved fraction at the end of the dissolution run. However, an estimate for θ cannot be obtained from visual inspection when $\theta \leq 1$ since $\varphi_A^* = 1$

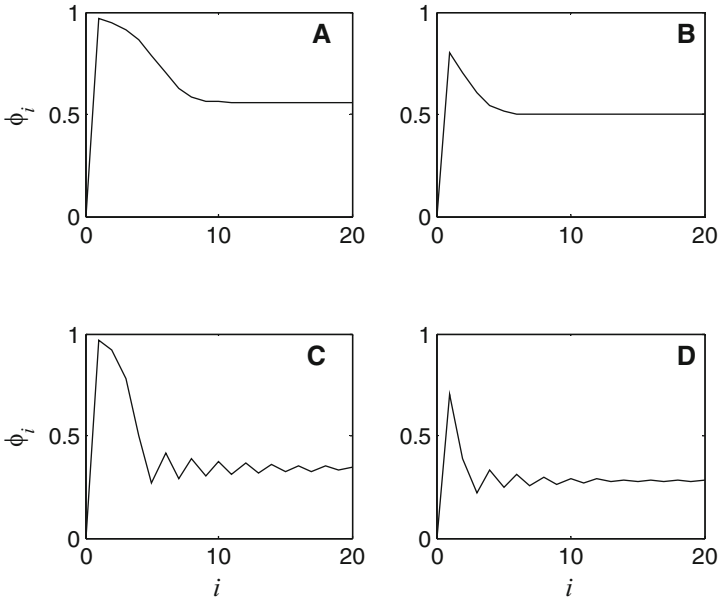


Fig. 5.8 Plots of the dissolved fraction φ_i as a function of generations i using (5.25) with k and θ values satisfying the inequality $1/\theta < k < 2/(\theta - 1)$: (**A**) $k = 0.97$, $\theta = 1.79$; (**B**) $k = 0.8$, $\theta = 2.0$; (**C**) $k = 0.97$, $\theta = 2.94$; (**D**) $k = 0.7$, $\theta = 3.57$

in all cases. The initial estimates for k and θ can be further used as starting points in a computer fitting program to obtain the best parameter estimates.

The population growth model of dissolution utilizes the usual information available in dissolution studies, i.e., the amount dissolved at certain fixed intervals of time. The time points of all observations need to be transformed to equally spaced values of time and furthermore to take the values $0, 1, 2, \ldots$. Since the model does not rely on diffusion principles it can be applied to both homogeneous and inhomogeneous conditions. This is of particular value for the correlation of in vitro dissolution data obtained under homogeneous conditions and in vivo observations adhering to the heterogeneous milieu of the gastrointestinal tract. The dimensionless character of k allows comparisons to be made for k estimates obtained for a drug studied under different in vitro and in vivo conditions, e.g., various dissolution media, fasted, or fed state.

Example 2. Danazol Data

For the continuous model, a fitting example of (5.21) to actual experimental data of danazol [129] is shown in Figure 5.9. For the discrete model, a number of fitting examples are shown in Figure 5.10 for danazol dissolution data obtained by using 15 minutes as a "time unit." Table 5.1 lists the estimates for k and θ obtained from the computer analysis of danazol data utilizing an algorithm minimizing the sum of squared deviations between experimental and theoretical values obtained from (5.25). ■

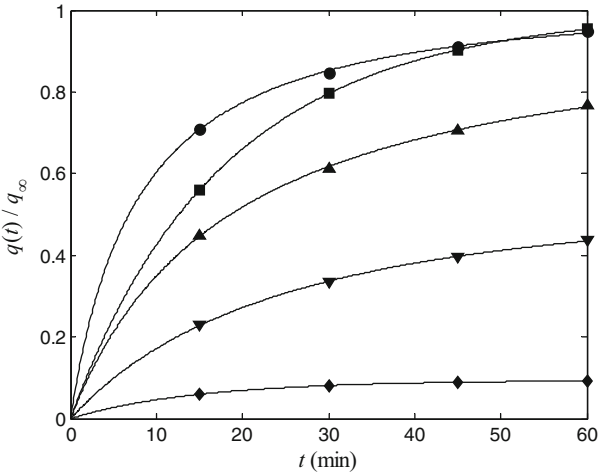


Fig. 5.9 The fraction of dose dissolved as a function of time for the danazol data [129]. Symbols represent experimental points and the lines represent the fittings of (5.21) to data. Key (% sodium lauryl sulfate in water as dissolution medium): ● 1.0; ■ 0.75; ▲ 0.50; ▼ 0.25; ♦ 0.10

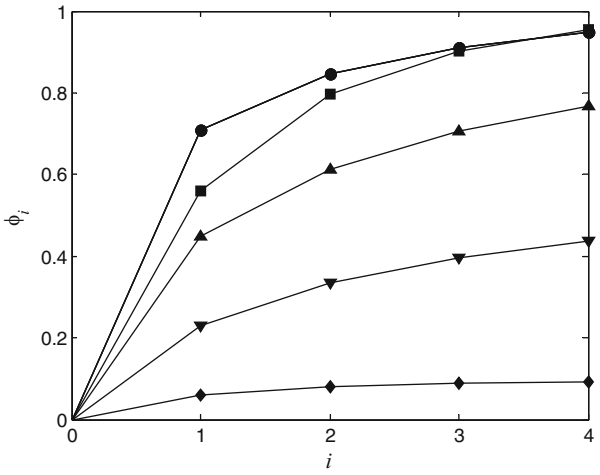


Fig. 5.10 The fraction of dose dissolved φ_i as a function of generations i, where the *solid line* represents the fittings of (5.25) to danazol data [129]. Symbols represent experimental points transformed to the discrete time scale for graphing and fitting purposes assigning one generation equal to 15 minutes. Key (% sodium lauryl sulfate in water as dissolution medium): ● 1.0; ■ 0.75; ▲ 0.50; ▼ 0.25; ♦ 0.10

Table 5.1 Estimates for k and θ obtained from the fitting of (5.24) to danazol data, Figure 5.10. (a) Percentage of sodium lauryl sulfate in water, (b) Determination coefficient.

Dissolution medium[a]	k	θ	R^{2b}
0.10	0.06	10	0.993
0.25	0.23	1.82	0.9993
0.50	0.45	0.75	0.9999
0.75	0.56	0.08	0.9995
1.00	0.71	0.47	0.9996

5.2.3 *Modeling Supersaturated Dissolution Data*

The dissolution data are basically of monotonic nature (the drug concentration or the fraction of drug dissolved is increasing with time) and therefore the corresponding modeling approaches rely on monotonic functions. However, nonmonotonic dissolution profiles are frequently observed in studies dealing with co-precipitates of drugs with polymers and solid dispersion formulations [130, 131]. The dissolution profiles in these studies usually exhibit a supersaturation phenomenon, namely, an initial rapid increase of drug concentration to a supersaturated maximum followed by a progressive decline to a plateau value. This kind of behavior cannot be explained with the classical diffusion principles in accord with the diffusion layer model of dissolution. It seems likely that the initial sudden increase is associated with a rapid reaction of the solid particles with the dissolution medium. The dynamics of the difference equation for the population growth model of dissolution, (5.25), can capture this behavior and therefore can be used to model supersaturated dissolution data [132].

Example 3. Nifedipine Data

An example of fitting (5.25) to experimental data of a nifedipine solid dispersion formulation [131] is shown in Figure 5.11. Initially, the drug concentration values are transformed to the corresponding dissolved fractions of dose φ_i and plotted as a function of the generations i, obtained by using a "time unit" of 5 minutes. The transformation of sampling times to generations i is achieved by adopting the time needed to reach maximum concentration (equivalent to maximum fraction of dose dissolved) as the time unit of (5.25). Reading the maximum and lowest values of φ_i, one obtains initial estimates for parameters k and $1/\theta$, respectively. These values are further used as starting points in a computer program minimizing the sum of squared deviations between observed and predicted values to determine the best parameter estimates. The estimated parameter values for k and θ were found to be 0.323 and 4.06, respectively. The value of k denotes the maximum fraction of dose that is dissolved in a time interval equal to the time unit used. The value of θ corresponds to the reciprocal of the plateau value, which is the fraction of dose remaining in solution at steady state. ■

However, the use of (5.25) should not be considered as a panacea for modeling nonmonotonic dissolution curves. Obvious drawbacks of the model (5.25) are

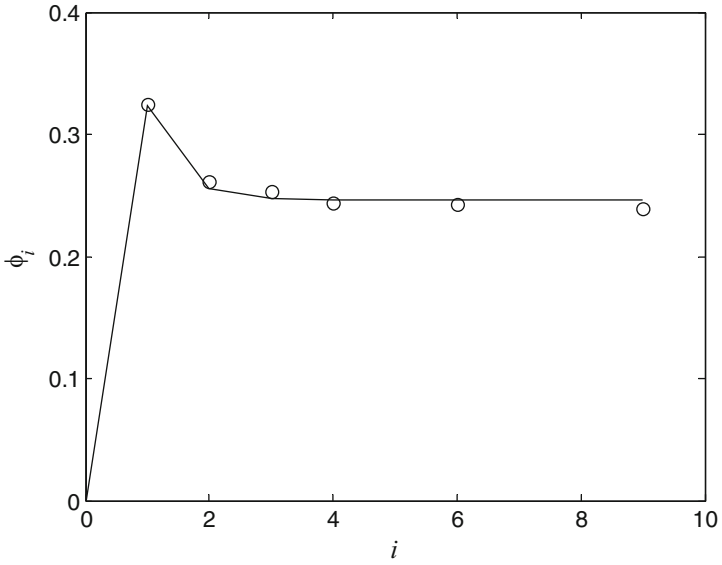


Fig. 5.11 Plot of the dissolved fraction φ_i as a function of generations i (time step 5 min) using (5.25) for the dissolution of nifedipine solid dispersion with nicotinamide and polyvinylpyrrolidone (1 : 3 : 1), in 900 ml of distilled water. Fitted line of (5.25) is drawn over the experimental data

1. The data on the ascending limb of the dissolution curve, if any, should be ignored.
2. The time required to reach the maximum value of the dissolved fraction of drug should be adopted as the time interval between successive generations.
3. The time values of the data points that can be used for fitting purposes should be integer multiples of the time unit adopted.

Further, when k takes values much larger than $1/\theta$, (5.25) exhibits chaotic behavior following the period-doubling bifurcation (cf. Chapter 3). For example, (5.25) leads to chaos when $1/\theta = 0.25$ and k is greater than 0.855. Despite the aforementioned disadvantages, the model offers the sole approach that can be used to describe supersaturated dissolution data. In addition, the derivation of (5.25) relies on a model built from physical principles, i.e., a reaction-limited dissolution model. Other approaches based on empirical models, e.g., polynomial functions, could provide better fittings for supersaturated dissolution data but these approaches will certainly lack in physical meaning.

In 2011 Charkoftaki et al. [133] modified the continuous reaction-limited model of dissolution [126] to describe classical experimental supersaturated dissolution data of carbamazepine in presence of d-alpha-tocopheryl polyethylene glycol 1000 succinate (TPGS). The model developed was based on a time-dependent expression for the forward microconstant of the bidirectional reaction carbamazepine-TPGS at the solid–liquid interface. The following modified version of equation 5.24 was

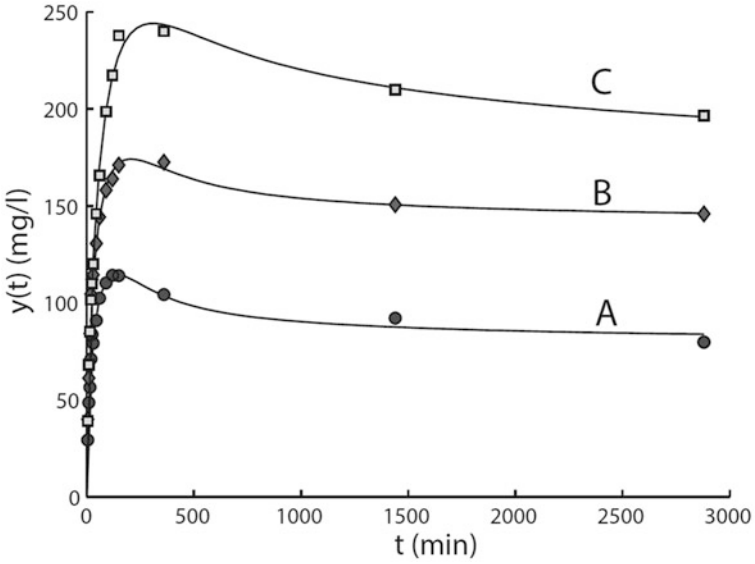


Fig. 5.12 Fitings of equation 5.27 to three data sets of carbamazepine tablets in presence of TPGS at 10° C exhibiting supersaturated dissolution profiles. Key: (**A**) 0.5 mM TPGS; (**B**) 2 mM TPGS; (**C**) 4 mM TPGS

fitted to the experimental data of carbamazepine dissolution in presence of TPGS at 10° C:

$$\dot{y}(t) = k\left[\lambda + (1+t)^{-h}\right]\left[\frac{q_0}{V} - y(t)\right]^a - k_{-1}y(t) \tag{5.27}$$

where k is a constant in $(\text{time})^{b-1}$ units and λ is a constant in $(\text{time})^{-h}$ units. Figure 5.12 shows the fittings of the above equation to three data sets of carbamazepine dissolution.

5.3 Modeling Random Effects

In all previous dissolution models described in Sections 5.1 and 5.2, the variability of the particles (or media) is not directly taken into account. In all cases, a unique constant (cf. Sections 5.1, 5.1.1, and 5.1.2) or a certain type of time dependency in the dissolution rate "constant" (cf. Sections 5.1.3, 5.2.1, and 5.2.2) is determined at the commencement of the process and fixed throughout the entire course of dissolution. Thus, in essence, all these models are deterministic. However, one can also assume that the above variation in time of the rate or the rate coefficient can take place randomly due to unspecified fluctuations in the heterogeneous properties

of drug particles or the structure/function of the dissolution medium. Lansky and Weiss have proposed [134] such a model assuming that the rate of dissolution $k(t)$ is stochastic and is described by the following equation:

$$k(t) = \overline{k} + \sigma\xi(t),$$

where $\overline{k}$ is the deterministic part of the dissolution rate "constant," $\xi(t)$ is Gaussian white noise, and $\sigma > 0$ is its amplitude. According to the definition of this equation, the "constant" $\overline{k}$ represents the mean of $k(t)$.

The stochastic nature of $k(t)$ allows the description of the fraction of dose dissolved, $\varphi(t)$, in the form of a stochastic differential equation if coupled with the simplest dissolution model described by (5.16), assuming complete dissolution ($\theta = 1$):

$$\mathrm{d}\varphi(t) = \overline{k}[1 - \varphi(t)]\,\mathrm{d}t + \sigma\xi(t)[1 - \varphi(t)]\,\mathrm{d}B(t), \tag{5.28}$$

where the symbol $\varphi(t)$ is used here to denote the random nature of the process, while $\mathrm{d}B(t)$ comes from the Brownian motion since the noise $\xi(t)$ is the formal derivative of the Brownian motion, $\dot{B}(t)$. The solution of (5.28) gives

$$\varphi(t) = 1 - \exp\left[-\left(\overline{k} + \frac{1}{2}\sigma^2\right)t - \sigma B(t)\right].$$

A discretized version of (5.28) can be used to perform Monte Carlo simulations using different values of σ and generate $\varphi(t)$-time profiles [134]. The random fluctuation of these profiles becomes larger as the value of σ increases.

Stochastic variation may be introduced in other models as well. In this context, Lansky and Weiss [134] have also considered random variation for the parameter k^* of the interfacial barrier model (5.20).

5.4 Homogeneity vs. Heterogeneity

Lansky and Weiss defined [135] the classical dissolution first-order model in terms of the fraction of dose dissolved, $\varphi(t)$ (equation 5.16 assuming $\theta = 1$),

$$\dot{\varphi}(t) = k[1 - \varphi(t)],\ \varphi(0) = 0,$$

as the simplest homogeneous case, since the fractional dissolution rate function $k(t)$ derived from the above equation,

$$k(t) = \frac{\dot{\varphi}(t)}{1 - \varphi(t)},$$

is constant throughout the dissolution process. In physical terms, the homogeneous model dictates that each drug molecule has equal probability to enter solution during the entire course of the dissolution process. Plausibly, the various dissolution models have different time-dependent functional forms of $k(t)$. Accordingly, all these models were termed heterogeneous since the time dependence of the functions $k(t)$ denotes that the probability to enter solution is not identical for all drug molecules. To quantify the departure from the homogeneous case, Lansky and Weiss proposed [135] the calculation of the Kullback–Leibler information distance $Dist\,(f, \varphi)$ as a measure of heterogeneity of the function $f(t)$ from the homogeneous exponential model $\varphi\,(t)$ derived from the previous equation:

$$Dist\,(f, \varphi) = \int_0^\infty f(t) \ln \frac{f(t)}{\varphi\,(t)} \mathrm{d}t.$$

This measure of heterogeneity generalizes the notion of heterogeneity as a departure from the classical first-order model initially introduced [123] for the specific case of the Weibull function. In addition, the above equation can also be used for comparison between two experimentally obtained dissolution profiles [135].

The comparison of dissolution curves based on the calculation of $Dist\,(f, \varphi)$ is model-independent; however, other model-dependent comparative approaches have been proposed [136]. Caution should be exercised, though, when comparison of estimates of the parameters obtained from various models is attempted in the context of heterogeneity assessment. For example, the valid use of (5.15) for the homogeneous case presupposes that the amount needed to saturate the medium is exactly equal to the dose used in actual practice, i.e., $\theta = 1$ [136]. Recently, Lansky and Weiss presented [137] in a concise form the results of their recent studies [124, 134]. The empirical and semiempirical models for drug dissolution were reviewed and classified in five groups: first-order model with a time lag, models for limited solubility of drug, models of heterogeneous compound, Weibull and inverse Gaussian models, and models defined on a finite time window. In this contribution, the properties of models were investigated, the parameters were discussed, and the role of drug heterogeneity was studied.

5.5 Comparison of Dissolution Profiles

The comparison of dissolution profiles is of interest for both research and regulatory purposes. Several methods, which can be roughly classified as (1°) statistical approaches, (2°) model-dependent, and (3°) model-independent methods, have been reported in the literature for the comparison of dissolution profiles [138–140]. The statistical approaches are based on the analysis of variance, which is used to test the hypothesis that the two profiles are statistically similar. The model-dependent methods are mainly used for clarifying dissolution or release mechanisms under various experimental conditions and rely on the statistical

comparison of the estimated parameters after fitting of a dissolution model (e.g., the Weibull model) to the raw data. The model-dependent methods can be applied to dissolution profiles with nonidentical dissolution sampling schemes, while the model-independent methods require identical sampling points since they are based on pairwise procedures for the calculation of indices (factors) from the individual raw data of two profiles. Two of these factors, namely, the *difference factor* f_1 and the *similarity factor* f_2, have been adopted by the regulatory agencies and have been included in the relevant dissolution Guidances for quality control testing [141–143]. Each one of these factors is calculated from the two mean dissolution profiles and is being used as a point estimate measure of the (dis)similarity of the dissolution profiles.

The difference factor f_1 [141] measures the relative error (as a percentage) between two dissolution curves over all time points:

$$f_1 = 100 \frac{\sum_{i=1}^{m} |R_i - T_i|}{\sum_{i=1}^{m} R_i}. \tag{5.29}$$

where m is the number of data points, R_i and T_i are the percentage of drug dissolved for the reference and test products at each time point i, respectively.

The similarity factor f_2 [141–143] is a logarithmic reciprocal transformation of the sum of squared errors and is a measurement of the percentage similarity in the dissolution between the two dissolution curves:

$$f_2 = 50 \log \left\{ 100 \left[1 + \frac{1}{m} \sum_{i=1}^{m} (R_i - T_i)^2 \right]^{-0.5} \right\}. \tag{5.30}$$

Both factors take values in the range $0-100$ assuming that the percentage dissolved values for the two products are not higher than 100%. When no difference between the two curves exist, i.e., at all time points $R_i = T_i$, then $f_1 = 0$ and $f_2 = 100$. On the other hand, when the maximum difference between the two curves exists, i.e., at all time points $|R_i - T_i| = 100$, then $f_1 = 100$ and $f_2 = 0$.

The calculation of the factors from the mean profiles of the two drug products presupposes that the variability at each sample time point is low. Thus, for immediate release formulations, the FDA guidance [141] allows a coefficient of variation of no more than 20% for the early data points (e.g., 10 or 15 min), while a coefficient of variation less than 10% is required for the other time points. According to the guidances [141, 143], when batches of the same formulation are compared, a difference up to 10% at all sample points is considered acceptable. On the basis of this boundary, the acceptable range of values derived from (5.29) and (5.30) for f_1 is $0 - 15$ [141] and for f_2 is $50 - 100$ [141, 143]. From a technical point of view, the following recommendations are quoted in the guidances [141, 143] for the calculation of f_1 and f_2 as point estimates:

1. a minimum of three time points (zero excluded),
2. 12 individual values for every time point for each formulation,
3. not more than one mean value of $> 85\%$ dissolved for each formulation.

Note that when more than 85% of the drug is dissolved from both products within 15 minutes, dissolution profiles may be accepted as similar without further mathematical evaluation. For the sake of completeness, one should add that some concerns have been raised regarding the assessment of similarity using the direct comparison of the f_1 and f_2 point estimates with the similarity limits [144–146]. Attempts have been made to bring the use of the similarity factor f_2 as a criterion for assessment of similarity between dissolution profiles in a statistical context using a bootstrap method [145] since its sampling distribution is unknown.

Although there are some differences between the European [143] and the US guidance [141, 142], e.g., the composition of the dissolution media, it should be pointed out that both recommend dissolution studies as quality assurance tests as well as for bioequivalence surrogate inference. The latter aspect is particularly well developed in the FDA guidance [142] in the framework of the Biopharmaceutics Classification System (BCS), which is treated in Section 6.6.1.

Chapter 6
Oral Drug Absorption

The right drug for the right indication in the right dosage to the right patient.

Anonymous

The understanding and the prediction of oral drug absorption are of great interest for pharmaceutical drug development. Obviously, the establishment of a comprehensive framework in which the physicochemical properties of drug candidates are quantitatively related to the extent of oral drug absorption will accelerate the screening of candidates in the discovery/preclinical development phase. Besides, such a framework will certainly help regulatory agencies in developing scientifically based guidelines in accord with a drug's physicochemical properties for various aspects of oral drug absorption, e.g., dissolution, in vitro–in vivo correlations, biowaivers of bioequivalence studies.

However, the complex interrelationships among drug properties and processes in the gastrointestinal tract make the prediction of oral drug absorption a difficult task. In reality, drug absorption is a complex process dependent upon drug properties such as solubility and permeability, formulation factors, and physiological variables including regional permeability differences, pH, luminal and mucosal enzymes, and intestinal motility, among others. Despite this complexity, various qualitative and quantitative approaches have been proposed for the estimation of oral drug absorption. In all approaches discussed below the drug movement across the epithelial layer is considered to take place transcellularly since transcellular passive diffusion is the most common mechanism of drug transport.

The absorption models described in this chapter can be divided as follows:

- pseudoequilibrium models,
- mass balance approaches,
- dynamic models,
- heterogeneous approaches, and
- models based on chemical structure.

The last section of this chapter is devoted to the regulatory aspects of oral drug absorption and in particular to the BCS and the relevant FDA guideline. The recent regulatory-scientific advances related to BCS as well as the Biopharmaceutics Drug

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_6

Disposition Classification System (BDDCS) are described too. At the very end of the chapter, we mention the difference between randomness and chaotic behavior as sources of the variability encountered in bioavailability and bioequivalence studies.

6.1 Pseudoequilibrium Models

These models assume that oral drug absorption takes place under equilibrium conditions. Spatial or temporal aspects of the drug dissolution, transit and uptake and the relevant physiological processes in the gastrointestinal tract are not taken into account. Only drug-related properties are considered as the key parameters controlling the absorption process.

6.1.1 The pH-Partition Hypothesis

Back in the 1940s, physiologists were the first to realize that in contrast to the capillary walls, with their large and unselective permeability, cell membranes present a formidable barrier to the diffusion of small molecules. A prominent scientist, M.H. Jacobs, in 1940 [147] was the first who studied the cell permeability of weak electrolytes and described quantitatively the nonionic membrane permeation of solutes. This observation initiated a number of specific studies [148–153] during the 1950s on the mechanisms of gastrointestinal absorption of drugs. The results of these studies formed the basis for the pH-partition hypothesis, which relates the dissociation constant, lipid solubility, and the pH at the absorption site with the absorption characteristics of various drugs throughout the gastrointestinal tract. Knowledge of the exact ionization of a drug is of primary importance since the un-ionized form of the drug, having much greater lipophilicity than the ionized form, is much more readily absorbed. Consequently, the rate and extent of absorption are principally related to the concentration of the un-ionized species. Since the pH in the gastrointestinal tract varies, the Henderson–Hasselbach equations for the ionization of acids,

$$\mathrm{pH} = pK_a + \log\left[\frac{\text{ionized-concentration}}{\text{un-ionized-concentration}}\right],$$

and bases,

$$\mathrm{pH} = pK_a + \log\left[\frac{\text{un-ionized-concentration}}{\text{ionized-concentration}}\right],$$

relate the fraction of the un-ionized species with the regional pH and the pK_a of the compound.

Most of the gastrointestinal absorption studies were found to be in accord with the principles of the pH-partition hypothesis. However, several deviations were noted and attributed to the unstirred water layer, the microclimate pH, and the mucus coat adjacent to the epithelial cell surface [154–156].

Although the pH-partition hypothesis relies on a quasi-equilibrium transport model of oral drug absorption and provides only qualitative aspects of absorption, the mathematics of passive transport assuming steady diffusion of the un-ionized species across the membrane allows quantitative permeability comparisons among solutes. As discussed in Chapter 2, (2.19) describes the rate of transport under sink conditions as a function of the permeability P, the surface area $\mathcal{A}$ of the membrane, and the drug concentration $c\,(t)$ bathing the membrane:

$$\dot{q}\,(t) = P\mathcal{A}c\,(t)\,. \tag{6.1}$$

The proportionality between the rate of transport and permeability in (6.1) shows the importance of the latter parameter in the transcellular passive gastrointestinal absorption of drugs. Strictly speaking, one should utilize an estimate of the effective permeability (P_{eff}) [157] in (6.1) for predicting oral absorption potential of compounds. However, the methods for the estimation of P_{eff} are invasive, laborious, and time-consuming. Alternatively, various measures of lipophilicity such as the octanol/water partition coefficient ($\log P_c$) [158] and the distribution coefficient ($\log D$) [159] have been used as surrogates for predicting the oral absorption potential of compounds since permeability is mainly dependent on membrane partitioning.

6.1.2 Absorption Potential

In 1985 a major step in the theoretical analysis of oral drug absorption phenomena took place [160], when solubility and dose were also taken into account for the estimation of the absorption potential AP of a drug apart from the pH-partition hypothesis related parameters (lipophilicity, and degree of ionization). According to this concept, the AP is related proportionally to the octanol/water partition coefficient P_c, the fraction of the un-ionized species f_{un}, at pH= 6.5, and the physiological solubility c_s of the drug and inversely proportional to the dose q_0:

$$AP = \log\left[\frac{P_c f_{un} c_s V}{q_0}\right] = \log\left[\frac{P_c f_{un}}{\theta}\right]. \tag{6.2}$$

The logarithmic expression in the definition of AP has no physicochemical basis and is used for numerical reasons only; pH= 6.5 was selected as the representative pH of small intestines, where most of the absorption of drugs takes place. The incorporation of the terms P_c and f_{un} in the numerator of (6.2) means that the pH-partition hypothesis governs gastrointestinal absorption. Plausibly, AP was

considered proportional to solubility and inversely proportional to the dose in accord with classical dissolution–absorption considerations. The volume term V corresponds to the small-intestinal volume content, which was set arbitrarily equal to 250 ml; moreover, the use of the term V makes the AP dimensionless. The ratio q_0/c_sV was defined as the dimensionless solubility–dose ratio in Section 5.1.4 and it was denoted by θ.

The validity of the approach based on (6.2) was proven when the fraction of dose absorbed, F_a, was found to increase with AP for several drugs with a wide variety of physicochemical properties and various degrees of extent of absorption [160]. Additional support for the AP concept was provided by a three-dimensional plot of F_a as a function of the ionization–solubility/dose term (f_{un}/θ) and the octanol/water partition coefficient P_c [161]. In fact, because of the recent interest in the apparent permeability estimates P_{app} measured in the in vitro Caco-2 monolayer system, it was suggested that P_{app} can replace the octanol/water partition coefficient P_c in (6.2) [161].

Although the AP concept is a useful indicator of oral drug absorption, its qualitative nature does not allow the derivation of an estimate for F_a. A quantitative version of F_a as a function of AP was published in 1989. It was based on the equilibrium considerations used for the derivation of AP and the fundamental physicochemical properties in (6.2) with the implied competing intestinal absorption and nonabsorption processes [162]. This quantitative AP concept relies on (6.3), where a nonlogarithmic expression for AP is used:

$$F_a = \frac{(AP)^2}{(AP)^2 + f_{un}\,(1 - f_{un})}. \tag{6.3}$$

Based on physiological-physicochemical arguments, constraints were proposed for P_c, i.e., to be set equal to 1000 when $P_c > 1000$ and θ equal to 1 when $\theta < 1$. Equation (6.3) is, in reality, the first ever published explicit relationship between F_a and physicochemical drug properties. It was used to classify drugs according to their solubility, permeation, and ionization characteristics [162]. Moreover, (6.3) was monoparameterized:

$$F_a = \frac{Z^2}{1 + Z^2}, \tag{6.4}$$

where

$$Z = \frac{AP}{\sqrt{f_{un}\,(1 - f_{un})}}.$$

Equation (6.4) was used for fitting purposes using AP and F_a data reported in the literature and applying the constraints mentioned above for P_c and θ in the calculation of AP, Figure 6.1.

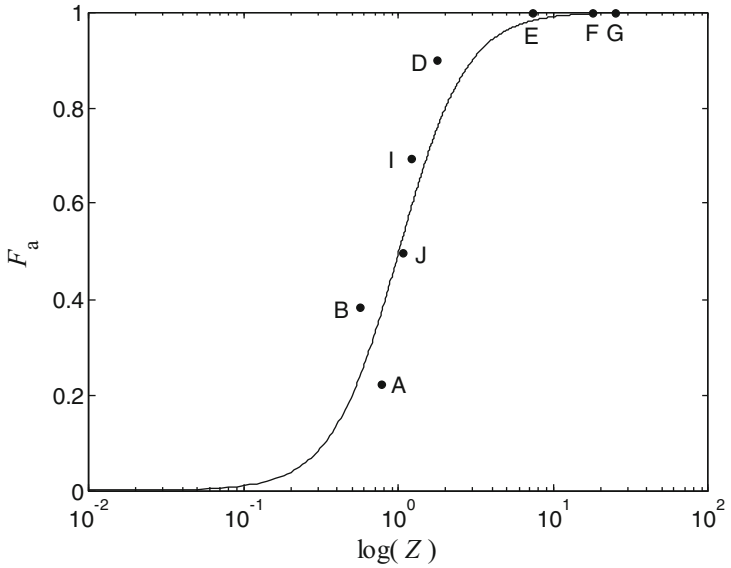


Fig. 6.1 Plot of the fraction of dose absorbed for various drugs as a function of Z. Key: **A**, acyclovir; **B**, chlorothiazide solution; **D**, hydrochlorothiazide; **E**, phenytoin; **F**, prednisolone; **G**, digoxin (Lanoxicaps); **I**, cimetidine; **J**, mefenamic acid

A number of modifications in the solubility and the partition coefficient terms of the *AP* have also been proposed in the literature [163–165]. According to these authors the modified absorption potential expressions can be considered better predictors of the passive absorption of drugs than the original *AP*. The most recent approach [165] relies on a single absorption parameter defined as the ratio of the octanol/water partition coefficient to the luminal oversaturation number. The latter is equal to the solubility-normalized dose for suspensions and equal to unity for solutions.

A relevant simple model was used to estimate the maximum absorbable dose (*MAD*) [166]. It takes into account the permeability, expressed in terms of a first-order rate constant k_a, the solubility c_s of the drug, and two physiological variables, the dissolution-intestinal volume V arbitrarily set to 250 ml, and the duration of gastrointestinal absorption t_a for 6 h:

$$MAD = Vc_s k_a t_a.$$

This model assumes gastrointestinal absorption from a saturated solution of the drug (hypothetically maintained at a constant saturated value) for a time period equal to 6 h.

6.2 Mass Balance Approaches

These approaches place particular emphasis on the spatial aspects of the drug absorption from the gastrointestinal tract. The small intestine is assumed to be a cylindrical tube with fixed dimensions where the drug solution or suspension follows a homogeneous flow. Mass balance relationships under steady-state assumptions are used to estimate the fraction of dose absorbed as a function of the drug properties and of physiological parameters.

6.2.1 Macroscopic Approach

In the early 1990s the research group of G. Amidon in Ann Arbor applied mass balance approaches to the analysis of drug intestinal absorption [55, 56]. The small intestine is assumed to be a cylindrical tube with physiologically relevant dimensions (radius R and length L), while a constant volumetric flow rate Q describes the transit process of the intestinal contents, Figure 6.2. The *macroscopic approach* [55] refers mainly to highly soluble compounds. The incorporation of the dissolution step as an important part of the absorption process is treated in Section 6.2.2 under the heading *microscopic approach* [56].

The macroscopic approach under the steady-state assumption provides estimates for the fraction of dose absorbed F_a for the three cases, which refer to the magnitude of c_0 and c_{out} in Figure 6.2 relative to drug solubility c_s, namely:

1. Case I: $c_0 \leq c_s$ and $c_{out} \leq c_s$ (the drug is in solution throughout the transit process);
2. Case II: $c_0 > c_s$ and $c_{out} \leq c_s$ (solid drug at inlet; concentration reaches solubility at a certain point and diminishes thereafter);
3. Case III: $c_0 > c_s$ and $c_{out} > c_s$ (solid drug exists at both ends of the tube).

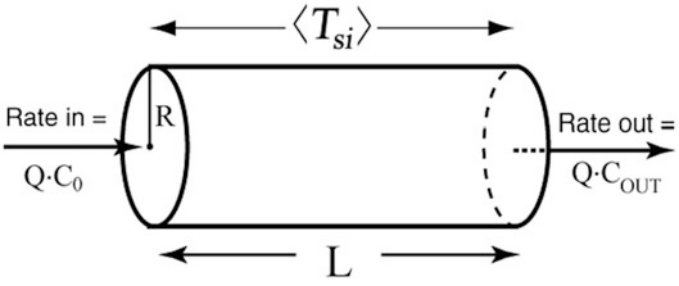


Fig. 6.2 The small intestine is modeled as a homogeneous cylindrical tube of length L and radius R. c_0 and c_{out} are the inlet and outlet drug concentrations, respectively. The other symbols are defined in the text

Irrespective of the specific case considered, the general mass balance relationship for the system depicted in Figure 6.2 under the steady-state assumption is

$$Q\left(c_0 - c_{out}\right) = 2\pi R P_{eff} \int_0^L c\left(z\right) \mathrm{d}z,$$

where P_{eff} is the effective permeability of drug and $\mathrm{d}z$ the infinitesimal axial length. The fraction of dose absorbed F_a can be expressed in terms of c_0 and c_{out} using the previous equation:

$$\begin{aligned} F_a &= 1 - \frac{c_{out}}{c_0} = \frac{2\pi R P_{eff}}{Q c_0} \int_0^L c\left(z\right) \mathrm{d}z \\ &= \frac{2\pi R P_{eff} L}{Q} \int_0^1 c^*\left(z^*\right) \mathrm{d}z^* = 2A_n \int_0^1 c^*\left(z^*\right) \mathrm{d}z^*. \end{aligned} \tag{6.5}$$

The last two integrals of the previous equation are expressed in dimensionless variables, $c^* = c/c_0$, $z^* = z/L$ with normalized limits $(0, 1)$, while the symbol A_n is the *absorption number* of the drug:

$$A_n \triangleq \frac{\pi R L}{Q} P_{eff} = \langle T_{si} \rangle \frac{P_{eff}}{R}. \tag{6.6}$$

The first fraction of the previous equation shows that A_n is exclusively determined by the effective permeability P_{eff} of drug since all other variables are species-dependent physiological parameters. In terms of characteristic times, the A_n of a drug can also be defined as the ratio of the mean small-intestinal transit time $\langle T_{si} \rangle$, to its absorption time R/P_{eff}.

For the calculation of F_a, one should first express the dimensionless concentration $c^*\left(z^*\right)$ as a function of z^*, for each one of the three cases considered above, and then integrate (6.5).

1. For case I, the concentration profile $c^*\left(z^*\right)$ diminishes exponentially as a function of distance z^* assuming the complete radial mixing model [167] in the tube,

$$c^*\left(z^*\right) = \frac{c_{out}}{c_0} = \exp\left(-2A_n z^*\right),$$

and for the fraction of dose absorbed,

$$F_a = 1 - \exp\left(-2A_n\right).$$

This last equation shows that when the drug is in solution throughout the transit process and $c_0 \leq c_s$ and $c_{out} \leq c_s$, then F_a is dependent exclusively and exponentially on A_n. According to this equation, large values of A_n ensure complete absorption for this type of drugs.

2. For case III, the concentration c_{out} can be considered equal to the solubility since $c_0 > c_s$ and $c_{out} > c_s$; therefore

$$c^*\left(z^*\right) = \frac{c_s}{c_0} = \frac{1}{\theta}, \tag{6.7}$$

and for the fraction of dose absorbed,

$$F_a = \frac{2A_n}{\theta}. \tag{6.8}$$

Although this equation indicates that F_a is proportional to A_n and inversely proportional to θ, this should be judged with caution since the conditions of case III, expressed in terms of concentration, are physically irrelevant ($c_0 > c_s$ and $c_{out} > c_s$). In addition, the use of (6.7) for the derivation of (6.8) assumes instantaneous dissolution in order to maintain the value of c_s constant throughout the transit process.

3. Case II can be viewed as a hybrid of cases I and III. As long as $c_0 > c_s$, the conditions assumed for case III are prevailing. Then, using a simple mass balance equation up to the temporal (spatial) point when c_0 reaches solubility ($c_0 = c_s$) and (6.7), the fraction absorbed F_{a1} can be calculated as

$$F_{a1} = \frac{c_0 V - c_s V}{c_0 V} = 1 - \frac{1}{\theta}.$$

Beyond this spatiotemporal point until the drug exits from the tube, the inequality $c_0 < c_s$ holds and therefore the fraction absorbed F_{a2} in this region follows the results obtained for case I conditions:

$$F_{a2} = \frac{1}{\theta}\left[1 - \exp\left(-2A_n + \theta - 1\right)\right].$$

Consequently, the total fraction of dose absorbed F_a is the sum of F_{a1} and F_{a2}:

$$F_a = 1 - \frac{1}{\theta}\exp\left(-2A_n + \theta - 1\right).$$

The most significant result of the macroscopic approach was derived from the analysis of case I conditions. It was found that the absorption number A_n and in particular its major determinant, the effective permeability, control the intestinal absorption of drugs. This observation triggered a large number of studies, and in recent years several attempts have been made to model the fraction of dose absorbed, F_a, with experimental in situ and in vitro models such as cell cultures (Caco-2, HT-29, and MDCK) [168–170] and artificial membranes (IAM, PAMPA) [171]. The aim of these studies is to find a correlation between the apparent permeability estimates P_{app} measured in these systems and the experimental F_a values. The most popular among these systems is the in vitro Caco-2 monolayer system [172], which is a donor–receptor compartment apparatus separated by a cell monolayer grown

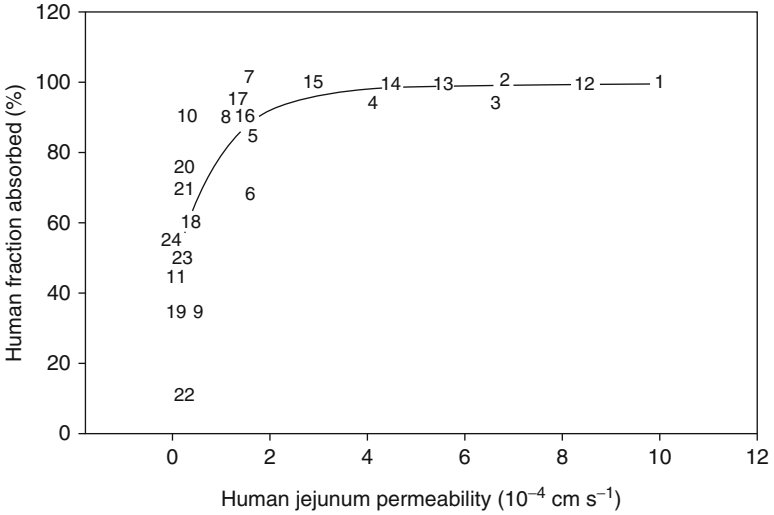


Fig. 6.3 Plot of the fraction of dose absorbed (in %) of various drugs as a function of the permeability estimates in the Caco-2 system. Key: 1, D-glucose; 2, verapamil; 3, piroxicam; 4, phenylalanine; 5, cyclosporin; 6, enalapril; 7, cephalexim; 8, losartan; 9, lisinopril; 10, amoxicillin; 11, methyldopa; 12, naproxen; 13, antipyrine; 14, desipramine; 15, propanolol; 16, amiloride; 17, metoprolol; 18, terbutaline; 19, mannitol; 20, cimetidine; 21, ranitidine; 22, enalaprilate; 23, atenolol; 24, hydrochlorothiazide

on a porous polycarbonate filter and is used to estimate the apparent permeability of compounds. In reality, an estimate for P_{app} is obtained from the experimental permeation data using (6.1) and solving it in terms of P; the flux rate $\dot{q}(t)$ is obtained from the slope of the receptor chamber solute mass vs. time plot, while $\mathcal{A}$ is the cross-sectional area of cell surface and $c(t) = c_0$ is the initial solute concentration in the donor compartment. Extensive research in the passive transport mechanisms of a great number of compounds in cell culture monolayers indicates that an apparent permeability estimate in the range of $2 \times 10^{-6} - 10^{-5}$ cm s^{-1} [172–174] ensures complete absorption of the solute provided that absorption is not solubility- and/or dissolution-limited, Figure 6.3.

6.2.2 *Microscopic Approach*

This approach deals with the analysis of intestinal absorption of poorly soluble drugs, administered as suspensions, assuming that drug particles are spheres of the same initial radius size ρ_0. The resulting mathematical model [56] assumes complete radial mixing, takes into account drug dissolution, transit, and uptake, and relies on the homogeneous cylindrical intestinal tube depicted in Figure 6.2. Under the steady-state assumption, mass balance relationships for the drug processes in both

solid and solution phase are considered in a volume element of the intestine of axial length dz. Two differential equations expressed in dimensionless variables govern the reduction of the radius $\rho(z)$ of the particles from their initial value $\rho_0$and the change of the luminal concentration of the drug $c(z)$:

$$\begin{aligned} \frac{\mathrm{d}\rho^*(z^*)}{\mathrm{d}z^*} &= -\frac{D_n}{3}\frac{1-c^*(z^*)}{\rho^*(z^*)} && \rho^*(0) = 1, \\ \frac{\mathrm{d}c^*(z^*)}{\mathrm{d}z^*} &= \theta D_n \rho^*(z^*)\left[1 - c^*(z^*)\right] - 2A_n c^*(z^*) && c^*(0) = 0, \end{aligned} \tag{6.9}$$

where $z^* = z/L$, $c^*(z^*) = c(z^*)/c_s$, $\rho^*(z^*) = \rho(z^*)/\rho_0$, and D_n is the dissolution number defined by the following relationship:

$$D_n \triangleq \frac{(\mathcal{D}/\rho_0)\, c_s \left(4\pi\rho_0^2\right)\left(\pi R^2 L\right)}{Q\left(4\pi\rho_0^3\varrho\right)},$$

where $\mathcal{D}$ is the diffusivity and ϱ is the density of the drug. Using a mass balance relationship for the solid and solution phases at the outlet of the tube ($\rho^* = 1$), the following equation is obtained for the fraction of dose absorbed, F_a:

$$F_a = 1 - \left(\rho^*|_{z^*=1}\right)^3 - \frac{1}{\theta}\left(c^*|_{z^*=1}\right).$$

This equation can be used in conjunction with (6.9) for the estimation of F_a. The microscopic approach points out clearly that the key parameters controlling drug absorption are three dimensionless numbers, namely, absorption number A_n, dissolution number D_n, and θ. The first two numbers are the determinants of membrane permeation and drug dissolution, respectively, while θ reflects the ratio of the dose administered to the solubility of drug.

6.3 Dynamic Models

These models are dependent on the temporal variable associated with the drug transit along the small intestine. Drug absorption phenomena are assumed to take place in the time domain of the physiological mean transit time. For those dynamic models that rely on diffusion/dispersion principles both the spatial and temporal variables are important in order to simulate the spatiotemporal profile of the drug in the intestinal lumen.

6.3.1 Compartmental Models

The compartmental approach to the process of a drug passing through the gastrointestinal tract has been used to simulate and explain oral drug absorption. The simplest approach relies on a single mixing tank model of volume V where the

drug has a uniform concentration while a flow rate Q is ascribed to the contents of the tank. Thus, the ratio V/Q corresponds to the time period beyond which drug dissolution and/or absorption is terminated. In other words, it is equivalent to the small-intestinal transit time for the homogeneous tube model. Similarly, the ratio Q/V indicates the first-order rate constant for drug removal from the absorption sites. One or two mixing tanks in series have been employed for the study of various oral drug absorption phenomena [54, 175, 176].

Mixing tanks in series with linear transfer kinetics from one to the next with the same transit rate constant k_t have been utilized to obtain the characteristics of flow in the human small intestine [177, 178]. The differential equations of mass transfer in a series of m compartments constituting the small intestine for a nonabsorbable and nondegradable compound are

$$\dot{q_i}(t) = k_t q_{i-1}(t) - k_t q_i(t), \quad i = 1 : m, \tag{6.10}$$

where $q_i(t)$ is the amount of drug in the i-th compartment. The rate of exit of the compound from the small intestine is

$$\dot{q_m}(t) = -k_t q_m(t). \tag{6.11}$$

Solving the system of (6.10) and (6.11) in terms of the fraction of dose absorbed, we obtain

$$F_a = \frac{q_m(t)}{q_0} = 1 - \exp(-k_t t)\left[1 + k_t t + \frac{(k_t t)^2}{2} + \ldots + \frac{(k_t t)^{m-1}}{(m-1)!}\right]. \tag{6.12}$$

Analysis of experimental human small-intestine transit time data collected from 400 studies revealed a mean small-intestinal transit time $\langle T_{si} \rangle = 199\,\text{min}$ [177]. Since the transit rate constant k_t is inversely proportional to $\langle T_{si} \rangle$, namely, $k_t = m/\langle T_{si} \rangle$, (6.12) was further fitted to the cumulative curve derived from the distribution frequency of the entire set of small-intestinal transit time data in order to estimate the optimal number of mixing tanks. The fitting results were in favor of seven compartments in series and this specific model, (6.10) and (6.11) with $m = 7$, was termed the *compartmental transit model*.

The incorporation of a passive absorption process in the compartmental transit model led to the development of the *compartmental absorption transit model* (CAT) [179]. The rate of drug absorption in terms of mass absorbed $q_a(t)$ from the small intestine of the compartmental transit model is

$$\dot{q_a}(t) = k_a \sum_{i=1}^{7} q_i(t),$$

where k_a is the first-order absorption rate constant. Then, the fraction of dose absorbed F_a, using the previous equation, is

$$F_a = \frac{q_a(t)}{q_0} = \frac{k_a}{q_0} \sum_{i=1}^{7} \int_0^{\infty} q_i(t)\, \mathrm{d}t. \tag{6.13}$$

The solution of (6.12) and (6.13) yields

$$F_a = 1 - \left(1 + \frac{k_a}{k_t}\right)^{-7}.$$

Recall that k_t is equal to $7/\langle T_{si} \rangle$, while k_a can be expressed in terms of the effective permeability and the radius R of the small intestine [56]:

$$k_a = \frac{2P_{eff}}{R}. \tag{6.14}$$

The previous equation can be written as

$$F_a = 1 - \left(1 + \frac{2P_{eff}\langle T_{si} \rangle}{7R}\right)^{-7}.$$

The CAT model presupposes that dissolution is instantaneous and therefore the kinetics of the permeation step control the gastrointestinal absorption of drug. This is reflected in the previous equation, which indicates that the effective permeability is the sole parameter controlling the intestinal absorption of highly soluble drugs.

Due to its compartmental nature, the CAT model can easily be coupled with the disposition of drug in the body using classical pharmacokinetic modeling. In this respect the CAT model has been used to interpret the saturable small-intestinal absorption of cefatrizine in humans [179].

The CAT model was further modified to include pH-dependent solubility, dissolution/precipitation, absorption in the stomach or colon, first-pass metabolism in gut or liver, and degradation in the lumen. Physiological and biochemical factors such as changes in absorption surface area, transporter, and efflux protein densities have also been incorporated. This advanced version of CAT, called ACAT [180], has been formulated in a commercially available simulation software product under the trademark name GastroPlus™. A set of differential equations, which is solved by numerical integration, is used to describe the various drug processes of ACAT as depicted in Figure 6.4.

6.3.2 Convection–Dispersion Models

The use of convection–dispersion models in oral drug absorption was first proposed in the early 1980s [181, 182]. The small intestine is considered a one-dimensional tube that is described by a spatial coordinate z that represents the axial distance

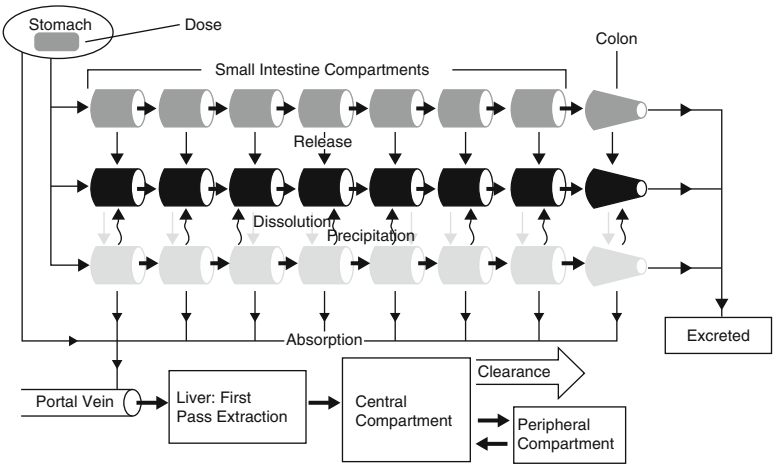


Fig. 6.4 Schematic of the ACAT model. Reprinted from [180] with permission from Elsevier

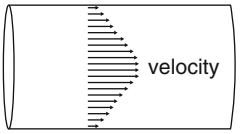


Fig. 6.5 The velocity of the fluid inside the tube is larger near the axis and much smaller near the walls. This is considered to be the main factor for the dispersion of the distribution of the drug

from the stomach. In addition, the tube contents have constant axial velocity v and constant dispersion coefficient D, which arises from molecular diffusion, stirring due to the motility of the intestines, and Taylor dispersion due to the difference of the axial velocity at the center of the tube compared with the tube walls (Figure 6.5). The small-intestine transit flow for a nonabsorbable and nondegradable compound in this type of model is described by [177, 182]

$$\frac{\partial c\,(z,t)}{\partial t} = D\frac{\partial^2 c\,(z,t)}{\partial z^2} - v\frac{\partial c\,(z,t)}{\partial z}, \tag{6.15}$$

where $c\,(z,t)$ is the concentration. An analytical solution of this equation can be obtained if one assumes that the stomach operates as an infinite reservoir with constant output rate in terms of concentration and volume. Under these assumptions, the following analytical solution was obtained [182]:

$$\frac{c\,(z,t)}{c_0} = \frac{1}{2}\left[\operatorname{erf} c\left(\frac{z}{\sqrt{4Dt}} - \sqrt{\frac{v^2 t}{4D}}\right) + \exp\left(\frac{vz}{D}\right)\operatorname{erf} c\left(\frac{z}{\sqrt{4Dt}} + \sqrt{\frac{v^2 t}{4D}}\right)\right], \tag{6.16}$$

where erf c is the complementary error function defined by

$$\operatorname{erf} c\,(x) \triangleq 1 - \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-z^2\right) \mathrm{d}z.$$

Equation (6.16) allows one to generate the axial profile of normalized concentration $c\,(z,t)\,/c_0$ at different times, Figure 6.6A. The second term in the parentheses

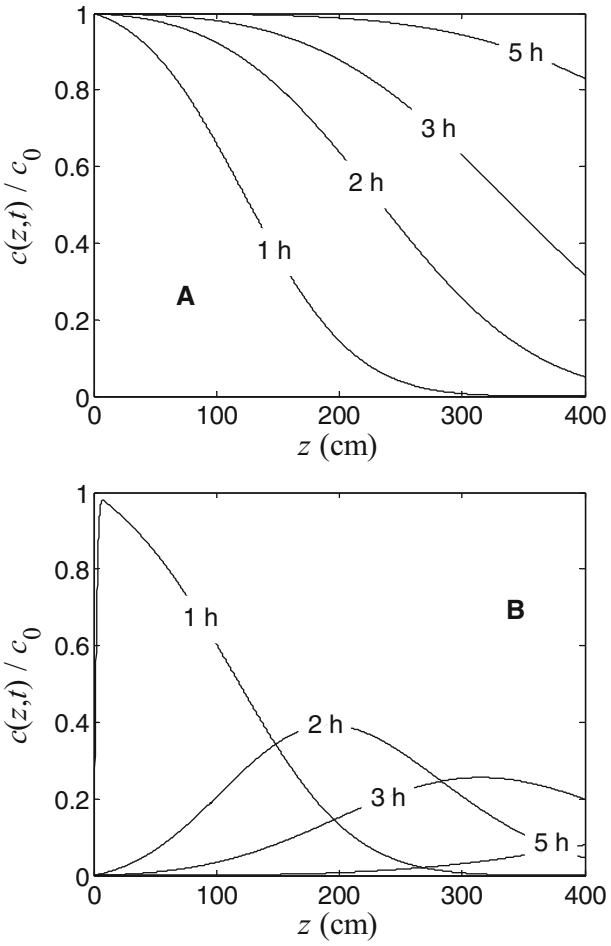


Fig. 6.6 Axial profile snapshots of normalized concentration (with respect to the constant input concentration) inside the intestinal lumen, at various times. (**A**) (6.16) is used, with $D = 0.78\,\mathrm{cm^2\,s^{-1}}$, $v = 1.76\,\mathrm{cm\,min^{-1}}$, and a constant-concentration infinite reservoir input. (**B**) the analytical solution of (6.17) with initial condition $c\,(z,0) = 0$ is used, with $D = 0.78\,\mathrm{cm^2\,s^{-1}}$, $v = 1.76\,\mathrm{cm\,min^{-1}}$, $k_a = 0.18\,\mathrm{h^{-1}}$, and a constant-concentration reservoir input, applied only for the first hour, $t_\circ = 1\,\mathrm{h}$

of (6.16) is relatively small compared to the first; therefore, (6.16) can be approximated by the following:

$$\frac{c\,(z,t)}{c_0} = \frac{1}{2}\,\mathrm{erf}\,c\left(\frac{z}{\sqrt{4Dt}} - \sqrt{\frac{v^2 t}{4D}}\right).$$

By replacing the spatial coordinate z with the length of the tube L in the previous equation, the fraction of dose exiting the small intestine as a function of time is obtained:

$$\frac{c\,(L,t)}{c_0} = \frac{1}{2}\,\mathrm{erf}\,c\left(\frac{L-vt}{\sqrt{4Dt}}\right).$$

This equation allows one to consider the cumulative distribution of small-intestinal transit time data with respect to the fraction of dose entering the colon as a function of time. In this context, this equation characterizes well the small-intestinal transit data [177, 178], while the optimum value for the dispersion coefficient D was found to be equal to $0.78\,\mathrm{cm}^2\,\mathrm{s}^{-1}$. This value is much greater than the classical order of magnitude $10^{-5}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$ for molecular diffusion coefficients since it originates from Taylor dispersion due to the difference of the axial velocity at the center of the tube compared with the tube walls, as depicted in Figure 6.5.

For absorbable substances, a first-order absorption term can be coupled with the convection–dispersion (6.15) to model both the fluid flow and the absorption process:

$$\frac{\partial c\,(z,t)}{\partial t} = D\frac{\partial^2 c\,(z,t)}{\partial z^2} - v\frac{\partial c\,(z,t)}{\partial z} - k_a c\,(z,t)\,, \tag{6.17}$$

where k_a is the first-order absorption rate constant. Although the previous equation is solved numerically, an analytical solution can be obtained [183] for appropriate initial and boundary conditions. More specifically, with a zero initial condition $c\,(z,0) = 0$ and boundary conditions that correspond to a constant reservoir for an initial period $t_\circ$ only,

$$c\,(0,t) = \begin{cases} c_0 \text{ for } 0 < t \le t_\circ, \\ 0 \quad \text{for } t_\circ < t, \end{cases} \qquad \left.\frac{\partial c(z,t)}{\partial z}\right|_{z\to\infty,t} = 0,$$

the analytical solution of (6.17) is

$$c\,(z,t) = \begin{cases} c_0\Phi\,(z,t) & \text{for } 0 < t \le t_\circ, \\ c_0\Phi\,(z,t) - c_0\Phi\,(z,t-t_\circ) & \text{for } t_\circ < t, \end{cases}$$

where

$$\Phi\,(z,t) = \frac{1}{2}\exp\left[\frac{(v-\alpha)\,z}{2D}\right]\mathrm{erf}\,c\left(\frac{z-vt}{\beta}\right) + \frac{1}{2}\exp\left[\frac{(v+\alpha)\,z}{2D}\right]\mathrm{erf}\,c\left(\frac{z+vt}{\beta}\right)$$

and

$$\alpha = v\sqrt{1 + 4k_a Dv^{-2}}, \quad \beta = 2\sqrt{Dt}.$$

Profiles of the analytical solution of (6.17) were plotted in Figure 6.6B.

In this category of dispersion models, one can also classify a "continuous plug flow with dispersion" model for the simulation of gastrointestinal flow and drug absorption [184]. In this model, the drug is passively absorbed, while the intestinal transit is described via a Gaussian function. The drug solution moves in a concerted fashion along the intestines, but with an ever-widening distribution about the median location in contrast to the time-distribution theoretical profiles of classical dispersion–convention models shown in Figure 6.6. The model described nicely the dose-dependent absorption of chlorothiazide in rats [184], and it has been used for the development of a physiologically based model for gastrointestinal transit and absorption in humans [185].

Recently, a novel convection–dispersion model for the study of drug absorption in the gastrointestinal tract, incorporating spatial heterogeneity, was presented [186]. The intestinal lumen is modeled as a tube (Figure 6.7), where the concentration of the drug is described by a system of convection–dispersion partial differential equations. The model considers:

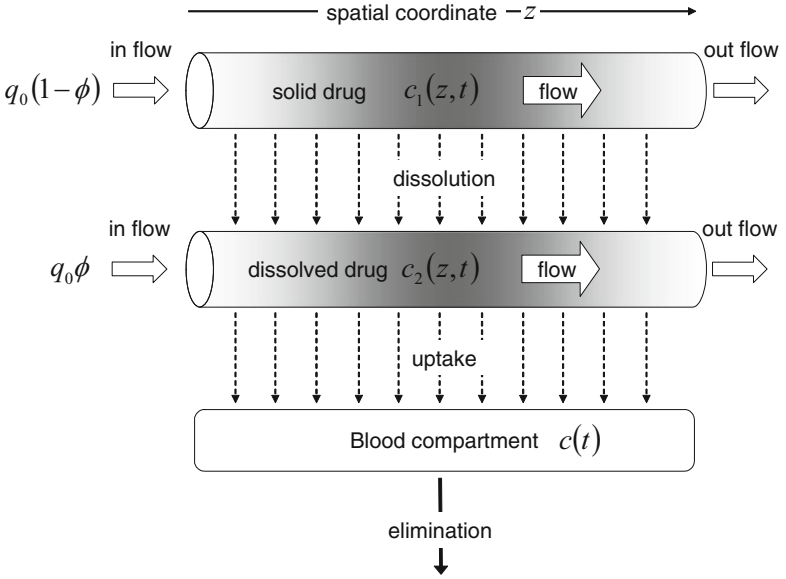


Fig. 6.7 A dispersion model that incorporates spatial heterogeneity for the gastrointestinal absorption processes. q_0 denotes the administered dose and φ is the fraction of dose dissolved in the stomach

- two drug concentrations, for the dissolved and the undissolved drug species, and
- spatial heterogeneity along the axis of the tube for the various processes included, i.e., axial heterogeneity for the velocity of the intestinal fluids, the constants related to the dissolution of the solid drug, and the uptake of the dissolved drug from the intestinal wall.

The model includes more realistic features than previously published dispersion models for the gastrointestinal tract, but the penalty for that is that it can be solved only numerically and includes a large number of parameters that are difficult to be estimated based solely on blood data.

6.4 Heterogeneous Approaches

The approaches discussed in Sections 6.1, 6.2, and 6.3 were based on the concept of homogeneity. Hence, the analysis of drug dissolution, transit, and uptake in the gastrointestinal tract was accompanied by the assumption of perfect mixing in the compartment(s) or the assumption of homogeneous flow. In the same vein, the convection–dispersion models [177, 178, 181–184, 186] consider the small intestine as a uniform tube with constant axial velocity, constant dispersion behavior, and constant-concentration profile across the tube diameter. The heterogeneous approaches attempt to incorporate the geometrically heterogeneous features of the internal structure of the intestinal tube, e.g., microvilli as well as the inhomogeneous flow of drug toward the lower end of the intestinal tube.

The assumptions of homogeneity and/or well-stirred media used in Sections 6.1 to 6.3 are not only not obvious, but they are also in fact contrary to the evidence given the anatomical and physiological complexity of the gastrointestinal tract. Both in vivo drug dissolution and uptake are heterogeneous processes since they take place at interfaces of different phases, i.e., the liquid–solid and liquid–membrane boundaries, respectively. In addition, both processes occur in heterogeneous environments, i.e., variable stirring conditions in the lumen. The mathematical analysis of all models described previously relies furthermore on the assumption that an isotropic three-dimensional space exists in order to facilitate the application of Fick's laws of diffusion. However, recent advances in physics and chemistry, as discussed in Chapter 2, have shown that the geometry of the environment in which the processes take place is of major importance for the treatment of heterogeneous processes. In media with topological constraints, well-stirred conditions cannot be postulated, while Fick's laws of diffusion are not valid in these spaces. Most of the arguments questioning the validity of the diffusion theory in a biological context seem to be equally applicable in the complex media of the gastrointestinal tract [187, 188]. However, advances in heterogeneous kinetics have led to the development of fractal-like kinetics that are suitable for processes taking place in heterogeneous media and/or involving complicated mechanisms. In the light of the above-mentioned gastrointestinal heterogeneity, the drug gastrointestinal processes are discussed below in terms of fractal concepts [189].

6.4.1 Heterogeneous Gastrointestinal Transit

Since gastrointestinal transit has a profound effect on drug absorption, numerous studies have focused on the gastric emptying and the intestinal transit of different pharmaceutical dosage forms. Gastric emptying is totally controlled by the two patterns of upper gastrointestinal motility, i.e., the interdigestive and the digestive motility pattern [190]. The interdigestive pattern dominates in the fasting state and is organized into alternating phases of activity and quiescence. Studies utilizing gamma scintigraphy have shown that gastric emptying is slower and more consistent in the presence of food [191, 192]. The transit through the small intestine, by contrast, is largely independent of the feeding conditions and physical properties of the system [191, 192], with an average transit time of $\approx$ 3 h [177]. Thus, normal transport seems to operate in the various segments of the small intestine and therefore a linear evolution in time of the mean position of the propagating packet of drug molecules or particles can be conceived.

Several studies with multiparticulate forms have indicated that the movement of pellets across the ileo–caecal junction involves an initial regrouping of pellets prior to their entrance and spreading in the colon [193–195]. According to Spiller et al. [196] the ileocolonic transit of 1 ml solution of a ^{99m}Tc-diethyl triamine-pentaacetic acid (DTPA) in humans is rapid postprandially and slow and erratic during fasting. Under fasting conditions the ileum is acting as a reservoir in several cases and the colonic filling curves of DTPA exhibit long plateaus and low slopes that are indicative of episodic colonic inflow and wide spreading of the marker in the colon [196]. Similarly, Krevsky et al. [197] have shown that an 8 ml bolus containing ^{111}In-DTPA installed into the cecum was fairly evenly distributed throughout all segments of the colon after 3 h. Finally, the colonic transit of different-sized tablets has also been shown to follow the same spreading pattern [198]. This type of marker movement is most likely due to the electrical activity of the proximal and distal parts of the colon [190]. The electrical waves in these regions are not phase locked and therefore random contractions of mixing and not propulsion of contents is observed. From a kinetic point of view, the wide spreading of the marker in the colon is reminiscent of what is known in physics as *dispersive transport* [199]. This conclusion can be derived if one compares time distribution analysis data of colonic transit (cf., for example, the data of the first 3 hours in Figure 3 of [197]) with the general pattern of dispersive transport (Figure 4 in [199]). These observations substantiate the view that dispersive transport [199] operates in the large intestine and therefore the mean position of the propagating packet of drug particles is a sublinear function of time. However, dispersive transport is a scale-invariant process with no intrinsic transport coefficients; in other words, a mean transit time does not exist since transport coefficients become subject- and time-dependent [199]. These observations provide an explanation for the extremely variable whole-bowel transit, i.e., $0.5 - 5$ d [198], since the greater part of the transit is attributable to residence time in the large intestine.

6.4.2 *Is in Vivo Drug Dissolution a Fractal Process?*

In the pharmaceutical literature there are several reports that demonstrate that flow conditions in the gastrointestinal tract do not conform to standard hydrodynamic models. Two investigations [200, 201] assessed the gastrointestinal hydrodynamic flow and the mechanical destructive forces around a dosage form by comparing the characteristics of in vitro and in vivo release of two different types of controlled-release paracetamol tablets. The results [200] indicate that the hydrodynamic flow around the dosage forms in the human gastrointestinal tract are very low, corresponding to a paddle speed of 10 rpm in the paddle method of dissolution or a velocity of about $1\,\mathrm{cm\,min^{-1}}$ ($1–2\,\mathrm{ml\,min^{-1}}$ flow rate) in the flow-through cell method. In parallel, low and high in vitro destructive forces were found to be physiologically meaningful and essential for establishing a useful in vitro dissolution testing system [200, 201].

Furthermore, data from gastrointestinal physiology have long since shown the heterogeneous picture of the gastrointestinal contents as well as the importance of mechanical factors in the gastrointestinal processes [190]. It is very well established that the gastric contents are viscous, while shearing forces in the chyme break up friable masses of food. Since chyme moves slowly down the intestine by segmentation and short, weak propulsive movements, the flow is governed by resistance as well as by pressure generated by contraction [190]. Thus, there is a progressive reduction of the transit rate from duodenum to the large intestine [202, 203].

All the above observations [190–203] substantiate the view that the flow is forced in the narrow and understirred spaces of the colloidal contents of the lower part of the gastrointestinal tract. Consequently, friction becomes progressively more important than intermolecular diffusion in controlling the flow as the drug moves down the intestine. The characteristics of this type of flow have been studied [204, 205] with Hele–Shaw channels ensuring a quasi-two-dimensional space using miscible fluids of different viscosities. These studies revealed that when a less-viscous fluid moves toward a fluid with higher viscosity (polymer solution or colloidal suspension), the interface ripples and very soon becomes extremely meandering (fractal). These viscous, fractal fingers have been observed in experiments mimicking the secretion of HCl and its transport through the mucus layer over the surface epithelium [206]. Confirmation of this type of morphology (channel geometry) in the mucus layer has been provided by an in vivo microscopic study of the acid transport at the gastric surface [207]. The results obtained with the dyes Congo red and acridine strongly suggest that secreted acid (and pepsin) moves from the gastric crypts across the surface mucus layer into the luminal bulk solution only at restricted sites [207].

In the light of these observations one can argue that the dissolution of sparingly soluble drugs should be performed in topologically constrained media since the drug particles traverse the larger part or even the entire length of the intestines and attrition is a significant factor for their dissolution. However, one can anticipate poor

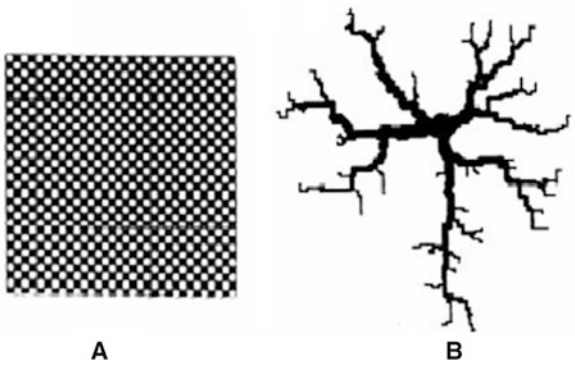


Fig. 6.8 Geometric representation of dissolution under (**A**) homogeneous and (**B**) heterogeneous conditions at a given time t. Reprinted from [189] with permission from Springer

reproducibility of dissolution results in topologically constrained media [208, 209] since the dissolution of particles will be inherently linked with the fractal fingering phenomenon, Figure 6.8:

1. The square in Figure 6.8A represents geometrically all currently used well-stirred dissolution media, which ensure at any time a homogeneous concentration of drug throughout their volume. Due to homogeneity a sample taken from a well-stirred dissolution medium can provide the amount of drug dissolved (white squares) after separation of the undissolved drug (black squares).
2. Dissolution in topologically constrained media gives rise to fractal fingering, Figure 6.8B (cf. also figures in [205, 208, 209]). The tree-like structure shown here indicates the flow of liquid where dissolution takes place. This structure is generated via the modified diffusion-limited aggregation (DLA) algorithm of [209] using the law $\rho = \alpha\,(m/N)^{\beta}$. Here, $N = 2,000$ (the number of particles of the DLA clusters), $\alpha = 10$ and $\beta = 0.5$ are constants that determine the shape of the cluster, ρ is the radius of the circle in which the cluster is embedded, $\rho_c = 0.1$ is the lower limit of ρ (always $\rho_c < \rho$), and m is the number of particles sticking to the downstream portion of the cluster. This example corresponds to a radial Hele–Shaw cell where water has been injected radially from the central hole. Due to heterogeneity a sample cannot be used to calculate the dissolved amount at any time, i.e., an average value for the percent dissolved amount at any time does not exist. This property is characteristic of fractal objects and processes.

According to van Damme [205], fractal fingering is in many respects a chaotic phenomenon because it exhibits a sensitive dependence on the initial conditions. Although this kind of performance for a dissolution system is currently unacceptable, it might mirror more realistically the erratic dissolution of drugs with very low extent of absorption.

6.4.3 Fractal-like Kinetics in Gastrointestinal Absorption

Derivation of the equations used in linear compartmental modeling relies on the hypothesis that absorption takes place from a homogeneous drug solution in the gastrointestinal fluids and proceeds uniformly throughout the gastrointestinal tract. Homogeneous gastrointestinal absorption is routinely described by the following equation [210]:

$$\dot{q_a}(t) = F_a q_0 k_a \exp(-k_a t),$$

where F_a is the fraction of dose (q_0) absorbed, and k_a is the first-order absorption rate constant. Nevertheless, the maximum initial absorption rate ($F_a q_0 k_a$) associated with the previous equation is not in accord with stochastic principles applied to the transport of drug molecules in the absorption process [210]. Theoretically, the absorption rate must be zero initially and increase to reach a maximum over a finite period of time. This type of time dependency for the input rate has been verified in deconvolution and maximum entropy studies of rapid-release dosage forms [210–212]. To overcome the discrepancies between the above equation and the actual input rates observed in deconvolution studies, investigators working in this field have utilized a cube-root-law input [213], polynomials [214], splines [212], and multiexponential [215] functions of time. In the same vein, but from a pharmacokinetic perspective, Higaki et al. [216] have considered models for time-dependent rate "constants" in oral absorption. Although these approaches [210, 212–216] are purely empirical, their capability in approximating the real input function indicates that power functions of time can be of value in describing the gastrointestinal drug absorption.

A more realistic approach to modeling drug absorption from the gastrointestinal tract should take into account the geometric constraints imposed by the heterogeneous structure and function of the medium. A diffusion process under such conditions is highly influenced, drastically changing its properties. For example, for a random walk in disordered media, the mean square displacement $\langle z^2(t) \rangle$ of the walker is given by (2.10):

$$\langle z^2(t) \rangle \propto t^{2/d_w},$$

where d_w is the random-walk dimension (cf. Section 2.2). The value of d_w is larger than 2, typically $d_w = 2.8$ (2 dimensions), and $d_w = 3.5$ (3 dimensions), so the overall exponent is smaller than 1. Furthermore, in understirred media, where reactions or processes take place in a low-dimensional space, the rate "constant" is in fact time-dependent at all times (cf. Section 2.5). Hence, the transit, dissolution, and uptake of drug under the heterogeneous gastrointestinal conditions can obey the principles of fractal kinetics [17, 217], where rate "constants" depend on time. For these heterogeneous processes, the time dependency of the rate coefficient k is expressed by

$$k = k_o t^{\lambda},$$

where $k_\circ$ is a constant, while the exponent λ is different from zero and is the outcome of two different phenomena: the heterogeneity (geometric disorder of the medium) and the imperfect mixing (diffusion-limit) condition. Therefore, k depends on time since $\lambda \neq 0$ in inhomogeneous spaces while in three-dimensional homogeneous spaces $\lambda = 0$ and therefore $k = k_\circ$, i.e., classical kinetics prevail and the rate constant does not depend on time. For "ideal" drugs having high solubility and permeability the homogeneous assumption ($\lambda = 0$, gastrointestinal absorption proceeds uniformly from a homogeneous solution) seems to be reasonable. In contrast, this assumption cannot be valid for the majority of drugs and in particular for drugs having low solubility and/or permeability. For these drugs a suitable way to model their gastrointestinal absorption kinetics under the inhomogeneous gastrointestinal conditions is to consider a time-dependent absorption rate coefficient k_a,

$$k_a = k_1 t^{\alpha},$$

and a time-dependent dissolution rate coefficient k_d,

$$k_d = k_2 t^{\beta}.$$

In reality, the exponents α and β determine how sensitive k_a and k_d are in temporal scale and the kinetic constants k_1 and k_2, determine whether the processes happen slowly or rapidly. The dimensions of k_1 and k_2 are time$^{-(1+\alpha)}$ and time$^{-(1+\beta)}$, respectively. Thus, the absorption rate $\dot{q_a}(t)$ is

$$\dot{q_a}(t) = k_a q_a(t) = k_1 t^{\alpha} q_a(t),$$

where $q_a(t)$ is the dissolved quantity of drug in the gastrointestinal tract. Since the change of $q_a(t)$ is the result of dissolution and uptake, which are both taking place under heterogeneous conditions ($\alpha \neq 0$ and/or $\beta \neq 0$), the previous equation exhibits a nonclassical time dependency for the input rate. Consequently, this equation provides a theoretical basis for the empirical power functions of time utilized in deconvolution studies [210, 212–215].

The values of the parameters α and β for drugs exhibiting heterogeneous absorption kinetics are inherently linked with the physicochemical properties of the drug, the formulation, the topology of the medium (gastrointestinal contents), and the initial distribution of drug particles in it [17]. It is worthy of mention that the initial conditions (the initial random distribution of the reactants: solid drug particles and gastrointestinal contents) are very important in fractal kinetics [17]. For all these reasons, population parameters for drugs having $\alpha \neq 0$ and/or $\beta \neq 0$ are unlikely since the topology of the medium and the initial conditions are by no means consistent or controlled, being dependent on subject and time of day. For the sake of completion, one should add that under homogeneous conditions ($\alpha = \beta = 0$) both k_a and k_d are independent of time and therefore classical kinetics can be applied.

6.4.4 The Fractal Nature of Absorption Processes

Relying on the above considerations one can argue that drugs can be classified with respect to their gastrointestinal absorption characteristics into two broad categories, i.e., homogeneous and heterogeneous. Homogeneous drugs have satisfactory solubility and permeability, and are dissolved and absorbed mostly prior to their arrival to the large intestine. It seems likely that the gastrointestinal absorption characteristics of the homogeneous group of drugs are adequately described or modeled with the homogeneous approach, i.e., well-stirred in vitro dissolution systems and classical absorption kinetics. In contrast, drugs with low solubility and permeability can be termed heterogeneous, since they traverse the entire gastrointestinal tract, and are most likely to exhibit heterogeneous transit, dissolution, and uptake and therefore heterogeneous absorption kinetics. In this context, the following remarks can be made for the heterogeneous drugs that exhibit limited bioavailability and high variability, and most of them can be classified in categories II and IV of the BCS [157] (cf. also Section 6.6.1):

- Mean or median values should not be given for the whole bowel transit since most of the dissolved and/or undissolved drug traverses the entire gastrointestinal tract. The complex nature of transit involving normal and dispersive transport [199] as well as periods of stasis would be better expressed by reporting the range of the experimental values.
- Dissolution testing with the officially used in vitro systems ensuring homogeneous stirring conditions, should be solely viewed as a quality control procedure and not as a surrogate for bioequivalence testing. According to the current view [157], limited or no in vitro–in vivo correlations are expected using conventional dissolution tests for the category IV drugs and the drugs of category II used in high doses. Since this unpredictability is routinely linked with our inability to adequately mimic the in vivo conditions, one should also consider whether the chaotic character of in vivo dissolution is a valid hypothesis for the failure of the in vitro tests. It is advisable, therefore, to perform physiologically designed dissolution experiments in topologically constrained media [205, 208, 209] for drugs of categories II and IV [157] in order to determine potential cutoffs for dose and solubility values as well as flow characteristics for drug classifications (homogeneous and heterogeneous drugs). Further, these cutoffs could be used for setting standards for in vitro drug dissolution methodologies of drugs classified as heterogeneous.
- A notion that routinely accompanies oral absorption studies is that the mathematical properties of the underlying processes have a Gaussian distribution where the moments, such as the mean and variance, have well-defined values. Relying on this notion, drugs and/or formulations are categorized as low or highly variable. Thus, any drug that generates an intraindividual coefficient of variation greater than 30% as measured by the residual coefficient of variation (from analysis of variance) is arbitrarily characterized as highly variable. The use of a statistical measure of dispersion for drug classification is based on the law of large numbers,

which dictates that the sample means for peak blood concentration, c_{max}, and the area under the blood time–concentration curve, AUC, converge to fixed values while the variances decrease to nonzero finite values as the number used in averaging is increased. The conventional assessment of bioequivalence relies on the analysis of variance to get an estimate for the intraindividual variability prior to the construction of the 90% confidence interval between 80 and 125% for AUC and c_{max}. The basic premise of this approach is that errors are normally distributed around the estimated mean values and two one-sided t-tests can be performed. Although the validity of this assumption seems to be reasonable for drugs following classical kinetics, concern is arising for the parameters c_{max} and t_{max} (time corresponding to c_{max}) when fractal-like kinetics govern absorption since for many fractal time-dependent processes [4, 199] the mean and the variance may not exist. Under heterogeneous conditions, both c_{max} and t_{max} will depend on α and β, and therefore mean values for these parameters cannot be justified when fractal kinetics are operating. Apparently, a significant portion of variability with the heterogeneous drugs can be mistaken as randomness and can be caused by the time dependency of the rate coefficients of the in vivo drug processes. These observations provide a plausible explanation for the high variability in c_{max} values and the erroneous results obtained in bioequivalence studies [218]. From the above it appears that is inappropriate to apply rigorous statistical tests in bioequivalence studies for heterogeneous drugs using parameter estimates for c_{max} and t_{max} that do not actually represent sample means. The suggested [219] comparison of the time–concentration curve profiles of test and reference products in bioequivalence studies seems to be in accord with the reservations pointed out regarding use of specific parameters for the assessment of the absorption rate.

6.4.5 Modeling Drug Transit in the Intestines

The small-intestinal transit flow is a fundamental process for all gastrointestinal absorption phenomena. However, the structure of the gastrointestinal tract is highly complex and it is practically impossible to explicitly write and solve the equations of motion for the drug flow. Instead, numerical computer-simulation techniques that incorporate the heterogeneous features of the gastrointestinal wall structure and of the drug flow are used in this section to characterize the intestinal transit process in humans.

An algorithm is built from first principles, where the system structure is recreated and subsequently the drug flow is simulated via Monte Carlo techniques [220]. This technique, based on principles of statistical physics, generates a microscopic picture of the intestinal tube. The desired features of the complexity are built in, in a random fashion. During the calculation all such features are kept frozen in the computer memory (in the form of arrays), and are utilized accordingly. The principal characteristic of the method is that if a very large number of such units is built, then the average behavior of all these will approach the true system behavior.

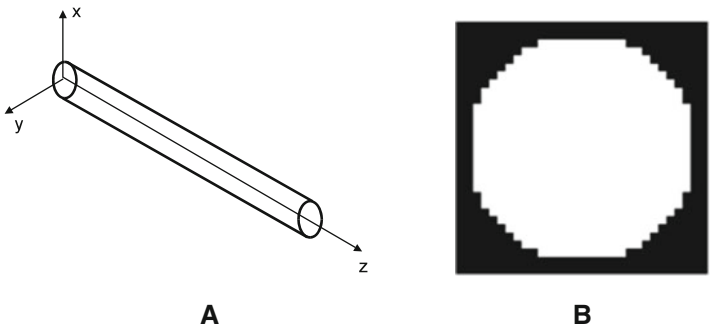


Fig. 6.9 (**A**) The cylinder used for the tube construction. (**B**) Cross section of the tube. Reprinted from [220] with permission from Springer

6.4.5.1 Construction of the Heterogeneous Tube

The model is based on a cylinder whose length is several orders of magnitude larger than its radius. Thus, any entanglements that are present are ignored, since they do not influence the dynamics of the phenomena. Initially, a three-dimensional parallelepiped with a square cross section, of size $x : y : z$ equal to $31 : 31 : 3000$ is constructed, Figure 6.9A. Inside it a cylinder with a radius of 14 units is built, a cross section of which appears in Figure 6.9B. Hence, the quotient of [radius/length]$=$ $R/L = 14/3000$ in the tube model is quite similar to the ratio of physiological data 1.3 cm/3 m for the human small intestine.

For convenience in the calculations, an underlying lattice of discrete spacing forming in effect a three-dimensional grid is used. This grid covers the entire cylinder, while for all spatial considerations the grid sites are utilized. The interior of the cylinder has a finite concentration of villi attached to the cylinder wall, which have the property that they may absorb the dissolved drug particles flowing through the cylinder. The villi have the usual random dendritic structure, and they are formed by the DLA method [209]. The absorption of the drug particles in the model takes place when a flowing particle happens to have a position right next to the villi coordinates, implying that when a particle comes in contact with a villi structure it can be absorbed. The probability for absorption by the villi or walls is p_a. Since the present model focuses on the tube structure and the characteristics of flow, $p_a = 0$, while the case of $p_a \neq 0$ is treated in the following section.

The villi have a random dendritic-type structure, and they are formed initially by use of an algorithm based on the well-known DLA [209] model from solid-state physics. At random positions, $2z$ seed particles (z the cylinder length, Figure 6.9A) are placed on the cylinder surface by positioning 2 particles on each z value. Following the DLA model, another particle, starting at a random point of each cross section, makes a three-dimensional random walk (diffusion) inside the cylinder.

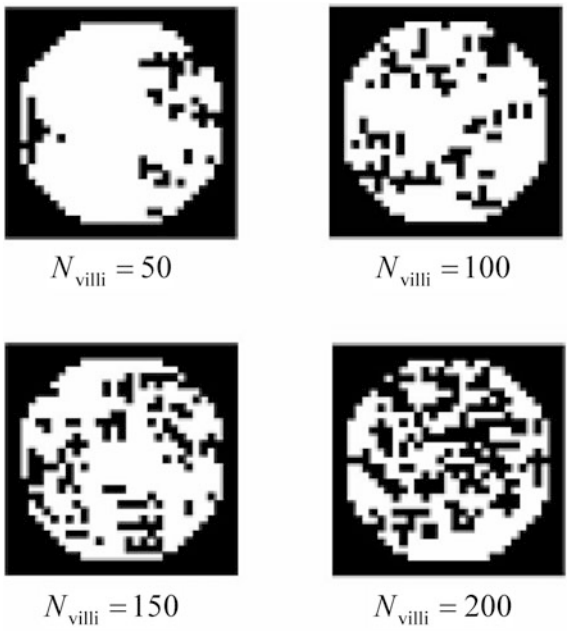


Fig. 6.10 Cross sections of the tube at random positions for various concentrations of villi, $N_{villi} = 50, 100, 150, 200$. Reprinted from [220] with permission from Springer

The walk stops when the moving particle visits any of the neighbor sites of the original seed particles. At this point it stops and becomes attached to the neighboring seed particle. The particle is constrained to move inside the cylinder. Then a second particle starts a random walk, until it meets either one of the seeds or the already *frozen* particle. The process continues and the internal structure of the tube, which can be of varying complexity, is built using a total of N_{villi} particles per unit length. The size of each villi cluster is limited to the value $1.5N_{villi}$. This is done in order to achieve a uniform distribution of villi cluster sizes. The higher the N_{villi} value, the more ramified is the ensuing structure. Some examples for various values of N_{villi} are shown in Figure 6.10. This figure shows typical two-dimensional cross sections of the cylinder, for four different N_{villi} values, $N_{villi} = 50, 100, 150$, and 200, at random places. It is clearly seen how the villi complexity is built up with increasing N_{villi}. Some squares appear not to be connected to any others in these pictures. In fact, these are indeed connected to adjacent (first neighbor) squares in the next or previous cross section of the tube (i.e., with $z' = z + 1$ or $z' = z - 1$), which are not shown in Figure 6.10.

6.4.5.2 Dynamics

The dynamics of the system are also followed utilizing the Monte Carlo technique. This includes motion of the particles through the tube, dissolution in the solvent flow, and absorption by the villi or the tube walls. Time is incremented by arbitrary time units, the MCS, which is the time it takes for a particle to move to one of its

neighbor positions. A "tablet" can be inserted in one end of the tube (input end) at predefined time increments expressed in MCS. The "tablet" is modeled as an aggregate of drug particles of mass $q_0 = 100$. This means that one "tablet" can later be broken down successively into 100 units, which represent the solid drug particles. These can be further dissolved in the encompassing solution. But as long as the "tablet" has a mass larger than one it cannot be dissolved in the solution. All diffusing species (dissolved and undissolved) flow through the cylinder from the input end toward the direction of the other end (output end). This is accomplished by using a diffusion model of a biased random walk that simulates the fluid flow.

A simple random walk is the prototype model of the regular Brownian motion. Such a model is modified here, by including a bias factor, which makes the motion ballistic rather than simply stochastic. This bias factor, ε, increases the probability for motion in the z-direction, i.e., toward the output end, as compared to the probabilities in all other directions. This makes the flow of the particles and the dissolved drug molecules possible. If $\varepsilon = 0$, there is a motion but it is rather stationary and in all possible directions. If $\varepsilon > 0$, this makes the flow possible. The rate of the flow is also directly affected by the numerical value of ε, with increasing ε values resulting in increasing flow rates. With this statistical model the diffusing species can momentarily go against the flow, or sideways. This is a realistic feature, but it occurs with reduced probability.

Two different models of the biased random walk were envisaged. In model I the three directions of space, x, y, and z, are all equally probable, but in the z direction, the probability toward the output end (z_+) is now $(1/z) + \varepsilon$, while the corresponding probability toward the input end (z_-) is $(1/z) - \varepsilon$ (where z is the coordination number of the underlying space, e.g., $z = 6$ in a three-dimensional space). This model has the characteristic that diffusion is equally probable in all possible directions, the species spending equal times in all of them, but due to the ε factor, when the z direction is chosen a positive flow drives the solution to the output end.

In a second model II, more emphasis is given to the motion toward the output and less to the other directions. The probabilities for motion in the different directions are now defined differently. While in the simple random walk the probability for motion in a specific direction is $1/z$, here the probability for motion in the output direction is $(1/z) + \varepsilon$, while the probability in any of the other five directions is

$$\frac{1 - \left(\frac{1}{z} + \varepsilon\right)}{z - 1}.$$

Thus, the values that ε can take are in the range

$$0 < \varepsilon < 1 - \frac{1}{z},$$

while the overall forward probability p_f, i.e., the probability toward the output end, is in the range

$$\frac{1}{z} < p_f < 1.$$

At each time step there is a probability p_d for the "tablet" to dissolve, i.e., $0 < p_d < 1$. In the Monte Carlo method the "tablet" is tested at every step to determine whether a fragment (one new particle) is to be released. When this happens a fragment of the "tablet" with mass $\psi = 1$ breaks off, and gets separated from the larger mass. It is understood that this $\psi = 1$ particle is immediately dissolved, and it is never reattached to the original mass. This dissolved particle now performs a random walk of its own, with the same characteristics (bias) as the main "tablet." The mass q_0 of the "tablet" is then reduced by ψ. The virtual experiment of the flow starts when a large number of drug particles (e.g., 10, 000) with mass $\psi = 1$ are inserted simultaneously at time $t = 0$ in the tube and are allowed to diffuse. To concentrate on the transit process exclusively, dissolution is considered instantaneous and p_d is set equal to 1, while absorption is not allowed by setting $p_a = 0$. When the fragments of the "tablet" reach the end of the tube, they are discarded. At the end of the simulation time the mass that has exited from the end of the tube is computed. The mean transit time is also computed by keeping track of the time it took for the particles to reach the end of the tube.

When the diffusing species come in contact with a closed site (such as the villi sites of the model) they have two options. In the first option, the particle does not "feel" the presence of the closed site, and it may attempt, unsuccessfully, to go to it. This model is called the *blind ant model*. In the second model, the particle feels the presence of the closed site, and thus it never attempts to land on it. This is called the *myopic ant model*. The difference between these two models is that the blind ant consumes long times in unsuccessful attempts, and thus its motion is slower than the myopic ant case.

6.4.5.3 Simulated vs. Experimental Data

The details of the flow of particles in the heterogeneous tube were studied using a model II biased random walk. In Figure 6.11, the mean transit time of the drug particles vs. the forward probability p_f (i.e., the probability toward the output along the z-axis) is plotted for various villi concentrations, for the two cases of the blind ant (part A), and the myopic ant (part B). For no villi structures, $N_{\text{villi}} = 0$, and for $N_{\text{villi}} = 50$ we observe that for larger p_f values the transit times of the particles were shorter, as one would expect. For larger villi concentrations the transit time became longer as p_f was increased. This behavior may seem inconsistent, but can easily be explained if we consider that when a drug fragment meets an obstacle (villi) then its forward motion is hampered, and it must move in the x or y direction (sideways) in order to circumvent it and continue moving toward the end of the tube. What happens is that when p_f values are large, then the probability for movement along the x- or y-axis is reduced. This does not give the particle the freedom to easily pass the obstacle, so it wastes time trying to move in the z direction. This explains the rise in the transit times, which is larger for larger villi concentrations. This qualitative picture is valid for both models in parts (A) and (B) of Figure 6.11. Plausibly, in comparing the two figures, the transit times are always longer in the blind ant case,

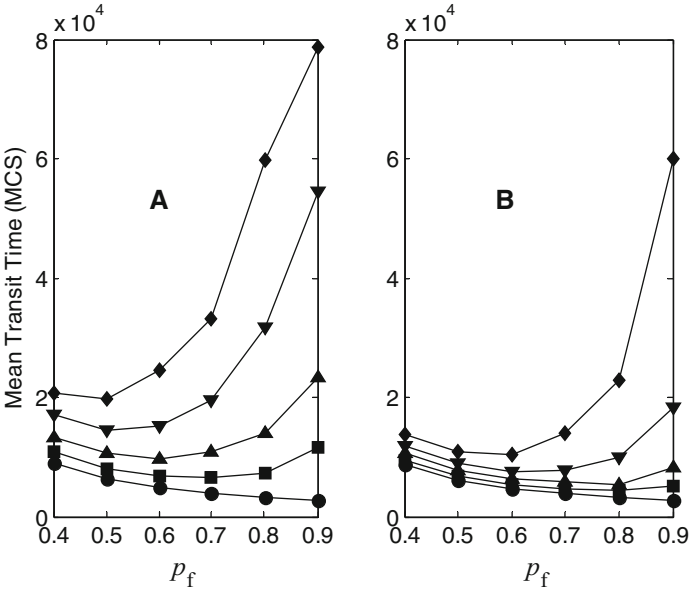


Fig. 6.11 Mean transit times vs. the forward probability for various concentrations of villi, (**A**) blind ant model; (**B**) myopic ant model. Key (N_{villi}): ● 0; ■ 50; ▲ 100; ▼ 150; ◆ 200

for any villi concentration. The system behavior as shown in Figure 6.11 implies that the interplay of these two factors, namely the villi structure and the bias probability (flow rate), is important in determining the dynamics of the flow.

The frequency of transit times that result from the simulations for various values of villi and forward probability p_f are also compared to experimental data [177]. Model I consistently produces narrower frequencies than do model II and the experiments. This is because in model I, motion in the preferred z direction occurs with the same frequency as motion in the other directions. The effect of the flow along the tube length is downplayed, as opposed to the other model (II), in which it is emphasized. In Figure 6.12 the results for model I of the biased diffusion, together with the experimental data are presented. A wide range of variation for the two parameters, i.e., the bias factor ε and the villi concentration N_{villi}, was used, and the best resemblance between simulation and experimental data was achieved for the values of $N_{\text{villi}} = 190$ and forward probability $p_f = 0.65$, Figure 6.12. The x-axis here is in units of minutes. This is done by establishing a correspondence of 1 s = 1.5 MCS, since this is the value that produces the best possible fit.

Overall, the biased random walk, which places more emphasis on the motion toward the output end and less on the other directions, mimics more closely the transit profile of the experimental data. Both diffusion models, i.e., the blind and the myopic ant models, can reproduce the basic features of the real small-intestinal transit profile.

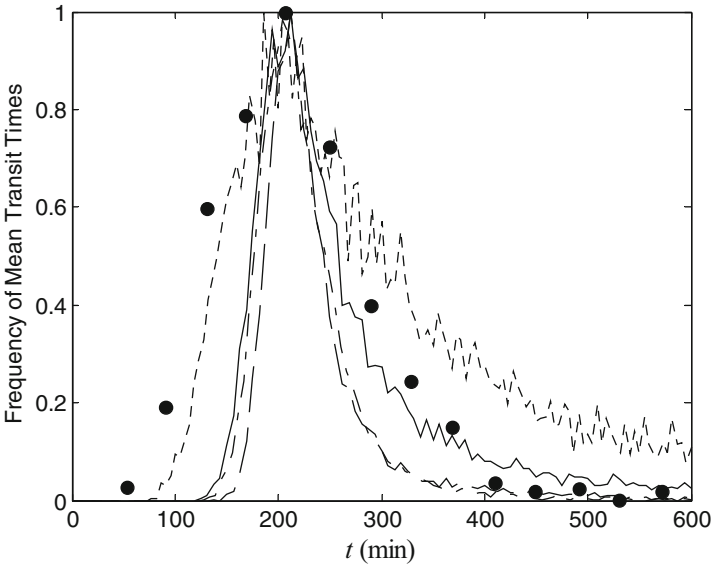


Fig. 6.12 Frequency of mean transit times vs. time (min) using the diffusion model II for the blind ant model positions for various concentrations of villi and forward probabilities p_f values. Key: • experimental data; *solid line*, $N_{\text{villi}} = 200$ and $p_f = 0.6$; *dashed line*, $N_{\text{villi}} = 200$ and $p_f = 0.5$; *dotted line*, $N_{\text{villi}} = 180$ and $p_f = 0.7$; *dashed-dotted line*, $N_{\text{villi}} = 180$ and $p_f = 0.5$

6.4.6 Probabilistic Model for Drug Absorption

The probabilistic absorption model described herein [221] was based on the cylinder built in [220] that incorporates all the random heterogeneities that make up the gastrointestinal tube. The optimal heterogeneous characteristics found in [220] were assigned to the number of villi and the type of the biased random walk. Thus, the parameter number of villi N_{villi} was set equal to 190, while the blind ant model for the biased random walk with forward probability $p_f = 0.65$ was used to simulate the motion of the dissolved and undissolved drug species. The dissolved species are tagged and continue the random walk and can be absorbed by the cylinder wall structure, or exit the tube if they reach its end. The quantities input and exiting through the tube, their transit time, and the fraction of the species absorbed and dissolved during the flow are monitored.

6.4.6.1 Simulation of Dissolution and Uptake Processes

A "tablet," which is modeled as an aggregate of drug particles of mass q_0, is inserted in one end of the tube (input end). At each time step a portion of the mass of the "tablet" can be dissolved. The rate of dissolution is considered to be dependent on three factors, which are all expressed in probability values.

1. The first factor, k_d, mimics the conventional dissolution rate constant; it is inherent for every drug and takes values in the range $0 < k_d < 1$. A value close to unity denotes a drug with rapid dissolution characteristics. Thus, a specific k_d value is conceived for a given drug under certain experimental conditions. As a probability value, k_d corresponds to p_d and it expresses the number of events occurring in a time unit. Consequently, k_d has dimension of time^{-1}.
2. The second factor, k_c, is related to the first-order concentration dependence of the dissolution rate. As dissolution proceeds the amount of drug in solution increases exponentially and therefore the value of k_c is reduced exponentially. This reduction is controlled by the relative amount dissolved, $q(t)/q_s$, as defined in Section 5.1.4, at each time point:

$$k_c \rightarrow k_c(t) = \exp\left[-\ln(10)\, q(t)/q_s\right],$$

   where $q(t)$ is the mass of the dissolved drug at any moment during the simulation and q_s is the dissolved mass at saturation. q_s is computed by multiplying the minimum physiologic solubility $c_{s,\min}$ of the drug by the luminal volume, which is assumed to be 250 ml. The $\ln(10)$ factor was chosen so that the magnitude of k_c, when the dissolved mass was equal to the dissolved mass at saturation, should arbitrarily be one-tenth of the value of k_c when the dissolved mass is equal to zero. Thus, k_c is reduced exponentially as dissolution proceeds. Of course, at saturation ($q(t) = q_s$) no more material is allowed to dissolve.
3. The third factor, k_s, depends on the surface area of the drug particles. It is known that the reduction of the surface area is related nonlinearly to the reduction of mass as dissolution proceeds. Since the nonlinear relationship between the undissolved mass, $q_0 - q(t)$, and surface area is dependent on the geometric characteristics of the drug particles, the value of k_s is considered to decrease proportionally to $\exp\{[q_0 - q(t)]/q_0\} = \exp[1 - \varphi(t)]$ in order to avoid any shape assumptions. Therefore, k_s is not computed directly in the simulation, but is calculated from the undissolved drug mass at any moment during the simulation. The exact equation that gives k_s is

$$k_s \rightarrow k_s(t) = 0.01 \exp\{4.5\,[1 - \varphi(t)]\}.$$

   The constants in the last equation are chosen so that k_s arbitrarily equals 0.9 when $q(t)$ is close to zero and $k_s = 0.01$ when $q(t)$ equals q_0. In essence, the probability factor k_s is related to the diminution of the surface area of drug particles during the dissolution process.

The quantities k_c and k_s in the last two equations result from a calculation of an exponential, and thus have no physical dimensions. The effective dissolution probability rate "constant" $k_{d,eff}$ is calculated by multiplying the above three factors, so that $k_{d,eff} = k_d k_c k_s$. Thus, $k_{d,eff}$ has dimension of time^{-1} and denotes the fraction of the total number of drug particles that can be dissolved per MCS. The mass of the "tablet" that will break off at any moment is given by multiplying the value

of $k_{d,eff}$ by the undissolved mass of the tablet. If $q_d(t)$ is this mass, then $q_d(t) = [q_0 - q(t)]\, k_{d,eff}$ and $q_d(t)/\psi$ particles of the "tablet" with mass ψ will break off, and will get separated from the larger mass. The dissolved particles now flow on their own, with the same characteristics (forward probability) as the undissolved particles. The mass $q_0 - q(t)$ of the undissolved drug is then reduced by $q_d(t)$.

Dissolved particles are tagged in the calculation at all times, so their location relative to all other particles and the tube walls is known. When one of the dissolved particles comes "in contact" (when it is in a lattice site adjacent to villi or tube wall) with the tube walls or the villi there is a probability k_a' that it will be absorbed. It is obvious that the higher the value of k_a', the higher the probability of a dissolved particle of being absorbed. This proportionality implies that only passive mechanisms are considered. If a dissolved particle is absorbed it is immediately removed from the system. If it is not absorbed, it remains on its site and continues the flow. When a dissolved or undissolved particle reaches the end of the tube, it is discarded.

At the end of the simulation time, the mass that was absorbed and the mass that has exited from the end of the tube can be computed. Further, the dimensionless absorption number A_n can be computed [157] from

$$A_n = \frac{1}{2} \langle T_{si} \rangle\, k_a$$

using (6.6) and (6.14). In this relation $\langle T_{si} \rangle$ is equal to 24, 500 MCS, i.e., the mean intestinal transit time found in [220]. It must be noted that k_a as it appears above is not identical to the one used as a parameter in the simulation. While they both describe probabilities, k_a is a first-order macroscopic rate constant expressed in dimension of time^{-1}, while the k_a' in the simulations describes the microscopic probabilistic events of the simulation model.

6.4.6.2 Absorption of Freely Soluble Drugs

The absorption of freely soluble drugs having various values of k_a' was studied. Initially, the relationship between the simulated k_a' values and the corresponding conventional k_a values, which are computed from the simulation assuming first-order absorption, was explored. An amount of instantly dissolved mass of $q_0 = 20,000$ was inserted in the input end of the tube and both profiles of the fraction of the mass that was absorbed and exited the tube were recorded. To find out the relationship between k_a' and k_a, the following exponential equation was used to fit the simulated data of the fraction of dose absorbed F_a vs. time:

$$F_a = 1 - \exp(-k_a t),$$

where the fitting parameter is k_a in MCS^{-1} units, and time t is also expressed in MCS. Focusing on k_a' values, which ensure that most of the drug is absorbed and does not exit the tube, the following relation between k_a' and k_a was found:

$$k_a = 0.885 k_a'.$$

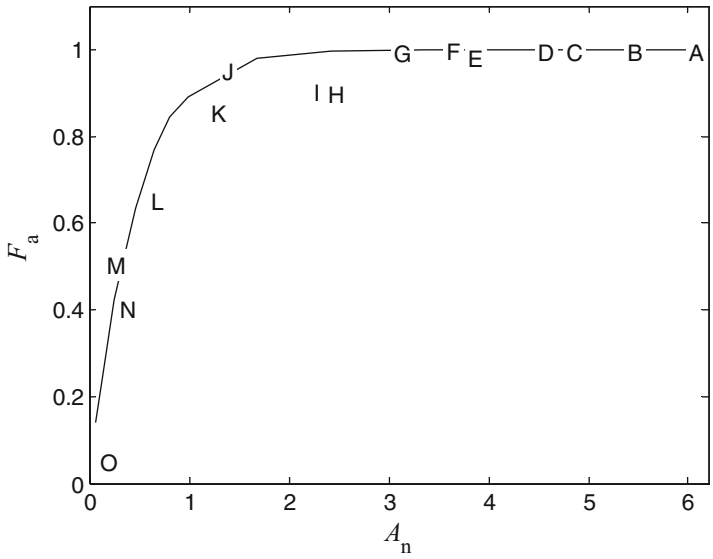


Fig. 6.13 Fraction of dose absorbed vs. A_n. The *solid line* represents results for 24, 500 MCS and the points the experimental data. Key: **A**, D-glucose; **B**, ketoprofen; **C**, naproxen; **D**, antipyrine; **E**, piroxicam; **F**, L-leucine; **G**, phenylalanine; **H**, beserazide; **I**, L-dopa; **J**, propranolol; **K**, metoprolol; **L**, terbutaline; **M**, furosemide; **N**, atenolol; **O**, enalaprilate

This relationship shows the proportionality between the first-order macroscopic rate constant k_a and the k'_a that describes the microscopic probabilistic events (the "successful" visits of the dissolved species to the villi). Similar simulations for instantly dissolved 20, 000 drug particles were carried out using various values of k'_a, and the fraction of the drug dose absorbed, F_a, at 24, 500 MCS was calculated. The k'_a values were then translated to MCS^{-1} values using the last equation, and the absorption number A_n was computed as delineated above. The fraction of the dose that was absorbed vs. the absorption number A_n is shown in Figure 6.13. The symbols represent the experimental data of various drugs [56], while the line gives the simulation results obtained from the model by adjusting the intestinal transit time to 24, 500 MCS. From the different intestinal transit times evaluated it was found that 24, 500 MCS gave the best description of the experimental data. Using the correspondence between MCS and real time units [220], the 24, 500 MCS are 16, 333 s or 4.5 h. The duration of 4.5 h is physiologically sound as an effective intestinal transit time to study gastrointestinal drug absorption in the model.

6.4.6.3 Absorption of Sparingly Soluble Drugs

The model was also applied to the study of low-solubility drugs. Numerical results of the system of differential equations reported in [56] were compared to the simulations based on the heterogeneous tube. In the simulations the z^* variable is

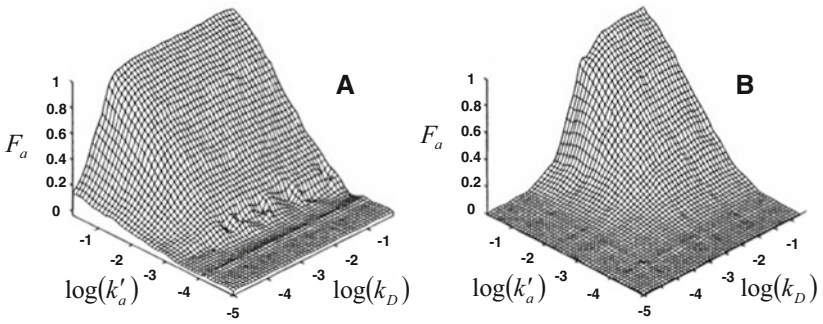


Fig. 6.14 Three-dimensional graph of fraction dose absorbed vs. k'_a and k_d. Dose and $c_{s,\min}$ values [157] correspond to those of digoxin (**A**) and griseofulvin (**B**)

computed using the mean transit time of the particles, $\langle T_{si} \rangle = 24,500$ MCS, and $z^* = t/\langle T_{si} \rangle$, expressing both t and $\langle T_{si} \rangle$ in MCS. The "tablet" was inserted in the tube entrance as a bolus of a given weight q_0 (e.g., 200 or 500 mg) and it was arbitrarily set that the bolus may break up eventually into a large number of particles, each weighing 0.01 mg. Thus, each "tablet" of mass q_0 can be finally broken down to $q_0/0.01$ particles. The values of k_c and k_s were continuously computed during the simulation–fitting procedure. Various values of the parameter k_d were used to get a good matching of the simulation and the theoretical curves obtained from the solution of equations [56] for the normalized concentration profile in the tube.

Finally, a three-dimensional plot of the fraction of dose absorbed F_a at $24,500$ MCS for various values of the parameters k'_a and k_d is shown in Figure 6.14 using values for dose and $c_{s,\min}$ corresponding to those of digoxin and griseofulvin. The plots of Figure 6.14 are indicative of the effect of dose on the fraction of dose absorbed for sparingly soluble drugs. For example, for a highly permeable drug ($k'_a \approx 0.5$) given in a large dose (500 mg) and having the dissolution characteristics of griseofulvin, $\approx$ 25% of the administered dose will be absorbed according to Figure 6.14B. In contrast, a drug like digoxin, which exhibits the same permeability and dissolution characteristics as griseofulvin, given at a low dose (0.5 mg) will be almost completely absorbed, Figure 6.14A.

6.5 Absorption Models Based on Structure

The ability to predict the fraction of dose absorbed F_a and/or bioavailability is a primary goal in the design, optimization, and selection of potential candidates in the development of oral drugs. Although new and effective experimental techniques have resulted in a vast increase in the number of pharmacologically interesting compounds, the number of new drugs undergoing clinical trial has not increased at the same pace. This has been attributed in part to the poor absorption of the

compounds. Thus, computer-based models based on calculated molecular descriptors have been developed to predict the extent of absorption from chemical structure in order to facilitate the lead optimization in the drug discovery process. Basically, the physicochemical descriptors of drug molecules can be useful for predicting absorption for passively absorbed drugs. Since dissolution is the rate-limiting step for sparingly soluble drugs, while permeability becomes rate-controlling if the drug is polar, computer-based models are based on molecular descriptors related to the important drug properties solubility and permeability across the intestinal epithelium.

A rapid popular screen for compounds likely to be poorly absorbed is Lipinski's [222] "rule of 5," which states that poor absorption of a compound is more likely when its structure is characterized by:

- molecular weight > 500,
- $\log P > 5$,
- more than 5 H-bond donors expressed as the sum of NHs and OHs, and
- more than 10 H-bond acceptors expressed as the sum of Ns and Os.

However, compounds that are substrates for biological transporters are exceptions to the rule. Based on the analysis of 2, 200 compounds in the World Drug Index that survived Phase I testing and were scheduled for Phase II evaluation, Lipinski's "rule of 5" revealed that less than 10% of the compounds showed a combination of any two of the four parameters outside the desirable range. Accordingly, the "rule of 5" is currently implemented in the form "if two parameters are out of range, a poor absorption is possible." However, compounds that pass this test do not necessarily show acceptable absorption.

Although various computational approaches for the prediction of intestinal drug permeability and solubility have been reported [223], recent computer-based absorption models utilize a large number of topological, electronic, and geometric descriptors in an effort to take both aqueous drug solubility and permeability into account. Thus, descriptors of "partitioned total surface areas" [172], Abraham molecular descriptors [224, 225], and a variety of structural descriptors in combination with neural networks [226] have been shown to be determinants of oral drug absorption.

Overall, the development of a robust predictor of the extent of absorption requires a careful screening of a large number of drugs that undergo passive transport to construct well-populated training and external validation test sets. The involvement in the data sets of compounds with paracellular, active transport, carrier-mediated transport mechanisms, or removal via efflux transporters can complicate the problem of in silico prediction of the extent of absorption. Another problem arises from the fact that published drug data for F_a or bioavailability are skewed toward high values (≈ 1), while the compounds in the training and external validation data sets should evenly distributed across the complete range of oral absorption.

6.6 Regulatory Aspects

Over the past fifteen years the advances described in the previous sections of this chapter have enhanced our understanding of the role of:

- the physicochemical drug properties,
- the physiological variables, and
- the formulation factors in oral drug absorption.

As a result, the way in which regulatory agencies are viewing bioavailability and bioequivalence issues has undergone change. In this section, we discuss the scientific basis of the regulatory aspects of oral drug absorption.

6.6.1 Biopharmaceutics Classification of Drugs

As mentioned in Section 6.1.2, the first attempts to quantitatively correlate the physicochemical properties of drugs with the fraction of dose absorbed were based on the absorption potential concept in the late 1980s [160, 162]. The elegant analysis of drug absorption by Amidon's group in 1993 based on a microscopic model [56] using mass balance approaches enabled Amidon and his colleagues [157] to introduce the Biopharmaceutics Classification System (BCS) in 1995. According to BCS a substance is classified on the basis of its aqueous solubility and intestinal permeability, and four drug classes were defined as shown in Figure 6.15. The properties of drug substance were also combined with the dissolution characteristics of the drug product, and predictions with regard to the in vitro–in vivo correlations for each of the drug classes were pointed out.

This important achievement affected many industrial, regulatory, and scientific aspects of drug development and research. In this context, the FDA guidance [227] on BCS issued in 2000 provides regulatory benefit for highly permeable drugs that are formulated in rapidly dissolving solid immediate release formulations. The guidance [227] defines a substance to be highly permeable when the extent of absorption in humans is 90% or more based on determination of the mass balance or in comparison to an intravenous reference dose. In parallel, the guidance [227] classifies a substance to be highly soluble when the highest dose strength is soluble in 250 ml or less of aqueous media over the pH range $1-7.5$, while a drug product is defined as rapidly dissolving when no less than 85% of the dose dissolves in 30 min using USP Apparatus 1 at 100 rpm in a volume of 900 ml in 0.1N HCl, as well as in pH 4.5 and pH 6.8 buffers.

It has been argued [228] that the use of a single solubility value in the original BCS article [157], Figure 6.15, for solubility classification is inadequate since drugs are administered in various doses. Moreover, solubility is a static equilibrium parameter and cannot describe the dynamic character of the dissolution process. Both aspects are treated in the guidance on biowaivers [227]; solubility is related to

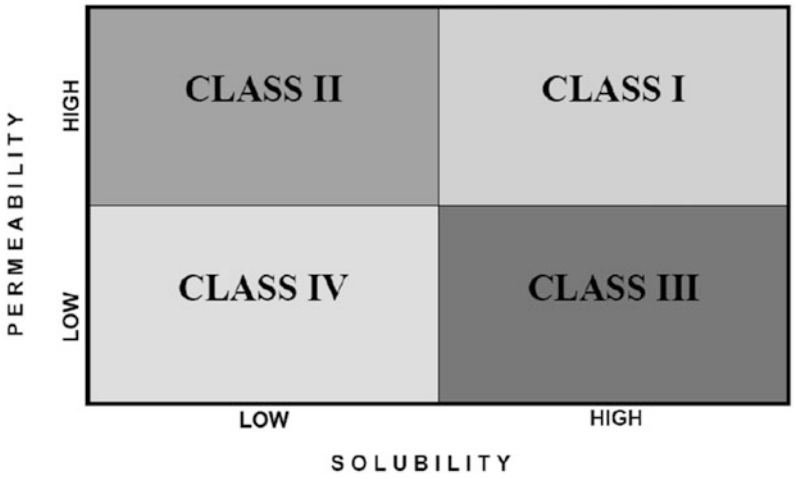


Fig. 6.15 The Biopharmaceutics Classification System (BCS). The reader should note that the original presentation of BCS [157] has been modified here and follows a kind of a "Cartesian plot," i.e., Class I (high solubility, high permeability) lies in the upper right corner while Class IV (low solubility, low permeability) lies at the lower left corner

dose, while dissolution criteria are specified. However, the reference of the FDA guidance exclusively to "the highest dose strength" for the definition of highly soluble drugs implies that a drug is always classified in only one class regardless of possible variance in performance with respect to solubility of smaller doses used in actual practice. This is not in accord with the dose dependency (non-Michaelian type) of oral drug absorption, which consistently has been demonstrated in early [160, 162] and recent studies [164, 165] related to the absorption potential concept and its variants as well as in the dynamic absorption models [56, 184, 185]. Moreover, the dissolution criteria of the FDA guidance [227], which unavoidably refer to a percentage of dose dissolved within a specific time interval:

- are not used as primary determinants of drug classification,
- have been characterized as conservative [229],
- have had pointed out suggestions for broadening them [230], and
- suffer from a lack of any scientific rationale.

In parallel, the current dissolution specifications [227] are not correlated with the drug's dimensionless solubility–dose ratio $1/\theta$, which has been shown [92] to control both the extent of dissolution and the mean dissolution time, *MDT*, which is a global kinetic parameter of drug dissolution.

The latter finding prompted the development of the Quantitative Biopharmaceutics Classification System (QBCS) [228] in which specific cutoff points are used for drug classification in the solubility–dose ratio permeability plane, Figure 6.16. Unity was chosen as the critical parameter for the dimensionless solubility–dose ratio axis because of the clear distinction between the two cases of complete dissolution (when

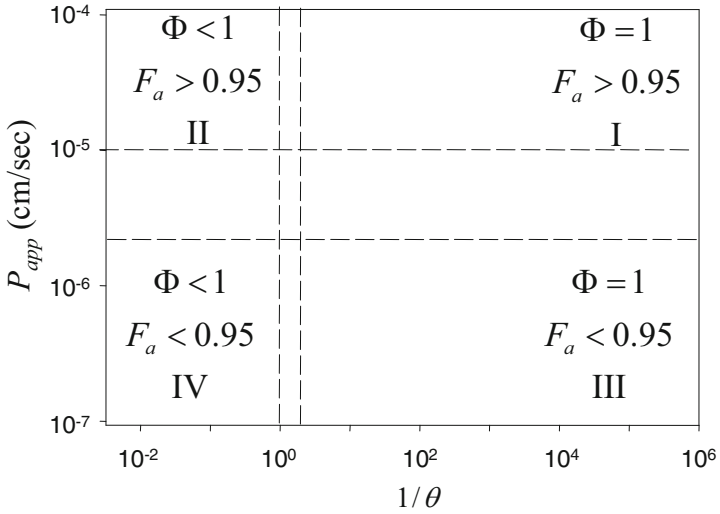


Fig. 6.16 The Quantitative Biopharmaceutics Classification System (QBCS) utilizes specific cutoff points for drug classification in the solubility–dose ratio ($1/\theta$), apparent permeability (P_{app}) plane. Each class of the QBCS can be characterized on the basis of the anticipated values for the fraction of dose absorbed, F_a and the fraction of dose dissolved, Φ at the end of the dissolution process assuming no interplay between dissolution and uptake. In essence the classification system is static in nature

$(1/\theta) \geq 1$) and incomplete dissolution (when $(1/\theta) < 1$) [92]. To account for variability related to the volume content, a boundary region of 250 to 500 ml was assumed and thus a boundary region for $1/\theta$ was set from 1 to 2. The boundary region of highly permeable drugs, P_{app} values in the range $2 \times 10^{-6} - 10^{-5}$ cm s^{-1} on the y-axis of Figure 6.16, can ensure complete absorption. It was based on experimental results [172–174], which indicate that drug absorption in Caco-2 monolayers can model drug transport in vivo.

In full analogy with BCS [157], the QBCS [228] classifies drugs into four categories based on their permeability (P_{app}) and solubility–dose ratio $1/\theta$ values defining appropriate cutoff points. For category I (high P_{app}, high $1/\theta$), complete absorption is anticipated, whereas categories II (high P_{app}, low $1/\theta$) and III (low P_{app}, high $1/\theta$) exhibit solubility–dose ratio- and permeability-limited absorption, respectively. For category IV (low P_{app}, low $1/\theta$), both permeability and solubility–dose ratio are controlling drug absorption. A set of 42 drugs was classified into the four categories of QBCS [228] and the predictions of their intestinal drug absorption were in accord with the experimental observations, Figure 6.17. However, some of the drugs classified in category II of the QBCS (or equivalently Class II of the BCS) exhibit a greater extent of absorption than the theoretically anticipated value based on a relevant semiquantitative analysis of drug absorption [228].

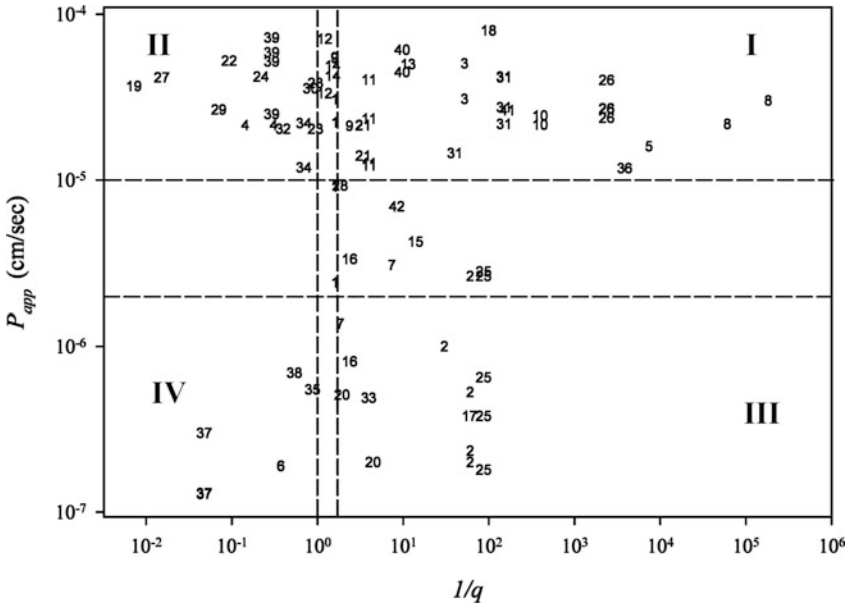


Fig. 6.17 The classification of 42 drugs in the (solubility–dose ratio, apparent permeability) plane of the QBCS. The intersection of the *dashed lines* drawn at the cutoff points form the region of the borderline drugs. Key: 1, acetyl salicylic acid; 2, atenolol; 3, caffeine; 4, carbamazepine; 5, chlorpheniramine; 6, chlorothiazide; 7, cimetidine; 8, clonidine; 9, corticosterone; 10, desipramine; 11, dexamethasone; 12, diazepam; 13, digoxin; 14, diltiazem; 15, disopyramide; 16, furosemide; 17, gancidovir; 18, glycine; 19, griseofulvin; 20, hydrochlorothiazide; 21, hydrocortisone; 22, ibuprofen; 23, indomethacine; 24, ketoprofen; 25, mannitol; 26, metoprolol; 27, naproxen; 28, panadiplon; 29, phenytoin; 30, piroxicam; 31, propanolol; 32, quinidine; 33, ranitidine; 34, salicylic acid; 35, saquinavir; 36, scopolamine; 37, sulfasalazine; 38, sulpiride; 39, testosterone; 40, theophylline; 41, verapamil HCl; 42, zidovudine

6.6.2 The Problem with the Biowaivers

According to the FDA guidance [227], petitioners may request biowaivers for high solubility–high permeability substances (Class I of BCS) formulated in immediate release dosage forms that exhibit rapid in vitro dissolution as specified above. The scientific aspects of the guidance as well as issues related to the extension of biowaivers using the guidance have been the subjects of extensive discussion [229, 230]. Furthermore, Yazdanian et al. [231] suggested that the high solubility definition of the FDA guidance on BCS is too strict for acidic drugs. Their recommendation was based on the fact that several nonsteroidal anti-inflammatory drugs (NSAID) exhibit extensive absorption and, according to the current definition of the FDA guidance, are classified in Class II (low soluble–high permeable) of the BCS. An important concluding remark of this study [231] is "an inherent limitation in the solubility classification is that it relies on equilibrium solubility determination, which is static and does not take into account the dynamic nature

Table 6.1 Dose and human bioavailability data of NSAIDs [227].

no.	**Drug**	**Highest Dose** (mg)	**Bioavailability** (%)
1	Diclofenac	50	54
2	Etodolac	400	> 80
3	Indomethacin	50	98
4	Ketorolac	20	100
5	Sulindac	200	88
6	Tolmetin	600	> 90
7	Fenoprofen	600	85
8	Flurbiprofen	100	92
9	Ibuprofen	800	> 80
10	Ketoprofen	75	100
11	Naproxen	500	99
12	Oxaprozin	600	$95 - 100$
13	Mefenamic acid	250	Rapidly absorbed
14	Acetylsalicylic acid	975	68 (unchanged drug)
15	Diflunisal	500	90
16	Salicylic acid	750	100
17	Meloxicam	15	89
18	Piroxicam	20	Rapidly absorbed
19	Celecoxib	200	-
20	Rofecoxib	25	93

of absorption." Moreover, the measurement of intrinsic dissolution rates [232] or the use of dissolution–absorption in vitro systems [233] appears more relevant than solubility to the in vivo drug dissolution dynamics for regulatory classification purposes. Also, the development of QBCS [228] is based on the key role of the solubility–dose ratio for solubility classification, since it is inextricably linked to the dynamic characteristics of the dissolution process [92]. All these observations point to the need for involvement of the dynamics of dissolution and uptake processes for the regulatory aspects of biopharmaceutical drug classification.

Recently, this type of analysis was attempted [234] for several nonsteroidal anti-inflammatory drugs listed in Table 6.1, which are currently classified as Class II drugs. The dynamics of the two consecutive drug processes, dissolution and wall permeation, were considered in the time domain of the physiologic transit time using a tube model that considers constant permeability along the intestines, a plug flow fluid with the suspended particles moving with the fluid, and dissolution in the small-particle limit. The radius of the spherical drug particles, ρ, and the concentration of dissolved drug in the intestinal tract, $c\,(z)$, are modeled as suggested by Oh et al. [56] for the development of BCS [157] by a system of differential equations, with independent variable the axial intestinal distance z, which is considered to be proportional to time, since the fluid flow rate is constant:

$$\begin{aligned} \frac{\mathrm{d}\rho(z)}{\mathrm{d}z} &= -\frac{\mathcal{D}\pi R^2}{Q\varrho}\frac{c_s - c(z)}{\rho(z)}, & \rho\,(0) &= \rho_0, \\ \frac{\mathrm{d}c(z)}{\mathrm{d}z} &= \frac{\mathcal{D}(n/V)4\pi^2 R^2}{Q}\rho\,(z)\,[c_s - c\,(z)] - \frac{2P_{eff}\pi R}{Q}c\,(z)\,, & c\,(0) &= 0, \end{aligned}$$

where $\mathcal{D}$ is the diffusion coefficient of the drug, ϱ is the density of the solid drug, R is the radius of the intestinal lumen, c_s is the solubility of the drug, Q is the volumetric flow rate, n is the number of drug particles in the dose, V is the luminal volume, and P_{eff} is the effective permeability of the drug.

These equations can be rewritten with respect to time if one multiplies both sides by $L/MITT$ (where L is the length of the tube and $MITT$ is the mean intestinal transit time) and simplifies:

$$\begin{aligned}\dot{\rho}(t) &= -\frac{\mathcal{D}}{\varrho}\frac{c_s - c(t)}{\rho(t)},\\ \dot{c}(t) &= \frac{3\mathcal{D}}{\varrho V}\frac{q_0}{\rho_0^3}\rho(t)\left[c_s - c(t)\right] - \frac{2P_{eff}}{R}c(t),\end{aligned}$$

where q_0 is the dose and ρ_0 is the initial radius of the drug particles.

Both sides of the last two equations are divided by q_0/V, and $c(t)$ and c_s are substituted with the fraction $\varphi(t)$ of dose dissolved and the dimensionless solubility–dose ratio θ, respectively, yielding

$$\begin{aligned}\dot{\rho}(t) &= \begin{cases} -\frac{\mathcal{D}}{\varrho}\frac{q_0}{V\rho(t)}\left[\frac{1}{\theta} - \varphi(t)\right] & \text{if } \rho(t) > 0, \\ 0 & \text{if } \rho(t) = 0, \end{cases} \quad \rho(0) = \rho_0,\\ \dot{\varphi}(t) &= \frac{3\mathcal{D}}{\varrho V}\frac{q_0}{\rho_0^3}\rho(t)\left[\frac{1}{\theta} - \varphi(t)\right] - \frac{2P_{eff}}{R}\varphi(t), \quad \varphi(0) = 0.\end{aligned} \tag{6.18}$$

The mass balance equation for the fraction F_a of dose absorbed at the end of the tube is

$$F_a = \frac{1}{q_0}\left[q_0 - q_{solid} - q_{dissolv}\right],$$

where q_{solid} and $q_{dissolv}$ denote the mass of the undissolved and dissolved drug, respectively, at the end of the intestine. This equation simplifies to the following:

$$F_a = 1 - \left[\frac{\rho(MITT)}{\rho(0)}\right]^3 \Phi, \tag{6.19}$$

where $\rho(MITT)$, and Φ refer to their values at $t = MITT = 199\,\text{min}$ [177].

The system of (6.18) and (6.19) describes the intestinal drug absorption as a function of four fundamental drug/formulation properties: dose q_0, solubility–dose ratio θ, initial radius of the particles ρ_0, and effective permeability P_{eff}. Typical values can be used for constants $\mathcal{D}$ ($10^{-4}\,\text{cm}^2\,\text{min}^{-1}$), ϱ ($1000\,\text{mg}\,\text{ml}^{-1}$), V (250 ml), and R (1 cm) [56]. Thus, one can assess, using (6.18) and (6.19), whether practically complete absorption ($F_a = 0.90$) of category II drugs of the QBCS is feasible by setting the permeability in (6.18) equal to $P_{eff} = 1.2 \times 10^{-2}\,\text{cm}\,\text{min}^{-1}$, which is equivalent [174] to the upper boundary limit $P_{app} = 10^{-5}\,\text{cm}\,\text{s}^{-1}$ of the apparent permeability borderline region of QBCS [228], Figure 6.16. The correlations developed [174] between effective permeability P_{eff}, values determined in humans and the Caco-2 system allowed the conversion of the Caco-2 to P_{eff} estimates. Figure 6.18 shows the simulation results in a graph of q_0 vs. $1/\theta$ for

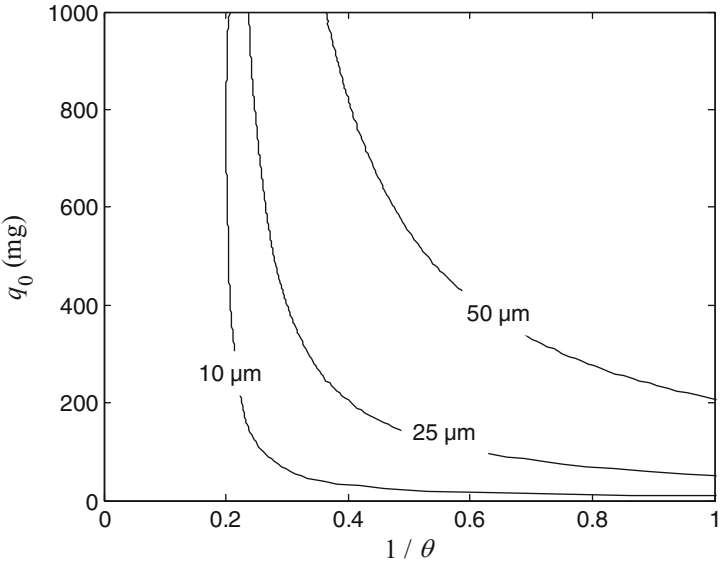


Fig. 6.18 Plot of dose q_0 vs. the dimensionless solubility–dose ratio $1/\theta$. The curves indicate 90% absorption for three radius sizes 10, 25, and 50 μm assuming $P_{eff} = 1.2 \times 10^{-2}$ cm min^{-1}. Since the value assigned to P_{eff} corresponds to the upper boundary limit (expressed in apparent permeability values, [174]) of the borderline permeability region of QBCS [228], compounds of category II of QBCS exhibiting complete absorption are located above the curves

the three particle sizes $\rho_0 = 10$, 25, and 50 μm. The areas above the lines, for each of the particle sizes considered, correspond to drug/formulation properties q_0, $1/\theta$, ensuring complete absorption, i.e., $F_a > 0.90$ for drugs classified in category II of the QBCS [228]. It is worth noting that for a given value of $1/\theta$, a higher fraction of dose is absorbed from a larger rather than a smaller dose. This finding is reasonable since the common $1/\theta$ value ensures higher solubility for the drug administered in a larger dose.

The underlying reason for a region of fully absorbed drugs in category II of the QBCS, shown in Figure 6.18, is the dynamic character of the dissolution–uptake processes. A global measure of the interplay between dissolution and uptake can be seen in Figure 6.19, which shows the mean dissolution time, *MDT*, in the intestines as a function of the effective permeability for a Class II drug ($1/\theta = 0.2$). Clearly, the *MDT* value is reduced as effective permeability increases. Needless to say that the *MDT* would be infinite for this particular drug ($1/\theta = 0.2$) if dissolution were considered in a closed system ($P_{eff} = 0$) [92]. The plot of Figure 6.19 verifies this observation since $MDT \to \infty$ as $P_{eff} \to 0$.

According to Yazdanian et al. [231] most of the NSAIDs listed in Table 6.1 are classified in Class II based on their solubility data at pH 1.2, 5.0, and fed state simulated intestinal fluid at pH 5.0. A series of simulations based on (6.18) and (6.19) revealed that the extensive absorption (Table 6.1) of the NSAIDs can

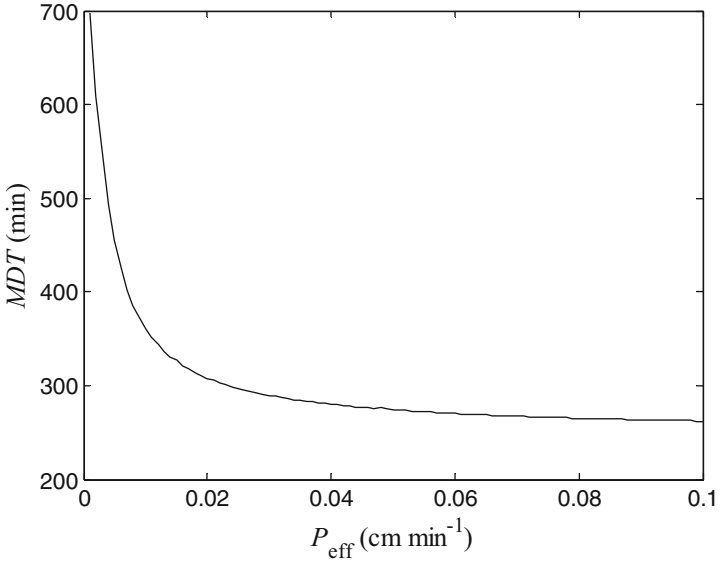


Fig. 6.19 The mean dissolution time *MDT* in the intestines as a function of P_{eff} for parameter values $q_0 = 10\,\text{mg}$, $(1/\theta) = 0.2$, and $\rho(0) = 10\,\mu\text{m}$. *MDT* is calculated as the area under the curve of the undissolved fraction of dose using the integral $MTD = \int_0^\infty \left[\frac{\rho(t)}{\rho(0)}\right]^3 \text{d}t$ in conjunction with (6.18)

be explained using the solubility–dose ratio values in buffer or fed state simulated intestinal fluid, both at pH 5.0, Figure 6.20. This plot shows the experimental data along with the curves generated from (6.18) and (6.19) assuming $F_a = 0.90$, radius sizes 10 and 25 μm, and assigning $P_{eff} = 2 \times 10^{-2}\,\text{cm}\,\text{min}^{-1}$, which corresponds [174] to the mean of the apparent permeability values of the NSAIDs ($P_{app} = 1.68 \times 10^{-5}\,\text{cm}\,\text{s}^{-1}$) [231]. Visual inspection of the plot based on the solubility at pH 5.0, Figure 6.20A, reveals that only the absorption of sulindac (no.5, $F_a = 0.88$) can be explained by the generated curve adhering to 25 μm, while flurbiprofen (no.8, $F_a = 0.92$) lies very close to the theoretical line of 10 μm.

In contrast, the extensive absorption of tolmetin (no.6, $F_a > 0.90$), sulindac (no.5, $F_a = 0.88$), etodolac (no.2, $F_a > 0.80$), diflunisal (no.15, $F_a = 0.90$), ibuprofen (no.9, $F_a > 0.80$), using the corresponding doses listed in Table 6.1, can be explained on the basis of the solubility data in fed state simulated intestinal fluid at pH 5.0, in conjunction with the generated curve assigning $\rho(0) = 25\,\mu$m, Figure 6.20B. Also, the curve generated from $\rho(0) = 10\,\mu$m and the solubility in the biorelevant medium of indomethacin (no.3) and piroxicam (no.18) explain their extensive absorption. Although naproxen (no.11, $F_a = 0.99$) lies very close and meloxicam (no.17, $F_a = 0.89$) in the neighborhood of the theoretical line of 10 μm, oxaprozin (no.12, $F_a = 0.95 - 1.00$) is located far away from the simulated curve of 10 μm, Figure 6.20B. Special caution is required in the interpretation for diclofenac

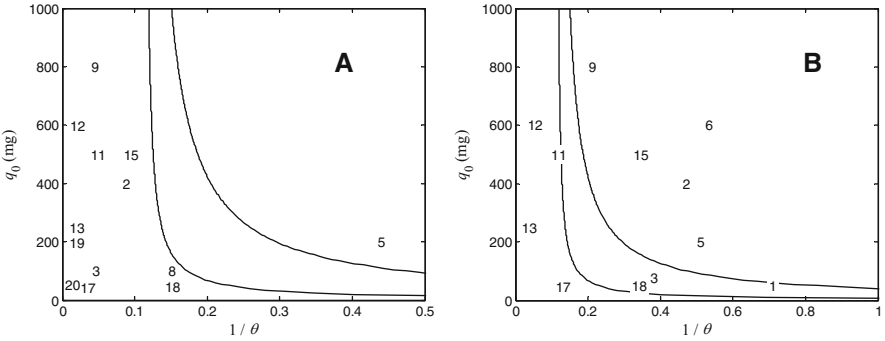


Fig. 6.20 Plot of q_0 vs. $1/\theta$, for the experimental data of Table 6.1 classified in Class II. The curves denote 90% absorption for two particle sizes (from left to right 10 and 25 μm) assigning $P_{eff} = 2 \times 10^{-2}$ cm min^{-1}, which corresponds [174] to the mean, $P_{app} = 1.68 \times 10^{-5}$ cm s^{-1} of the Caco-2 permeability values of the data [92]. Drugs located above the curves are fully absorbed ($F_a > 0.90$) Class II drugs. Key (solubility values in): (**A**) buffer, pH 5.0; (**B**) fed state simulated intestinal fluid, pH 5.0

(no.1, $F_a = 0.54$), which lies between the theoretical curves of 10 and 25 μm in Figure 6.20B. Some reports suggest that diclofenac undergoes first-pass metabolism ($F_a = 0.60$), while some others refer to absolute bioavailability 0.90 [235]. Explicit data for the extent of absorption of mefenamic acid (no.13), Figure 6.20B, are not reported [231], while solubility data in the fed state simulated intestinal fluid (pH 5.0) for the two nonacidic NSAIDs, celecoxib (no.19) and rofecoxib (no.20), have not been measured [231].

These results point out the importance of the dynamic nature of the absorption processes for those drugs classified in Class II. It should also be noted that a conservative approach was utilized for the interpretation of the NSAIDs' extensive absorption, Table 6.1. In fact, only the highest doses of drugs were analyzed, while the duration of absorption was restricted to the mean intestinal transit time, 199 min [177], i.e., absorption from the stomach or the large intestine was not taken into account. Moreover, the lower value for the volume of the intestinal content, 250 ml [228–230], was used in the simulations. This means that drugs like naproxen (no.11) and meloxicam (no.17) in Figure 6.20B would also have been explained if higher values of the two physiological parameters for time and volume had been used.

For the sake of completeness one should also add that Blume and Schug [236] suggested that Class III compounds (high solubility and low permeability) are better candidates for a waiver of bioavailability and bioequivalence studies since bioavailability is not so much dependent on the formulation characteristics as on the permeability of the compound. According to the European Medicines Agency guidance [237], petitioners may request biowaivers for Class III compounds; however, the most recent BCS-based FDA guideline issued in May 2015 adopts the biowaiver status for Class III compounds.

6.6.3 Biowaiver Monographs: BCS Considerations

Since the biowaiver status is a real alternative to in vivo pharmacokinetic bioequivalence studies, its importance for the pharmaceutical industry is more than obvious. The term "biowaivers" refers to all exceptions from the necessity to perform clinical studies. Accordingly, a large number of papers were published in this area of research after the publication of the FDA BCS guidance [227].

In fact, this guidance triggered off the development of biorelevant media in particular for solubility–dissolution studies of Class II compounds. Several review articles dealing with the use of biorelevant media for the in vitro testing of orally administered dosage forms have been published recently, e.g., [238]. Although the biorelevant media have not been officially adopted from the drug Agencies, they are useful for the assessment of drug dissolution in specific cases, e.g., locally acting drugs in the gastrointestinal tract.

In the same vein, twenty seven "biowaiver monographs" have published in the literature (a complete list can be found in [239]). Each one of the "biowaiver monographs" analyses a drug which can or cannot be considered as a Class I drug on the basis of the strict regulatory definitions [227, 237]. The authors of the "biowaiver monographs" use arguments based on the physicochemical and biopharmaceutical properties of drug and its pharmacokinetic-dynamic characteristics to substantiate drug's (in)eligibility for a biowaiver status.

However, the most important developments in the regulatory and scientific aspects of BCS are associated with the experimental–theoretical work exploring the limitations of the static–binary classification of all drugs in the four BCS classes as well as the definition of a Class I drug [240]. This review article provides an overview of the recent developments of science and regulation in oral drug absorption and focus on the various drug properties and processes in the milieu of gastrointestinal lumen, e.g., dose, solubility, permeability, dissolution, precipitation which are directly or indirectly associated with the application of BCS. One notable example is the importance of dose for the biopharmaceutic classification of drugs [92, 228, 231, 234, 241, 242]; thus, the WHO proposal on in vivo bioequivalence requirements does not only allow biowaivers for BCS Class II substances but also utilizes the solubility–dose ratio for biopharmaceutical classification purposes [243]. It should be noted that one of the requirements of both the previous EMA 2001 Guideline and the current FDA 2000 Guideline [237] specifies that "the (marketed) highest dosage strength" should be dissolved in 250 ml for getting the biowaiver status regarding the solubility criterion. However, the recent revised EMA 2010 Guideline [227] defines dose as the "highest single oral immediate release dose" recommended for administration in the summary of product's characteristics. The impact of this change has been analyzed recently [244] in terms of the biowaiver monographs for 27 active pharmaceutical ingredients published in the literature. Moreover, another recent study recommends that each dose strength be considered separately, i.e., whether or not it meets the solubility–dissolution regulatory criteria [239].

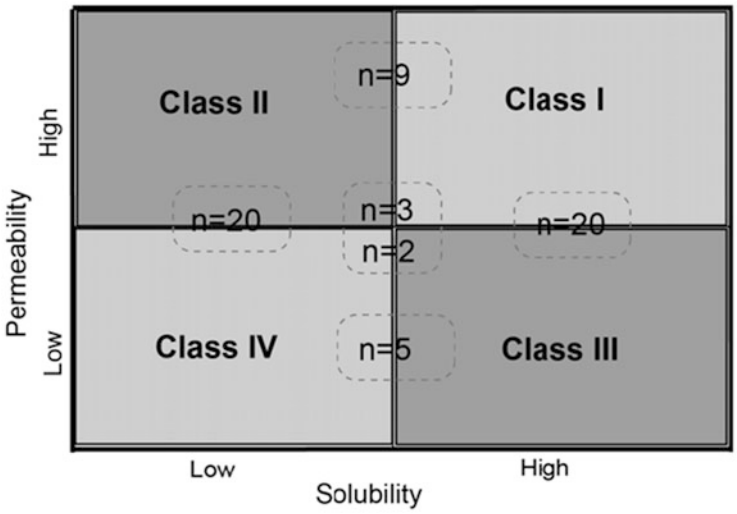


Fig. 6.21 A pictorial view of the 59 compounds, which have been assigned to multiple BCS classes in different papers [245]. The positioning of the encircled numbers of the plot corresponds to the boundary of the classes for each one of the 59 compounds reported in the literature. The numbers in the center of the rectangular for dual classification $n = 3$ and $n = 2$, refer to Classes I-IV and Classes II-III, respectively

Another notable example is the dual biopharmaceutic classification of a great number of drugs [245]. A pictorial view of the classification of 59 drugs in more than one BCS Classes is presented in Figure 6.21. According to this study [245], a more relevant pH restriction for acids and/or dissolution medium with lipids present better forecast solubility-limited absorption in vivo than the presently used BCS solubility criterion. Along these lines, Macheras and Karalis [246] introduced a non binary biopharmaceutic classification system, the so-called ABΓ system, Figure 6.22. This approach relies on the mathematical model used for the development of BCS [157], appropriately modified, to estimate the limiting values of drug solubility and permeability when the fraction of dose absorbed, F_a was 0.90 or 0.20. The continuity of the biopharmaceutical classification is ensured since the first category (A, alpha) includes drugs with $F_a \geq 0.90$, the B (beta) category consists of drugs with $F_a \leq 0.20$ while the area lying between the two boundaries of A and B defines the third category Γ (gamma), ($0.20 < F_a < 0.90$), Figure 6.22. It was found that most of the BCS classes II and III are included in category Γ which mainly consists of drugs with properties like moderate or low solubility and permeability; besides, the dynamic character of dissolution and uptake processes explains why category A is expanded toward BCS Class II, Figure 6.22.

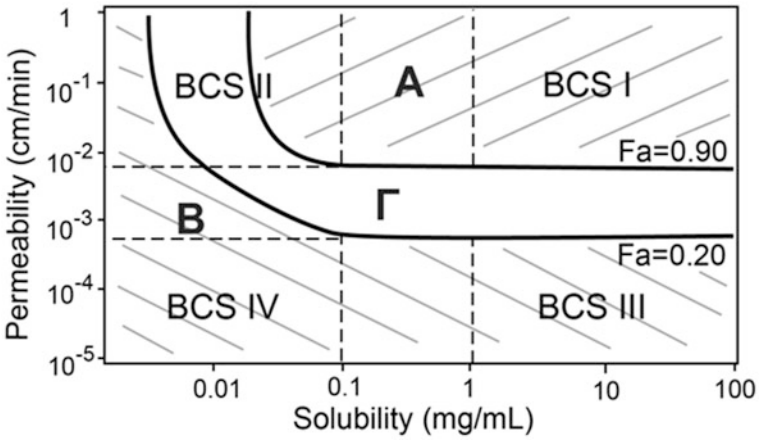


Fig. 6.22 The ABΓ system [246] is coplotted with a continuous version of BCS using upper and lower solubility limits 0.1 and 1 mg ml^{-1}, for Classes II, IV and I, III, respectively, while the lower and upper limits for permeability have been calculated numerically [246] applying the model used for the development of BCS [157] and assigning $F_a = 0.90$ or $F_a = 0.20$. In all cases, 'Dose' was equal to 10 mg, while the drug's particle size radius was 0.010 mm

6.6.4 Biopharmaceutics Drug Disposition Classification System

In the early-mid dates of 2000 decade, Professor Leslie Benet questioned, in a number of talks, the ability of a single permeability estimate to predict the extent of drug absorption along the lines of BCS. His arguments were based on the fact that permeability (expressed in velocity units) is a rate and not an extent parameter metric. As a matter of fact, a rate parameter like permeability can be also used as a predictor of extent of absorption; however, the morphological–functional–dynamical complexity of the gastrointestinal lumen and tract does not allow the reliable use of permeability as a sole parameter of drug's permeation across the gastrointestinal tract [248]. Based on these concerns, Wu and Benet [247] developed in 2005 the so-called Biopharmaceutics Drug Disposition Classification System (BDDCS), Figure 6.23. According to this Figure, the extent of metabolism (either low or high) replaces permeability in the four classes of BCS, Figure 6.15. It is worthy to mention that the EMA 2010 Guideline [227] adopted and assigned the "$\geq$ 90% metabolized" as cut-off limit, namely, an alternative criterion for the extent of absorption for Class I biowaivers. Strictly speaking the EMA 2010 Guideline [227] specifies "following a single oral dose to humans, administered at the highest dose strength, mass balance of Phase 1 oxidative and Phase 2 conjugative drug metabolites in the urine and feces, account for $\geq$ 90% of the dose administered." In parallel, Benet and Larregieu [249] stated that "although FDA-approved BCS Class I drugs are designated as high-permeability drugs, in fact, the criterion

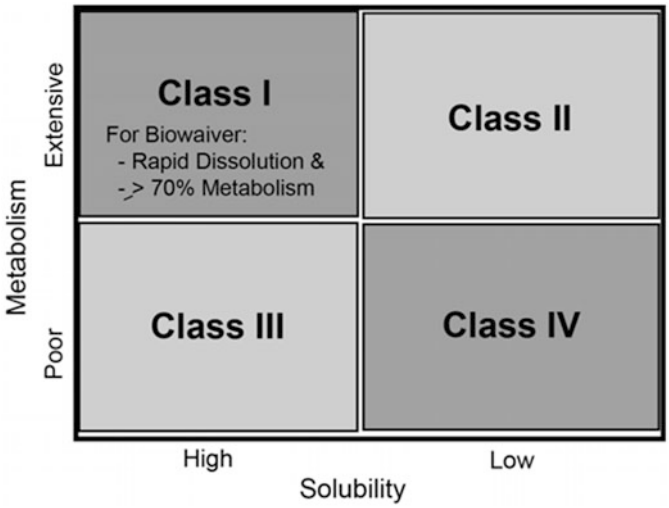


Fig. 6.23 The Biopharmaceutics Drug Disposition Classification System as proposed by Wu and Benet [247]

utilized is high extent of absorption. This ambiguity should be eliminated, and the FDA criterion should explicitly be stated as ≥ 90% absorption based on absolute bioavailability or mass balance."

The publication of BDDCS attracted the interest of scientists since it extends BCS toward drug elimination phenomena and the effects of efflux and transporters on oral drug absorption. Overall, the BDDCS is a useful tool in predicting (1°) drug disposition when transporter–enzyme interplay will yield clinically significant effects, (2°) the direction, mechanism, and importance of food effects, and (3°) the transporter effects on post-absorption systemic drug concentration following oral and intravenous dosing [250, 251]. A large number of studies followed the publication of the seminal BDDCS paper [247]. For example, in silico approaches were used in order to predict the BDDCS class for new compounds using molecular structure and available molecular descriptors and software [252] while classification of Class I marketed drugs was based on estimates of drug permeability vis a vis extent of drug metabolism (≥ 90% metabolized) [253]. Most importantly, Benet and coworkers [254, 255] classified over 900 drugs using BDDCS criteria and also applied a computational approach to predict BDDCS class of new molecular entities from molecular structures. Both studies revealed the importance of solubility–dose ratio for BDDCS classification. This finding not only coincides with the importance of dose for the classification of drugs in BCS mentioned above [92, 228, 231, 234, 241–243] but also emphasizes the utility of the concepts of critical dose, effective in vivo solubility, and dose-dependent BCS developed in [242]. Although the complimentary role of BCS and BDDCS in the improvement, simplification, and speed of drug development has been recognized [256], issues associated with differences in the drug's permeability considerations–mechanisms of the two systems are still open [257, 258].

6.7 Randomness and Chaotic Behavior

Pharmacokinetic studies are in general less variable than pharmacodynamic studies because simpler dynamics are associated with pharmacokinetic processes. According to van Rossum and de Bie [259], the phase space of a pharmacokinetic system is dominated by a point attractor since the drug leaves the body, i.e., the plasma drug concentration tends to zero. Even when the system is as simple as that, tools from the dynamic systems theory are still useful. When a system has only one variable a plot referred to as a *phase plane* can be used to study its behavior. The phase plane is constructed by plotting the variable against its derivative. The most classical phase plane, quoted even in textbooks, is the $\dot{c}(t)$ vs. $c(t)$ plot of the ubiquitous Michaelis–Menten kinetics. In the pharmaceutical literature the phase-plane plot was used by Dokoumetzidis and Macheras [260] for the discernment of absorption kinetics, Figure 6.24. The same type of plot was used for the estimation of the elimination rate constant [261].

A topic in which dynamic systems theory has a potential use is the analysis of variability encountered in bioavailability and bioequivalence studies with highly variable orally administered formulations [262–264]. For example, the dissolution of a sparingly soluble drug takes place in the continuously changing environment of the gastrointestinal lumen. Due to the interactive character of the three principal physiological variables that affect drug dissolution, i.e., the motility of intestines,

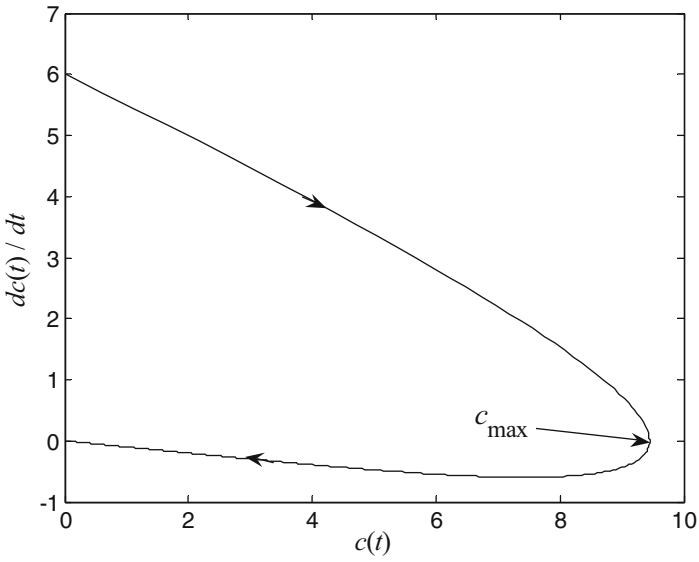


Fig. 6.24 Phase plane plot for a drug obeying one-compartment model disposition with first-order absorption and elimination. Time indexes each point along the curve. The time flow is indicated by the arrows, while the x-axis intercept corresponds to c_{max}

the composition and volume of gastrointestinal contents, a dynamic system of low dimension can be envisaged. If this is a valid hypothesis, a significant portion of the high variability encountered in the gastrointestinal absorption studies can be associated with the dynamics of the physiological variables controlling drug dissolution, transit, and uptake. However, the inaccessibility of the region and thus the difficulty of obtaining detailed information for the variables of interest compel one to infer that the observed variability originates exclusively from classical randomness.

Despite the hypothetical character of the previous paragraph, recent findings [265] have revealed the chaotic nature of the gastric myoelectrical complex. It seems likely that the frequently observed high variability in gastric emptying data should not be attributed exclusively to the classical randomness of rhythmic electrical oscillation in the stomach. Plausibly, one can argue that this will have an immediate impact on the absorption of highly soluble and permeable drugs from immediate release formulations since their absorption is controlled by the gastric emptying rate. Hence, the high variability of $c_{\max}$ values for this type of drug originates from both classical experimental errors and the chaotic dynamics of the underlying processes.

Finally, the heterogeneous dynamic picture of the gastrointestinal tract becomes even more complicated by the coexistence of either locally or centrally driven feedback mechanisms, e.g., avitriptan controlling drug absorption. Experimental observations indicate [266] that when avitriptan blood levels exceed a certain threshold level, a centrally driven feedback mechanism that affects gastric emptying is initiated. Consequently, the presence or absence of double or multiple peaks of avitriptan blood levels is associated with the dynamic system describing the dissolution and uptake of drug as well as the feedback mechanism controlling the functioning of the pylorus.

It can be concluded that the use of nonlinear dynamics in gastrointestinal absorption studies can provide a tool for:

- the interpretation of variability and
- the understanding of unpredictability in situations in which double, or multiple peaks are observed and classical explanations, e.g., enterohepatic cycling, are not applicable.

Part III
Modeling in Pharmacokinetics

The kinetic description of the overall disposition of drug in the body is based on experimental observations, most frequently time–drug plasma concentration data. In essence, the quantitative description of these kinetic patterns constitutes the discipline of pharmacokinetics.

Pharmacokinetic models are, or at least should be, created through a repeated *seesaw* process of monitoring against experimental evidence. Starting from a comparatively simple model, a comparison with experimental evidence will normally point to modifications of the original model, which in turn will call for new experimental evidence to assist a decision between alternative model modifications, and so on.

Mathematical models are typically classified as *phenomenological* or *empirical* [267, 268]. For the construction of phenomenological models, one must have some knowledge of the processes in terms of structural connectivity and functional mechanisms. On the other hand, when the underlying processes are unknown, one often relies on mathematical functional forms for the observed profile. These mathematical functions are empirical models. However, one can start with an empirical model and incorporate some mechanistic assumptions so that the model looks more "phenomenological." For this reason, a clear borderline cannot be drawn between phenomenological and empirical models. The full contrast between phenomenological and empirical models can be exemplified by the physiologically based and the input–output models. Neither of these extreme cases is suitable because:

- in the physiologically based models the compiled information is highly complex implying a decline in prediction performances and
- in the input–output models only a rough description of the external behavior of the process is established.

We believe that current science is a dynamic process of knowledge requiring updating the formal tools of analysis that are our mathematical models. For example, the starting point may be a purely phenomenological or purely empirical model,

but as the knowledge is accumulated, the phenomenological must become more empirical, or the empirical more phenomenological.

Most models should be considered "temporary." All require validation and many will be subject to change. Modeling may be viewed either as a screening process that employs variable selection methods to construct candidate models, or as a testing tool that validates a specific model. From a mathematical point of view, the art of good modeling relies on:

- a sound understanding and appreciation of the biological problem,
- a realistic mathematical representation of the important biological processes,
- finding useful solutions, preferably quantitative; and most importantly,
- a biological interpretation of the mathematical results in terms of insights and predictions.

In this part of the book, we first present in Chapter 7 the empirical pharmacokinetic models used to fit the observed kinetic data, placing particular emphasis on the relevance of power functions and the heterogeneous processes. The deterministic compartmental approach is described in Chapter 8 as the standard technique to analyze homogeneous processes. In Chapter 9, the application of fractional differential in equations in pharmacokinetics are presented. Chapter 10 focuses on the application of Modeling and Simulation approaches in bioequivalence assessment. The last chapter of this part, Chapter 11, is devoted to the stochastic modeling techniques, a powerful tool in mathematical biology suitable for the analysis of both homogeneous and heterogeneous processes.

Chapter 7
Empirical Models

It is through a few empirical functions that I am able to approach contemplation of the whole.

William A. Calder III (1934–2002)
Size, function and life history

In experimental or clinical pharmacokinetics, the simplest experiment consists in administering, in a rapid input, a large number of drug molecules having the same pharmacological properties and then in the subsequent time interval, sampling biological fluids in order to follow the decline in number of molecules or in drug concentration. The investigators are primarily interested in describing the observed decrease in time of the data by simple mathematical functions called *empirical models*. The most commonly employed model profiles are the negative exponential, the power-law, and the gamma profiles.

Exponential Profiles These have the form $c(t) = \gamma \exp(-\beta t)$. Differentiating with respect to time, one obtains

$$\dot{c}(t) = -\beta c(t), \qquad \text{or} \qquad \frac{[\mathrm{d}c(t)/c(t)]}{\mathrm{d}t} = \frac{\mathrm{d}\ln c}{\mathrm{d}t} = -\beta, \tag{7.1}$$

i.e., "the *relative* variation of the concentration c of the material divided by the *absolute* variation of time t is constant," which is the expression of Fick's law (cf. Section 2.3 and equation 2.14) under the assumption of constant volume of distribution V of the material in the medium. The constant β with dimension time^{-1} represents the ratio of the clearance CL to the volume V.

Power-Law Profiles These profiles follow the form $c(t) = \gamma t^{-\alpha}$. Differentiating with respect to time, one obtains

$$\dot{c}(t) = -\frac{\alpha}{t} c(t), \qquad \text{or} \qquad \frac{[\mathrm{d}c(t)/c(t)]}{[\mathrm{d}t/t]} = \frac{\mathrm{d}\ln c}{\mathrm{d}\ln t} = -\alpha, \tag{7.2}$$

i.e., "the *relative* variation of the concentration c of the material divided by the *relative* variation of time t is constant." Similarly, we can argue that the

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_7

dimensionless constant α relates to how many new molecules are eliminated from the experimental medium or from the body by a mechanism similar to the overall process as the time resolution becomes finer. Attention will be given below to clarifying the power law.

Gamma Profiles These profiles follow the form $c(t) = \gamma t^{-\alpha} \exp(-\beta t)$, which is reported in the literature as the gamma-function model [269]. This model was used to fit pharmacokinetic data empirically [270, 271]. Differentiating with respect to time, we obtain

$$\dot{c}(t) = -\left(\frac{\alpha}{t} + \beta\right) c, \tag{7.3}$$

i.e., the gamma profiles might be considered as the mixed exponential and power-law profiles; the general expression for the behavior of the process in specific cases becomes either exponential or power-law.

In the three profiles above, the coefficient γ is set according to the initial conditions. For instance, if $c(t_0) = c_0$ at $t_0 \neq 0$, γ is equal to

$$c_0 \exp(-\beta t_0) \qquad \text{or} \qquad c_0 (t_0)^{\alpha} \qquad \text{or} \qquad c_0 (t_0)^{\alpha} \exp(-\beta t_0)$$

for the exponential, power-law, or gamma model, respectively. Figure 7.1 illustrates, in linear, semilogarithmic, and logarithmic scales, the behavior of these basic

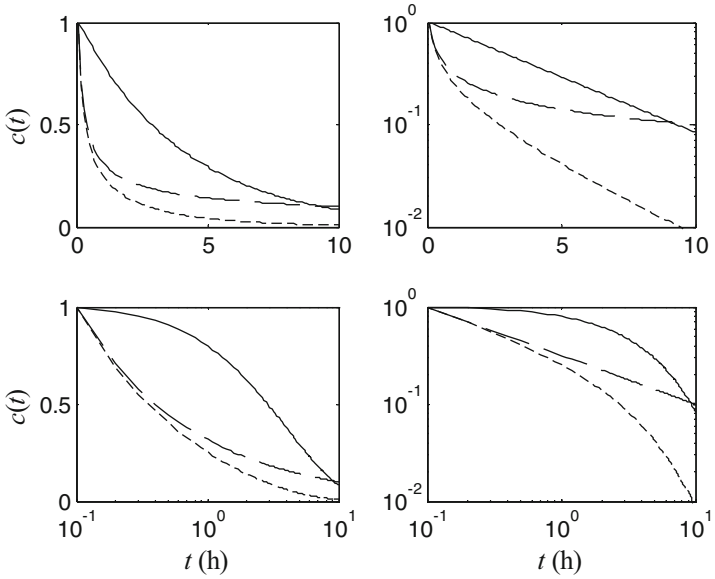


Fig. 7.1 Plots of the exponential, power-law, and gamma empirical models (*solid*, *dashed*, and *dotted lines,* respectively)

profiles with $\alpha = 0.5$, $\beta = 0.25$, and $c\,(0.1) = 1$. From these plots, we can decide in practice which empirical model we need to use:

- The y-semilogarithmic plot distinguishes the exponential model, which is depicted as one straight-line profile.
- The log–log plot distinguishes the power-law model, which is depicted as one straight-line profile.
- Both y-semilogarithmic and log–log plots are needed to decide for the gamma profile. It behaves like a power-law model in the early times (cf. the log–log plot) and as an exponential model in the later times (cf. the y-semilogarithmic plot).

The linear and x-semilogarithmic plots are uninformative for such decisions.

7.1 Power Functions and Heterogeneity

In a more realistic context, the observed data usually decay according to a sum of m negative exponentials

$$c\,(t) = \sum_{i=1}^{m} B_i \exp\,(-b_i t)\,,$$

which correspond to a series of well-stirred tanks where drug administration is in the first tank and the concentration is computed for the m-th tank.

In many cases, it was observed that when the fit of data improves as m increases, they would also be well fitted by a function of a negative power of time. It does seem extraordinary that the power function, with only two adjustable parameters, fits the data nearly as well as the sum of three or more exponential functions [269]. In fact, the scheme of the series of tanks corresponds to the states of a random walk that describes the retention of the molecules by movement of elements between nearest-neighbor sites from the administration to the sampling site. For large m, this random walk can be thought of as approximating a diffusion in a single heterogeneous site that is fitted by the empirical power-law model.

When the real process generates power-law data, alternatively a sum of exponentials and power function models may be used. But:

- power functions are defined by fewer parameters than the sums of exponentials;
- power functions seem to yield better long-term predictions;
- furthermore, the exponential parameters have little or no physiological meaning, under inhomogeneous conditions.

Overall, a large number of drugs that exhibit apparently multiexponential kinetics obey power-law kinetics. The cogent question is why many of the observed

time–concentration profiles exhibit power function properties. Although the origin of the power function remains unclear, some empirical explanations could elucidate its origin:

1. A power function can be related to the sum of an infinitely large number of exponential functions:

$$t^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^{\infty} u^{\alpha-1} \exp(-ut)\, \mathrm{d}u, \qquad \alpha > 0.$$

   Therefore, within a given range of time, the power functions can always be fitted by sums of negative exponentials within limits that are typical for experimental error. But the converse is not true: one cannot fit power functions to data generated by sums of negative exponentials.
2. Beard and Bassingthwaighte [272] showed that a power function can be represented as the sum of a finite number of scaled basis functions. Any probability density function may serve as a basis function. They considered as basis function a density corresponding to the passage time of a molecule through two identical well-stirred tanks in series. The weighted sum of such m models leads to the power function

$$t^{-\alpha} \propto \sum_{i=1}^{m} k_i^{\alpha+1} t \exp(-k_i t), \qquad \alpha > 0.$$

   This sum can also be viewed as the parallel combination of m pathways, each characterized by a different rate constant and a uniform distribution of flow in the input of these pathways. Then, the negative power function behavior can be attributed to the heterogeneity of the flow in the system.
3. Power functions can arise if the administered molecules undergo random walks with drift, as in the well-known Wiener process [273]. The concept of random walk in series can be expressed in terms of compartments in series that have one-way entrances and exits. Each series of compartments constitutes one region, and according to the inhomogeneous assumption the administered molecules move through such a region, while according to the homogeneous assumption they move randomly within it. The inhomogeneous process could be related to active transport, i.e., through membranes.

Therefore, it seems that when the response can be fitted by power-law empirical models, the underlying process is rather heterogeneous. This probably occurs because of inhomogeneous initial mixing and transport of the molecules by bloodstream that is understirred [274], or because of elimination of molecules by organs with structural heterogeneity. Perhaps the most obvious origin of the simple power function is a diffusion process that constitutes a rate-limiting step for removal of certain substances from the circulation [4]. Moreover, drug molecules can differ in their kinetic behavior because of inherent variability in their characteristics

such as molecular weight, chemical composition, or hepatic clearance involving a large number of metabolites. All these features introduce *functional heterogeneity*. Overall, homogeneity and heterogeneity can originate respectively when:

- Most substances intermix rapidly within their distribution spaces, and the rate-limiting step in their removal from the system is biochemical transformation or renal excretion. Substances of this nature are best described by compartmental models and exponential functions.
- Conversely, some substances are transported relatively slowly to their site of degradation, transformation, or excretion, so that the rate of diffusion limits their rate of removal from the system. Substances of this nature are best described by non-compartmental models and power functions.

7.2 Heterogeneous Processes

Description of distribution and elimination under homogeneous conditions can be done using classical kinetics, while fractal kinetics should be applied to describe distribution and elimination mechanisms under heterogeneous conditions. Classical transport theories, and the resulting mass-action kinetics, applicable to Euclidean structures do not apply to transport phenomena in complex and disordered media. The geometric constraints imposed by the heterogeneous fractal-like structure of the blood vessel network and the liver strongly modify drug dynamics [275]. Topological properties like connectivity and the presence of loops or dead ends play an important role. Hence, it is to be expected that media having different dimensions or even the same fractal dimension, but different spectral dimensions, could exhibit deviating behavior from that described by classical kinetics.

7.2.1 Distribution, Blood Vessels Network

According to Mandelbrot [276], fractal bifurcating networks mimic the vascular tree. Based on this observation, van Beek et al. [277] developed dichotomous branching fractal network models to explain the regional myocardium flow heterogeneity. Even though the developed models give overly simple descriptions of the fractal network, they describe adequately the dependence of the relative dispersion of flow distribution on the size of the supplied region of myocardium. These findings allow us to infer that such fractal approaches would be useful in describing other systems with heterogeneous flow distributions.

From a drug's site of administration, the blood is the predominant medium of transport of the molecules through the body to the drug's final destination. Conventionally, the blood is treated as a simple compartment, although the vascular system is highly complex and consists of an estimated 96, 000 km of vessels [278]. The key feature of the network is the continuous bifurcation of the parent vessels

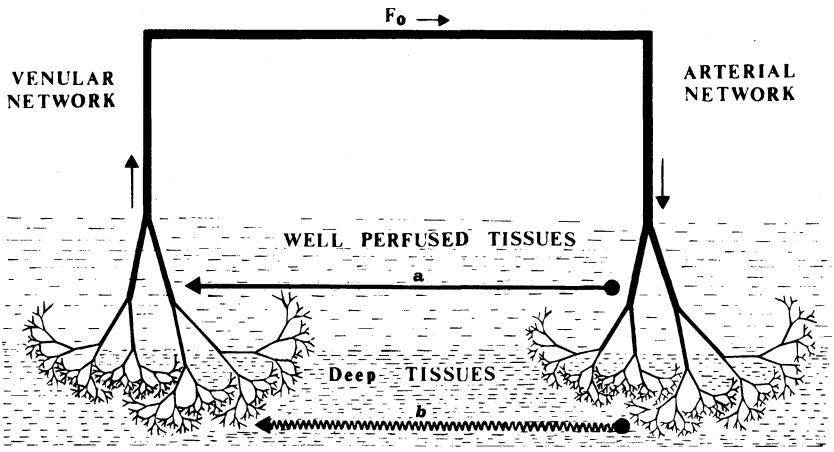


Fig. 7.2 A complete vascular dichotomous network used to describe the distribution of drug in the body. The black circle represents the drug molecules. (**a**) The distribution of drug in well-perfused tissues takes place under homogeneous (well-stirred) conditions. (**b**) The distribution of drug in deep tissues takes place under heterogeneous (understirred) conditions. Reprinted from [281] with permission from Springer

for many generations of branching. The vessels of one generation bifurcate to form vessels of the next generation in a continuous process toward smaller and smaller vessels. Some studies [279, 280] of the microvascular system have shown that the dimensions for vessel radii, branch length, and wall thickness in the mesenteric and renal arterial beds have fractal properties. The discovery of the fractal nature of the blood vessels, however, indicated that the distribution of flow within an organ might be fractal as well.

Building on the work of van Beek et al. [277], a dichotomous branching network of vessels representing the arterial tree connected to a similar venous network can be used to describe the distribution of the drug in the body, Figure 7.2. Thus, the general pattern of distribution of flow can also be assumed for the complete vascular system of Figure 7.2, envisaged for the distribution of drugs in the body. The flow will diverge in the arterial tree and converge in the venous tree, while at the ends of the arterial and venular networks the local flow will be slow and heterogeneous.

In the light of these network flow considerations, the distribution of drugs in the body can be classified into two broad categories. The distribution process of the drugs of the first category takes place under homogeneous (well-stirred) conditions. For the second category of drugs a significant part of the distribution process operates under heterogeneous (understirred) conditions.

- Drugs of the first category have physicochemical properties and permeability characteristics that allow them to leave the arteriole network and diffuse to the adjacent tissues under conditions of flow that ensure complete mixing (Figure 7.2 a). These drugs reach only the well- perfused tissues and return

rapidly to the venular draining network. The disposition of this category of drugs can be modeled with the "homogeneous model," which is identical mathematically to what we call the "one-compartment model." Obviously, the drug molecules obeying the homogeneous model permeate the walls of vessels prior to their arrival at the hugely dense ending of the networks; thus, the upper part of the vascular system and the well-perfused adjacent tissues comprise a homogeneous well-stirred "compartment."

- Based on the considerations of flow in the network, it is reasonable to argue that in close proximity with the terminal arteriolar ending, the blood flow and drug diffusion in the adjacent deep tissues will be so slow that the principle of the well-mixed system will no longer hold. Consequently, if a large portion of drug is still confined in the arterial system near its ending, the drug diffusion in the deep tissues will operate under heterogeneous (understirred) conditions (Figure 7.2 b). Transport limitations of drug in tissues have been dealt with so far with the flow- or membrane-limited physiological models [282] that maintain compartmental and homogeneity concepts. Albeit not specifying transport limitations, the previously developed description relies on the more realistic heterogeneous conditions of drug diffusion.

7.2.2 Elimination, Liver Structure

The liver is the major site of drug biotransformation in the body [283]. It is the largest composite gland of the body and weighs about 15 g kg^{-1} body weight. The physical structure of the liver exhibits unusual microcirculatory pathways [284]. Circulation in the liver can be divided into macrocirculation and microcirculation. The former comprises the portal vein, hepatic artery, and hepatic veins, while the latter consists of hepatic arterioles and sinusoids [284]. The sinusoids are the specialized capillaries of the liver that form an uninterrupted three-dimensional network and are fully permeable by substances. This macrocirculation spans the axes of the liver while branching into successively smaller vessels. At the anatomical level, there exist small histological units, called lobules, made up of an interlacing channel network of sinusoids supplied with blood and drug by the terminal ends of the portal venules and hepatic arterioles. Between the individual sinusoids of the interior of a lobule, one-cell-thick sheets of hepatocytes are interspersed [285, 286].

7.2.2.1 In Vitro–in Vivo Correlations in Liver Metabolism

The in vitro studies in this field of research attempt to assess the rate of metabolism at an early stage of drug development in order to:

- identify problematic substances and
- allow extrapolation of the in vitro findings to in vivo conditions.

The driving force for the execution of these studies is the reduction of cost, which is related to expensive animal testing. However, replacement of in vivo testing with in vitro approaches presupposes well-based understanding of the scaling factors associating the in vitro with the in vivo measurements. The establishment of relationships between in vitro and in vivo data are known as in vitro–in vivo correlations.

Both isolated rat hepatocytes and rat liver microsomes [287–289] have been advocated for the determination of the kinetic parameters $V_{\max}$ and k_M (cf. equation 2.20) under in vitro conditions. The development of in vitro–in vivo correlations is based on two essential steps. Initially, the units of the in vitro intrinsic clearance CL_{int} ($\mu\,\mathrm{l\,min^{-1}}$ per 10^6 liver cells or $\mu\,\mathrm{l\,min^{-1}}$ per mg microsomal protein) are converted to $\mathrm{ml\,min^{-1}}$ per standard rat weight of 250 g using scaling factors reported in the literature [290]. Next, a liver model that incorporates physiological processes such as hepatic blood flow, Q, and plasma protein binding is used to provide the hepatic clearance CL_h. Therefore, the liver modeling step of the in vitro–in vivo correlations is crucial in the scaling process from the in vitro to the in vivo estimates of clearances.

Due to its mathematical simplicity, most in vitro–in vivo correlations are based on a homogeneous, "well-stirred" model for the liver such that all metabolic enzymes in the liver are exposed to the same drug concentration [291]. Under steady-state conditions, the predicted hepatic clearance CL_h for this model is

$$CL_h = \frac{Qf_u CL_{int}}{Q + f_u CL_{int}},$$

where f_u is the blood unbound fraction. Alternatively, liver has also been viewed as a parallel tube model [292]. In this case, the liver is considered as an organ receiving a series of parallel blood flows carrying the drug in identical parallel tubes representing the sinusoids. Here, the hepatic clearance assuming linear kinetics and steady-state conditions is

$$CL_h = Q\left[1 - \exp\left(-\frac{f_u CL_{int}}{Q}\right)\right].$$

However, these two models assume either perfect mixing conditions (well-stirred model) or no mixing at all (parallel tube model) and cannot explain several experimental observations. Therefore, other approaches such as the distributed model [293], the dispersion model [294], and the interconnected tubes model [295, 296] attempt to capture the heterogeneities in flow and an intermediate level of mixing or dispersion. Despite numerous comparisons [289, 290, 297–299] of the use of various liver models [291–296] for predicting the in vivo drug clearance from in vitro measurements, there is still controversy regarding the most suitable liver modeling approach. This is so since drug-specific factors, like high- or low-cleared drugs, seem to have a major impact on the quality of the in vitro–in vivo correlations. For example, low-clearance drugs are rather independent of blood-flow

characteristics, while drugs with relatively higher clearance values show a more pronounced dependence on blood-flow properties.

7.2.2.2 Fractal Considerations in Liver Metabolism

Observations of the liver reveal an anatomically unique and complicated structure, over a range of length scales, dominating the space where metabolism takes place. Consequently, the liver was considered as a fractal object by several authors [4, 273] because of its self-similar structure. In fact, Javanaud [300], using ultrasonic wave scattering, has measured the fractal dimension of the liver as approximately $d_f \approx 2$ over a wavelength domain of $0.15 - 1.5$ mm.

While there is no performance advantage over a well-stirred classical compartment, one with a rate constant due to a uniformly random distribution of drug and enzyme, such a compartment may well be impossible to achieve under biological designs, and the implied comparison is therefore an ill-posed one [301]. It may be that the fractal liver design is the best design possible, so that comparisons against nonideal theoretical models, like a poorly stirred sphere with enzymes adhered along the inner wall, are favorable. For example, the fractal structure, with many layers of membrane at its interface, allows the organ to possess a high number (concentration) of enzymes, thus giving it a high reaction rate despite time-dependent (decay) fractal kinetics. Indeed, the intricate interlacing of a stationary, catalytic phase of hepatocytes with a liquid phase of blood along a fractal border is what reduces the required diffusional distances for reactions to take place with any appreciable celerity. Moreover, the complicated structure of the liver, which provides for a huge interface between drug and hepatocytes, may be generated simply during the growth of the liver. The fractal form may be parsimoniously encoded in the DNA, indirectly specified by means of a simple recursive algorithm that instructs the biological machinery on how to construct the liver. In this way, a vascular system made up of fine tubing with an effective topological dimension of one may fill the three-dimensional embedding space of the liver. These possibilities suggest that the structure of the liver may be that of a fractal.

In this context, Berry [302] studied the enzyme reaction using Monte Carlo simulations in two-dimensional lattices with varying obstacle densities as models of biological membranes. That author found that the fractal characteristics of the kinetics are increasingly pronounced as obstacle density and initial concentration increase. In addition, the rate constant controlling the rate of the complex formation was found to be, in essence, a time-dependent coefficient since segregation effects arise due to the fractal structure of the reaction medium. In a similar vein, Fuite et al. [303] proposed that the fractal structure of the liver with attendant kinetic properties of drug elimination can explain the unusual nonlinear pharmacokinetics of mibefradil [304, 305]. These authors utilized a simple flow-limited physiologically based pharmacokinetic model where clearance of the drug occurs in the liver by fractal kinetics [303]. The analytical solution of the proposed model was fitted to experimental dog data and the estimates for the spectral dimension d_s of

the dog liver were found to be in the range $1.78 - 1.91$. This range of values is consistent with the value found in ultrasound experiments on the liver, $d_f \approx 2$ [300]. Furthermore, special attention was given to mibefradil pharmacokinetics by studying the effect of species segregation on the kinetics of the enzyme reaction in fractal media using a microscopic pharmacokinetic model mimicking the intravenous and oral administration of the substrate [306]. This mathematical model coupled with Monte Carlo simulations of the enzyme reaction in a two-dimensional square lattice reproduced the classical Michaelis–Menten kinetics in homogeneous media as well as unusual kinetics in fractal media. Based on these findings, a time-dependent version of the classic Michaelis–Menten equation was developed for the rate of change of the substrate concentration in disordered media. This equation was successfully used to describe the experimental time–concentration data of mibefradil and to derive estimates for the model parameters.

7.3 Fractal Time and Fractal Processes

The concept of fractals may be used for modeling certain aspects of dynamics, i.e., temporal evolution of spatially extended dynamic systems in nature. Such systems exhibit fractal geometry and may maintain dynamic processes on all time scales. For example, the fractal geometry of the global cloud cover pattern is associated with fluctuations of meteorological parameters on all time scales from seconds to years. Temporal fluctuations exhibit structure over multiple orders of *temporal magnitude* in the same way that fractal forms exhibit details over several orders of *spatial magnitude*. Power-law behavior has been documented in the functioning of physiological systems [307, 308]. Long-range spatial correlations have also been identified at DNA level [309, 310]. Long-range correlations over time and space for geophysical records have also been investigated by Mandelbrot and Wallis [311] and, more recently, by Tang and Bak [312]. Recent studies have identified power laws that govern epidemiological phenomena [313]. All the reported long-range temporal correlations signify persistence or *memory*.

A major feature of this correlation is that the amplitudes of short-term and long-term fluctuations are related to each other by the scale factor alone, independent of details of growth mechanisms from smaller to larger scales. The macroscopic pattern, consisting of a multitude of subunits, functions as a unified whole independent of details of dynamic processes governing its individual subunits [314]. Such a concept, whereby physical systems consisting of a large number of interacting subunits obey universal laws that are independent of the microscopic details, is acknowledged as a breakthrough in statistical physics. The variability of individual elements in a system acts cooperatively to establish regularity and stability in the system as a whole [315]. Scale invariance implies that knowledge of the properties of a model system at short times or short length scales can be used to predict the behavior of a real system at large times and large length scales [316].

The spatiotemporal evolution of dynamic systems was not investigated as a unified whole, and fractal geometry of spatial patterns and fractal fluctuations in time of dynamic processes were investigated as two separate multidisciplinary areas of research till as late as 1987. In that year, Bak et al. [317, 318] postulated that fractal geometry in spatial patterns, as well as the associated fractal fluctuations of dynamic processes in time, are signatures of self-organized phase transition in the spatiotemporal evolution of dynamic systems. The relation between spatial and temporal power-law behavior was recognized much earlier in condensed-matter physics where long-range spatiotemporal correlations appear spontaneously at the critical point for continuous phase transitions. The amplitude of large- and small-scale fluctuations are obtained from the same mathematical function using an appropriate scale factor, i.e., ratio of the scale lengths.

Conversely, the relationship (7.2) expresses a time-scale invariance (self-similarity or fractal scaling property) of the power-law function. Mathematically, it has the same structure as (1.7), defining the capacity dimension d_c of a fractal object. Thus, α is the capacity dimension of the profiles following the power-law form that obeys the fundamental property of a fractal self-similarity. A fractal decay process is therefore one for which the rate of decay decreases by some exact proportion for some chosen proportional increase in time: the self-similarity requirement is fulfilled whenever the exact proportion, α, remains unchanged, independent of the moment of the segment of the data set selected to measure the proportionality constant.

Therefore, the power-law behavior itself is a self-similar phenomenon, i.e., doubling of the time is matched by a specific fractional reduction of the function, which is independent of the chosen starting time: self-similarity, independent of scale is equivalent to a statement that the process is fractal. Although not all power-law relationships are due to fractals, the existence of such a relationship should alert the observer to seriously consider whether the system is self-similar. The dimensionless character of α is unique. It might be a reflection of the fractal nature of the body (both in terms of structure and function) and it can also be linked with "species invariance." This means that α can be found to be "similar" in various species. Moreover, α could also be thought of as the reflection of a combination of structure of the body (capillaries plus eliminating organs) and function (diffusion characteristics plus clearance concepts).

7.4 Modeling Heterogeneity

From a kinetic viewpoint, the distribution of drugs operating under homogeneous conditions can be described with classical kinetics. When distribution processes are heterogeneous, the rate constant of drug movement in the tissues is not linearly proportional to the diffusion coefficient of the drug. Then, modeling of heterogeneity features should be based on fractal kinetics concepts [4, 9, 17].

7.4.1 Fractal Concepts

A better description of transport limitations can be based on the principles of diffusion in disordered media [319]. It has been shown [320] that in disordered media the value of the first-order rate constant is related to the geometry of the medium. In these media the diffusional propagation is hindered by its geometric heterogeneity, which can be expressed in terms of fractal and spectral dimensions. For our purposes, the propagation of the drug's diffusion front in the heterogeneous space of tissues can be viewed as a diffusion process in a disordered medium. Both the diffusion coefficient of the drug and the rate constant are dependent on the position of the radial coordinate of the diffusion front, and therefore both parameters are time-dependent. In these lower-dimensional systems, diffusion is inhibited because molecules cannot move in all directions and are constrained to locally available sites.

The description of these phenomena in complex media can be performed by means of fractal geometry, using the spectral dimension d_s. To express the kinetic behavior in a fractal object, the diffusion on a microscopic scale of an exploration volume is analyzed [303]. A random walker (drug molecule), migrating within the fractal, will visit $n(t)$ distinct sites in time t proportional to the number of random-walk steps. According to the relation (2.9), $n(t)$ is proportional to $t^{d_s/2}$, so that diffusion is related to the spectral dimension.

The case $d_s = 2$ is found to be a critical dimension value in the phenomena of self-organization of the reactants:

- For $d_s > 2$, a random walker has a finite escape probability-microscopic behavior conducive to re-randomize the distribution of reactants around a trap and deplete the supply of reactive pairs, and thus a stable macroscopic reactivity as attested by the classical rate constant [321, 322]. The scale of the self-organization is microscopic and independent of time, such that $n(t) \propto t$ (is linear) and $k = \dot{n}(t)$ is a constant, so the reaction kinetics are classical.
- For $d_s \leq 2$, a random walker (drug) is likely to stay at its original vicinity and will eventually recross its starting point, a microscopic behavior conducive to producing mesoscopic depletion zones around traps, e.g., enzymes. The compactness of the low-dimensional random walk implies ineffective diffusion, relevant mesoscopic density fluctuations of the drug, and an entailing aberrant macroscopic rate coefficient. Subsequently, the macroscopic reaction rate, which is given by the time derivative of $n(t)$, sometimes described as the efficiency of the diffusing, reacting random walker, will be

$$k(t) \propto \dot{n}(t) \propto t^{-(1-d_s/2)} = t^{-\lambda} \tag{7.4}$$

for transient reactions [303]. Since $0 < d_s \leq 2$, the parameter λ has values in the range $0 \leq \lambda < 1$. The minus sign in (7.4) is used to mimic the decrease of k with time as the walker (drug) has progressively less successful

visits. This time-dependent rate "constant" in the form of a power law is the manifestation of the anomalous microscopic diffusion in a dimensionally restricted environment leading to anomalous macroscopic kinetics [303].

The kinetic consequences that are associated with the time dependency of the rate "constant" are delineated in Section 2.5 under the heading, coined by Kopelman [9, 17], *fractal-like kinetics*.

7.4.2 Empirical Concepts

Heterogeneity could also be expressed and described by elementary operations with empirical models. The only difference between (7.1) and (7.2) lies in the coefficient of $c(t)$ on the right-hand side of the differential equations. This allows someone to infer empirically that these equations could be unified as

$$\dot{c}(t) = -\beta \left(\frac{\beta t}{\alpha}\right)^{-\lambda} c(t) \tag{7.5}$$

with initial condition $c(t_0) = c_0$ at $t_0 \neq 0$. The exponent λ takes integer 0 or 1 values corresponding to the exponential and power-law profiles, respectively, and α and β are as defined in (7.1) and (7.2). Since the gamma profile (7.3) is presented as the additive mixture of the previous ones, one wonders whether λ is allowed to attain fractional values between 0 and 1. Indeed, the previous equation could also be considered as a generalization of (7.1) and (7.2) assuming a fractional time exponent λ ($0 \leq \lambda \leq 1$). Under this assumption, (7.5) is similar to what we reported previously (equations 5 and 7 in [281]), obtained from the classical first-order rate kinetics assuming that the rate coefficient is a time-varying rate coefficient.

The solution of (7.5) is

$$c(t) = c_0 \exp\left\{-\frac{\alpha}{1-\lambda}\left[\left(\frac{\beta t}{\alpha}\right)^{1-\lambda} - \left(\frac{\beta t_0}{\alpha}\right)^{1-\lambda}\right]\right\} \tag{7.6}$$

for $\lambda \neq 1$ and

$$c(t) = c_0 \exp\left[-\alpha \ln\left(\frac{t}{t_0}\right)\right]$$

for $\lambda = 1$. Then, with fractional λ, the transition in output response is continuous between a homogeneous process ($\lambda = 0$) and a heterogeneous one ($\lambda \neq 1$) (or equivalently, how to generate multiexponential behavior starting from a monoexponential one). Inversely, after fitting observed data by empirical models such as (7.6), the estimated value of λ might help us classify drugs in two large groups:

- *Homogeneous drugs* with $\lambda \approx 0$: their kinetics can be described homogeneously with what we will call compartmental models. These drugs are characterized by small or medium volumes of distribution.

- *Heterogeneous drugs* with $\lambda \neq 0$: their kinetics are described with non-compartmental modeling, and in reality they approximate the true heterogeneous disposition, i.e., the time-dependent character of diffusion (flow). These drugs are characterized by high volumes of distribution.

Moreover, combinations of these models can also be used to roughly describe physiological considerations. For instance, if the drug is metabolized by the liver and simultaneously eliminated by the kidney, a gamma profile is obtained as the solution of (7.3), where the α/t term expresses the structural heterogeneity of the liver, and the term β, the homogeneous elimination process from the kidney.

7.5 Heterogeneity and Time Dependence

It has been stated that heterogeneous reactions taking place at interfaces, membrane boundaries, or within a complex medium like a fractal, when reactants are spatially constrained on a microscopic level, culminate in deviant reaction rate coefficients that appear to have a sort of *temporal memory*. Fractal kinetic theory suggested the adoption of a time-dependent rate "constant," with power-law form, determined by the spectral dimension. This time dependency could also be revealed from empirical models.

In fact, empirical models involve parameters without any physiological meaning. To obtain sound biological information from the observed data, these models should be converted to some more phenomenological ones, parameterized by volume of distribution, clearance, elimination rate constant, etc. In their simplest form, phenomenological models are based on Fick's first law (2.14), where the concentration gradient is the force acting to diffuse the material q through a membrane:

$$\dot{q}(t) = -CLc(t), \tag{7.7}$$

where CL is the clearance. Concentration and amount of material are also linked via the well-known relationship

$$q(t) = Vc(t), \tag{7.8}$$

where V is the volume of distribution of the material. We also explicitly denote the time dependency in each parameter, $CL(t)$ and $V(t)$, and define the rate constant $k(t)$ as

$$\frac{CL(t)}{V(t)} \triangleq k(t). \tag{7.9}$$

Differentiating (7.8) with respect to t and using expressions (7.7) and (7.9) to substitute $\dot{q}(t)$ and $CL(t)$, we obtain

$$c(t)\dot{V}(t) = -k(t)c(t)V(t) - V(t)\dot{c}(t). \tag{7.10}$$

According to the exponential, power-law, or gamma empirical model, $\dot{c}(t)$ may take the form of relation (7.1), (7.2), or (7.3), respectively. By introducing these relations in (7.10) we get, respectively,

$$\dot{V}(t) = [\beta - k(t)]V(t) \tag{7.11}$$

or

$$\dot{V}(t) = \left[\frac{\alpha}{t} - k(t)\right]V(t) \tag{7.12}$$

or

$$\dot{V}(t) = \left[\frac{\alpha}{t} + \beta - k(t)\right]V(t). \tag{7.13}$$

A time-invariant process has time-independent parameters. Therefore, a time-invariant process is that for which both V and k are invariant in time. From the three previous relationships, the only time-independent situation occurs in the exponential empirical model when $k(t) = \beta$. In this case, from (7.11) one has $V(t) = V_0$, a time-invariant volume. The processes fitted by the power-law and gamma empirical models are necessarily time-varying processes, because when either V or k is kept constant, the other becomes time-varying.

In these cases, two extreme situations may occur:

- k is time-invariant. If we assume $k(t) = \beta$ in (7.12) and (7.13), the time courses of the volume are

  $$V(t) = V_0 t^{\alpha}\exp(-\beta t) \qquad \text{and} \qquad V(t) = V_0 t^{\alpha},$$

  respectively, where V_0 is set according to the initial conditions. Taking into account this time dependence of volume, a unique form of the amount profile is obtained, $q(t) = Q_0\exp(-\beta t)$, irrespective of the exponential, power-law, or gamma concentration profiles.
- V is time-invariant. From (7.12) and (7.13) one obtains the time course of the rate constant:

  $$k(t) = \frac{\alpha}{t} \qquad \text{and} \qquad k(t) = \frac{\alpha}{t} + \beta,$$

  respectively. With time-invariant V, the amount profiles $q(t)$ will be proportional to the concentration profiles $c(t)$.

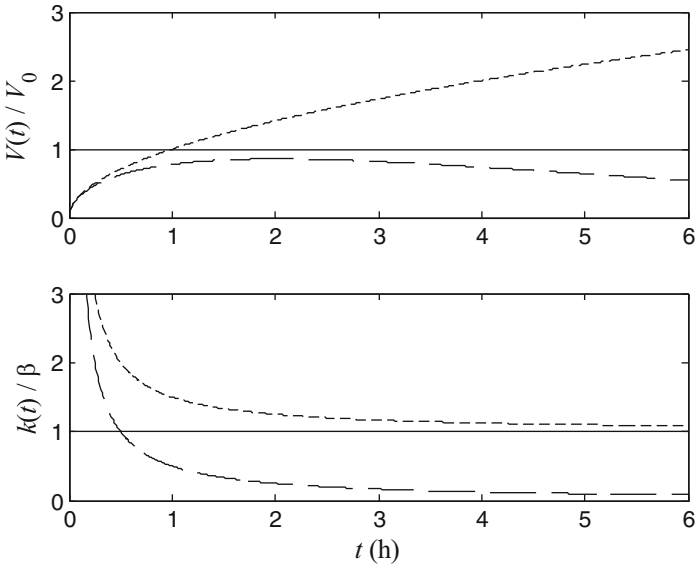


Fig. 7.3 Time courses of $V(t)/V_0$ (*up*) and of $k(t)/\beta$ (*down*) associated with the exponential, power-law, and gamma empirical models (*solid*, *dashed*, and *dotted lines*, respectively)

Figure 7.3 illustrates the time courses of the reduced volume of distribution $V(t)/V_0$ and of the reduced rate constant $k(t)/\beta$ with $\alpha = 0.5$ and $\beta = 0.25$. Certainly, mixed situations where both $k(t)$ and $V(t)$ are time-varying can be thought of.

This preliminary analysis highlights the difference between regular and irregular profiles associated with time-invariant and time-varying physiological parameters, respectively. Some authors have attempted to associate a functional physiological meaning to the gamma empirical model [323, 324] or to describe by stochastic modeling the real processes leading to power-law outputs [325, 326]. In contrast, in the case of calcium pharmacokinetics [281], the possible mechanisms underlying (7.3), where renal elimination of calcium was associated with the parameter β, and the other elimination mechanisms, with parameter α were discussed. Lastly, a simple approach for including, within a multicompartment model, time dependence of the transfer coefficients that vary continuously with the age of human patients was described by Eckerman et al. [327], but time dependence was over periods much greater than a single dose. This simplified the mathematics so that there was no time dependence of coefficients during the time course of a single dose. Within a physiological model, over a very long time scale of 98 days, Farris et al. [328] introduce time-dependent compartment volume changes due to growth in the studied rat model system.

Therefore, it is clear that when the outputs are optimally fitted by the power-law and gamma empirical models, the underlying processes are rather time-varying. The time-varying features of the observed processes are in fact the expression of functional or structural heterogeneities in the body.

7.6 Simulation with Empirical Models

The observed empirical models should now be employed to simulate and predict kinetic behaviors obtained with administration protocols other than those used for observation. Moreover, we must develop pharmacokinetics in a multicompartment system by including the presence of a fractal organ. We have argued that the liver, where most of the enzymatic processes of drug elimination take place, has a fractal structure. Hence, we expect transport processes as well as chemical reactions taking place in the liver to carry a signature of its fractality.

Little has been done so far to predict the effect of different modes of administration, in inhomogeneous conditions, on the observed $c(t)$ when this contains a power function. In fact, the availability of the drug in the process was simply expressed by an initial condition $c(t_0) = c_0$. Later on, exponential, power-law, or gamma profiles were observed according to the inherent heterogeneity of the process.

Empirical models helped us recognize heterogeneity in the process and express it simply by mathematical models with time-varying parameters V and k. Nevertheless, time in such time-varying parameters can be conceived only as a maturation time or as an age a associated with each administered molecule, i.e., $V(a)$ and $k(a)$. This time a must be distinguished from the exogenous time t associated with the evolution of the overall process. Several hypotheses based on fractal principles were formulated to explain heterogeneity and time dependency, but conceptual difficulties persist in explaining the time profiles of $V(a)$ and $k(a)$. Volume may represent the maximal space visited by a molecule and the elimination constant, the fragility of a molecule while it remains in the process. These parameters are dependent on the age a of each molecule, and they must be independent of the drug administration protocol, e.g., the repeated dosages, which are scheduled with respect to the exogenous time t of the process. Therefore, the relation between a and t must be resolved before integrating in the model the usual routes of administration. The heterogeneous process observed in several circumstances and the resulting complexity of the molecular kinetic behaviors, with respect to the actual experiments, required new techniques as well as modifications of Fick's law in order to comply with observations. In this way, two operational procedures may be retained:

- First operate at a molecular level and establish a probabilistic model for the behavior and the time spent by each molecule in the process. Second, take into account statistically all the molecules in the process. This stochastic formulation would be the most appropriate for capturing the structural and functional heterogeneity in the biological media. The resulting models supply tractable forms involving the time-varying parameters $V(a)$ and $k(a)$ [329]. This issue was greatly addressed in biological systems and only recently in pharmacokinetics [330, 331]. It will be developed, here, in Chapter 11.
- From a holistic point of view, the time-varying parameters $V(t)$ and $k(t)$ fitting the observed data could represent the dynamic behavior of a complex system involving feedback mechanisms implying the states $q(t)$. So, these parameters

can be assumed to be complex functions of $q(t)$, namely $V(q)$ and $k(q)$, leading to nonlinear kinetics (e.g., logistic saturable [332]), with time-varying coefficients [281], etc. For decades, this approach has had numerous applications in pharmacokinetics, and it allows any complex function to be assumed as V and k. Time variation in the parameters is treated in the next section.

7.7 Time-Varying Models

The fundamental working hypothesis is that time-varying parameters are expressions of feedback regulation mechanisms involving the process states. To reveal the dependence of the time-varying parameters on the process states, we propose the following procedure:

1. Roughly describe the process by means of a phenomenological continuous-time model according to the underlying physiological structure. The reason is that most physical laws are expressed in continuous time as differential equations. For instance, use compartmental configuration to sketch the fundamental mechanisms. This results in a linear state-space dynamic model described by a set of differential equations continuous in time

$$\dot{\underline{y}}(t) = \mathbf{A}\underline{y}(t) + \underline{b}u(t), \tag{7.14}$$

where $\underline{y}(t)$ and $u(t)$ are the states and inputs, respectively. The matrix $\mathbf{A}$ and the vector $\underline{b}$ involve the parameters $\underline{x}$ of this holistic description, e.g., exchange rates and volumes of distribution. The parameters $\underline{x}$, and therefore $\mathbf{A}$ and $\underline{b}$, may be constant or time varying.
2. Given a set of observed data, reveal the time dependency of $\underline{x}$. If x is demonstrated to be time-varying $\underline{x}(t)$, the issue is to obtain the time profile of $\mathbf{A}(t)$ and $\underline{b}(t)$ in (7.14). To perform this key operation, it is necessary to estimate the model "on-line" at the same time as the input–output data are received [333]. Identification techniques that comply with this context are called *recursive identification* methods. Other commonly used terms for such techniques are on-line or real-time identification, or sequential parameter estimation [334]. Using these techniques, it may be possible to investigate time variations in the process in a real-time context. However, tools for recursive identification are available for *discrete-time models*. Most common discrete-time models are difference-equation descriptions, such as the Auto-Regression with eXtra inputs (ARX) model. The basic relationship is the linear difference equation

$$y(t) + a_1 y(t-1) + \ldots + a_{n_a} y(t-n_a) = b_1 u(t-1) + \ldots + b_{n_b} u(t-n_b),$$

which relates the current output $y(t)$ to a finite number of past outputs $y(t-i)$ and inputs $u(t-i)$, with $i = 1, \ldots, (n_a, n_b)$. State-space and ARX models describe the functional relation between inputs and outputs. The order of the

state-space model relates to the number n_a and n_b of delayed outputs and inputs, respectively, used in the corresponding difference ARX model.

3. Analyze the time profile of $\mathbf{A}(t)$ and $\underline{b}(t)$ against the states $\underline{y}(t)$. For instance, one looks at the dependence of $a_k(t)$ on $y_j(t)$ by plotting $a_k(t)$ or $\log a_k(t)$ as a function of $y_j(t)$ or $\log y_j(t)$. This dependence can be expressed by a second-level model of the form $\mathbf{A}\left[\underline{y}(t)\right]$ and $\underline{b}\left[\underline{y}(t)\right]$ resulting in the nonlinear differential equation

$$\dot{\underline{y}}(t) = \mathbf{A}\left[\underline{y}(t)\right]\underline{y}(t) + \underline{b}\left[\underline{y}(t)\right]u(t). \tag{7.15}$$

This second-level model of the feedback mechanisms involving the states leads to nonlinear models for processes, which under some experimental conditions may exhibit chaotic behavior.

The transformation procedure of a time-varying parameter model to a nonlinear one has already been applied in other contexts. For instance in the simple case

$$\dot{y}(t) = -x(t)\,y(t), \tag{7.16}$$

if it is possible to approximate $\log x(t)$ linearly at any logarithmically transformed state $\log y(t)$, one obtains $\log x(t) = \lambda + \mu \log y(t)$. In terms of the original variables, this results in a power-law approximation

$$x(t) = \lambda y^{\mu}(t).$$

Subsequently, the differential equation with time-varying parameters (7.16) is transformed into a differential equation of the form

$$\dot{y}(t) = -\lambda y^{\mu+1}(t).$$

Another example is the diffusion-limited or dimensionally restricted homodimeric reaction presented in Section 2.5.3. Equation 2.23 is the traditional rate law with concentration squared and time-varying time "constant" $k(t)$, whereas (2.22) is the power law ($c^{\gamma}(t)$) in the state differential equation with constant rate.

In the presence of multiple states, the right-hand-side term consists of sums, products, and nesting of elementary functions such as y^{μ}, $\log y$, $\exp y$, and trigonometric functions, called the *S*-system formalism [335]. Using it as a canonical form, special numerical methods were developed to integrate such systems [336].

In the case where the input $u(t)$ is piecewise constant over time intervals (this condition is fulfilled in our context), then the conversion of (7.14) to a discrete-time model is

- possible without any approximation or additional hypothesis, if $\underline{x}$ is revealed not to be time-dependent;
- a very difficult task, if $\underline{x}$ is revealed to be time varying.

The following sections present the conversion of continuous- to discrete-time linear time-invariant (LTI) models. Finally, the recursive identification is presented for a model with time-varying parameters.

7.7.1 Discrete- vs. Continuous-Time LTI Models

To emphasize the continuous-time character in equation 7.14, $\mathbf{A}_C$ and $\underline{b}_C$ were introduced instead of $\mathbf{A}$ and $\underline{b}$, respectively. The equation is re-written

$$\dot{\underline{y}}(t) = \mathbf{A}_C \underline{y}(t) + \underline{b}_C u(t). \tag{7.17}$$

We assume state $\underline{y}(t)$ to be m-dimensional, so that A_C and $\underline{b}_C$ are $m \times m$ matrix and $m \times 1$ vector, respectively. This form can be converted in the following discrete-time form:

$$\underline{y}(t_{k+1}) = A_D \underline{y}(t_k) + \underline{b}_D u(t_k)$$

where the states $\underline{y}(t)$ are assumed to be evaluated or observed at the sampling times (not the biological sampling but the discretization sampling) $t_k = kT$ with $k = 1, 2, \ldots$. Also, T is the sampling interval, a sufficiently short time to obtain an adequate representation of the kinetic profile. Finally, the input $u(t)$ is assumed to be constant between two successive sampling times

$$u(k) = u(t) \quad kT \leq t < (k+1)T.$$

The A_D and $\underline{b}_D$ elements in the previous discrete version depend on the sampling interval T and on the parameters $\underline{x}$. They are given by the following relationships:

$$A_D = \exp(A_C T) \tag{7.18}$$

and

$$\underline{b}_D = \int_0^T \exp(A_C t)\, \underline{b}_C dt.$$

By using the derivation rule

$$\frac{d \exp(Xt)}{dt} = X \exp(Xt)$$

of a matrix exponentiation and, because $\underline{b}_C$ is not time dependent, the following holds:

$$\underline{b}_D = A_C^{-1}\left[\int_0^T d\exp\left(A_C t\right)\right]\underline{b}_C = A_C^{-1}\left[\exp\left(A_C T\right) - I\right]\underline{b}_C. \tag{7.19}$$

7.7.1.1 Closed-Form Solution of Discrete-Time Models

The commonly used expression of the discrete-time model is

$$\underline{y}\left(k+1\right) = A_D\underline{y}\left(k\right) + \underline{b}_D u\left(k\right) \tag{7.20}$$

where the sampling number k was only reported instead of the explicit notation $t_k = kT$. Briefly, the states $\underline{y}\left(k+1\right)$ at the time $k+1$ are expressed as a linear combination of the states $\underline{y}\left(k\right)$ and inputs $u\left(k\right)$ at the previous time k.

Nevertheless, the above form is useless because the states are implied in the same relationship at different sampling times kT and $kT+T$. The use of the *shift operator* q, defined by $qu\left(t\right) \triangleq u\left(t+T\right)$ or $q^{-1}u\left(t\right) \triangleq u\left(t-T\right)$, converts the set 7.20 of discrete-time equations to the following more relevant form:

$$\left[qI - A_D\right]\underline{y}\left(k\right) = \underline{b}_D u\left(k\right).$$

The ratio

$$\underline{\underline{\mathcal{F}}}\left(q,\underline{x}\right) \triangleq \frac{\underline{y}\left(k\right)}{u\left(k\right)} = \left[qI - A_D\right]^{-1}\underline{b}_D,$$

involves the shift operator q and it is named *discrete-time transfer function* of a multi-output model. The above relationship represents a rational function with polynomials in q for the numerator and denominator.

7.7.1.2 Discrete-Time Transfer Function

The elements A_D and $\underline{b}_D$ in the discrete-time transfer function are given by relationships 7.18 and 7.19, respectively. So

$$\underline{\underline{\mathcal{F}}}\left(q,\underline{x}\right) = \left[qI - \exp\left(A_C T\right)\right]^{-1} A_C^{-1}\left[\exp\left(A_C T\right) - I\right]\underline{b}_C. \tag{7.21}$$

In order to develop calculus involved in the above, the factorization of the $\exp\left(A_C T\right)$ is introduced. According to the second proposition in the Appendix G, the *Jordan factorization* involves the matrix G of eigenvectors $\underline{\eta}^{(i)}$ of A_C and the diagonal matrix Z of eigenvalues ζ_i of A_C

$$\exp\left(A_C T\right) = G\exp\left(ZT\right)G^{-1} = G\,diag\left(\exp\left(\zeta_i T\right)\right)G^{-1}$$

Here, $\exp\left(ZT\right) = diag\left(\exp\left(\zeta_i T\right)\right)$ because Z is a diagonal matrix, $Z = diag\left(\zeta_i\right)$.

By using the Jordan factorization and the *matrix inversion lemma* (cf. Appendix G), the pre- and post-multiplying factors involved in 7.21 become

$$\begin{aligned}[qI - \exp(A_C T)]^{-1} &= \left[qI - Gdiag\left(\exp(\zeta_i T)\right) G^{-1}\right]^{-1} \\ &= q^{-1}\left[I - Gdiag\left[\frac{1}{1 - q\exp(-\zeta_i T)}\right] G^{-1}\right]\end{aligned}$$

and

$$A_C^{-1}\left[\exp(A_C T) - I\right]\underline{b}_C = GZ^{-1}\left[diag\left(\exp(\zeta_i T)\right) - I\right] G^{-1}\underline{b}_C,$$

respectively. Given the above relationships, the transfer function 7.21 becomes

$$\underline{\mathcal{F}}(q, \underline{x}) = Gdiag(\sigma_i) G^{-1}\underline{b}_C \tag{7.22}$$

with

$$\sigma_i = \frac{\beta_i q^{-1}}{1 + \alpha_i q^{-1}}.$$

Here,

$$\alpha_i = -\exp(\zeta_i T) \qquad \text{and} \qquad \beta_i = -\frac{1 - \exp(\zeta_i T)}{\zeta_i}$$

are dummy variables with $i = 1 : m$. Therefore, the final form of discrete-time transfer function involves $\underline{b}_C$ and, the eigenvectors and eigenvalues of A_C. Since A_C and $\underline{b}_C$ are defined on the basis of $\underline{x}$, the transfer function depends on the shift operator q and the parameters $\underline{x}$.

7.7.2 Polynomial Form of Transfer Function

For a single-output discrete-time model, the transfer function 7.22 can be expressed as

$$\mathcal{F}(q, \underline{x}) \triangleq \frac{y(k)}{u(k)} = \frac{B(q, \underline{x})}{A(q, \underline{x})} \tag{7.23}$$

where $A(q, \underline{x})$ and $B_j(q, \underline{x})$ are polynomials depending on the parameters $\underline{x}$, and k is the sampling number. Parametrization of the polynomials could be obtained in several ways. The most popular is the ARX model whose parameters are the coefficients of polynomial terms in the numerator

$$B(q, \underline{x}) = b_1 q^{-1} + \ldots + b_{n_b} q^{-n_b} \tag{7.24}$$

and denominator

$$A\left(q,\underline{x}\right)=1+a_1q^{-1}+\ldots+a_{n_a}q^{-n_a}. \tag{7.25}$$

The order n_a of polynomial $A\left(q,\underline{x}\right)$ is equal to the number of first-order differential equations in the initial description 7.14. For the order n_b of $B\left(q,\underline{x}\right)$, $n_b \leq n_a$ holds in general. In the above expressions, the polynomial coefficients a_i and b_i are the new parametrization of the discrete-time model. These coefficients can be aggregated as follows:

$$\underline{\theta}=\left[\begin{array}{cccc}a_1 & \ldots & a_{n_a} & b_1 \ \ldots \ b_{n_b}\end{array}\right]^T.$$

Since polynomials 7.25 and 7.24 depend on $\underline{x}$, the new polynomial parameters $\underline{\theta}$ also depend on $\underline{x}$. For LTI models, conversion formulas allow computation of $\underline{\theta}$ given $\underline{x}$ or, inversely compute $\underline{x}$ given $\underline{\theta}$. From the relationships 7.23, 7.24, and 7.25, and after shifting by one the sampling interval

$$y\left(k\right)+a_1y\left(k-1\right)+\ldots+a_{n_a}y\left(k-n_a\right)=b_1u\left(k-1\right)+\ldots+b_{n_b}u\left(k-n_b\right). \tag{7.26}$$

This discrete expression of the model is the famous ARX form and that will be further used. The form advocates the designation "Auto-Regression" because the prediction $y\left(k\right)$ is composed of the previous "predictions" $y\left(k-i\right)$.

By introducing the *regression vector*

$$\underline{\varphi}\left(k\right)=\left[-y\left(k-1\right)\ \ldots\ -y\left(k-n_a\right)\ \ u\left(k-1\right)\ \ldots\ u\left(k-n_b\right)\right]^T, \tag{7.27}$$

the above relationship can be re-written as $y\left(k\right)=\underline{\varphi}^T\left(k\right)\underline{\theta}$. Again, to emphasize that the calculation of $y\left(k\right)$ depends on the parameters $\underline{\theta}$, we will rather call this calculated value

$$y\left(k|\underline{\theta}\right)=\underline{\varphi}^T\left(k\right)\underline{\theta}. \tag{7.28}$$

Therefore, concentrations $y\left(k|\underline{\theta}\right)$ are linear with respect both to the inputs $u\left(k\right)$ involved in the regression vector $\underline{\varphi}(k)$ and to the parameters $\underline{\theta}$.

7.7.2.1 Parameter Estimation

Now suppose that we do not know the values of parameters in $\underline{\theta}$, but that we have recorded inputs $u\left(k\right)$ and measured outputs $\widehat{y}\left(k\right)$ over m samples ($1 \leq k \leq m$). An obvious approach is to select $\widehat{\underline{\theta}}(m)$ in 7.28 so as to fit as well as possible the calculated values $y\left(k|\underline{\theta}\right)$ to the m measured outputs $\widehat{y}\left(k\right)$ by the weighted least squares method

$$\widehat{\underline{\theta}}\left(m\right)=\arg\min\frac{1}{m}\sum_{k=1}^{m}w\left(m,k\right)\left[\widehat{y}\left(k\right)-y\left(k|\underline{\theta}\right)\right]^2.$$

The factor $w(m,k)$ in the above expression is a weighting function. A pragmatic and useful way to use 7.28 is to view it as a way of determining the next *predicted* outputs given the previous *observations*. Therefore, the definition of regression vector 7.27 initially involving *predictions* now makes use of *observations*. So, for estimation purposes, the observed regression vector

$$\underline{\widehat{\varphi}}(k) = [-\widehat{y}(k-1) \ \ldots \ -\widehat{y}(k-n_a) \ \ u(k-1) \ \ldots \ u(k-n_b)]^T$$

will be used instead of the predicted regression vector 7.27. The estimator will now work as

$$\underline{\widehat{\theta}}(m) = \arg\min \frac{1}{m} \sum_{k=1}^{m} w(m,k) \left[\widehat{y}(k) - \underline{\widehat{\varphi}}^T(k)\, \underline{\theta}\right]^2. \tag{7.29}$$

In these developments, the measurement error in outputs was assumed negligible.

Off-Line Estimation Since predictions are linear with respect to parameters $\underline{\theta}$, the sum of weighted squared residuals is a quadratic form. Its minimum value can easily be obtained by setting at zero the derivative of the sum with respect to parameters $\underline{\theta}$. The expression for the resulting estimate is

$$\underline{\widehat{\theta}}(m) = Q^{-1}(m)\, \underline{f}(m) \tag{7.30}$$

with

$$Q(m) = \sum_{k=1}^{m} w(m,k)\, \underline{\widehat{\varphi}}(k)\, \underline{\widehat{\varphi}}^T(k) \quad \underline{f}(m) = \sum_{k=1}^{m} w(m,k)\, \underline{\widehat{\varphi}}(k)\, \widehat{y}(k). \tag{7.31}$$

According to the previous discussion, observations $\widehat{y}(k)$ are involved in $\underline{\widehat{\varphi}}(k)$.

Consequently, the description of the real process by means of a linear discrete-time model conveniently uses a single step estimation of parameters $\underline{\theta}$ and therefore, the iterative algorithm that is needed for estimation of parameters $\underline{x}$ involved in the nonlinear continuous-time model is circumvented. The inverse conversion formulas allow computation of $\underline{x}$ given $\underline{\theta}$ for the LTI models.

On-Line Estimation The model should be based on observations up to the current time. The on-line computation of the model must also be done in such a way that the processing of the measurements from one sample can be completed during one sampling interval. Such a recursive algorithm can be derived from an off-line version using the philosophy of performing one iteration in the numerical search at the same time as a new observation is included in the criterion. In this sense, apart from possible initial-value effects, recursive estimates coincide with its off-line version.

In the weighted least-squares criterion, relationships 7.30 and 7.31 hold for any k among the m sampling times. For the recursive algorithm, we assume the weighting sequence has the property

$$w(k,i) = \lambda(k)\, w(k-1,i) \quad 0 \le i \le k-1 \quad w(k,k) = 1. \tag{7.32}$$

This implies recursive relationships

$$Q(k) = \lambda(k)\, Q(k-1) + \underline{\widehat{\varphi}}(k)\, \underline{\widehat{\varphi}}^T(k)\, f(k) = \lambda(k) f(k-1) + \underline{\widehat{\varphi}}(k)\, \widehat{y}(k)$$

corresponding to the relationships 7.31. Also, $\underline{\widehat{\theta}}(k) = Q^{-1}(k) f(k)$ and taking into account the above relationships

$$\underline{\widehat{\theta}}(k) = \underline{\widehat{\theta}}(k-1) + Q^{-1}(k)\, \underline{\widehat{\varphi}}(k) \left[\widehat{y}(k) - \underline{\widehat{\varphi}}^T(k)\, \underline{\widehat{\theta}}(k-1) \right].$$

To avoid inverting $Q(k)$ at each step, it is convenient to introduce $P(k) = Q^{-1}(k)$. By applying the matrix inversion lemma, the algorithm is

$$\begin{aligned} z(k) &= \lambda(k) + \underline{\widehat{\varphi}}^T(k)\, P(k-1)\, \underline{\widehat{\varphi}}(k) \\ P(k) &= \left[P(k-1) - P(k-1)\, \underline{\widehat{\varphi}}(k)\, \underline{\widehat{\varphi}}^T(k)\, P(k-1)/z(k) \right] / \lambda(k) \\ L(k) &= P(k-1)\, \underline{\widehat{\varphi}}(k)/z(k) \\ \underline{\widehat{\theta}}(k) &= \underline{\widehat{\theta}}(k-1) + L(k) \left[\widehat{y}(k) - \underline{\widehat{\varphi}}^T(k)\, \underline{\widehat{\theta}}(k-1) \right] \end{aligned} \tag{7.33}$$

To use the recursive algorithms, initial values for their start-up are required. A possibility is to start recurrence after the time instant t_0 when $Q(t_0)$ has become invertible (typically $\frac{t_0}{T} > \dim(\theta)$). So starting conditions will be $Q(t_0), f(t_0)$, and $\underline{\widehat{\theta}}(t_0)$ in accordance with the off-line relationships 7.30 and 7.31.

A reason for using recursive identification in practice is that the properties of the system may be time-varying and that we want the identification algorithm to track the variations. This is handled in a natural way in the weighting criterion 7.29 by assuming less weight to older measurements that are no longer representative for the system. In terms of 7.32, this means that we choose $\lambda(i) < 1$. In particular for a system that changes gradually, the most common choice is $\lambda(i) \equiv \lambda$, then $w(k,i) = \lambda^{k-i}$ and old measurements in the criterion are exponentially discounted. In that case, λ is often called the *forgetting factor*.

The forgetting factor is chosen slightly less than 1 so that

$$w(k,i) = \exp\left[(k-i)\ln(\lambda)\right] \approx \exp\left[-(k-i)(1-\lambda)\right].$$

This means that measurements that are older than $T_0 = 1/(1-\lambda)$ samples are included in the criterion with a weight that is $\exp(-1) \approx 36\%$ of that of the most recent measurements. T_0 could be called the memory time constant of the criterion.

If the system remains approximately constant over T_0 samples, a suitable choice of λ can be made from the previous relationship, i.e., $\lambda = 1 - 1/T_0$. Typical choices of λ are in the range between 0.96 and 0.995.

Conversely, if 64% of sampling times and measurements should be supported over a time horizon $h_0 = TT_0$, the forgetting factor is $\lambda = 1 - T/h_0$. If $\lambda = 0.99$ is set, the sampling interval T must be the hundredth of horizon h_0.

It is worth mentioning that the conversion formulas between $\underline{x}$ (continuous time) and $\underline{\theta}$ (discrete time) based on the relationships 7.18 and 7.19 do not apply for the on-line estimation because these relationships assume LTI models. For applications with recursive identification, the starting model is the discrete-time model 7.26 and the time profiles of its parameters $\underline{\theta}$ are recursively obtained given the data. At each sampling time k, the $\underline{\theta}(k)$ are (1°) converted to parameters $A_D(k)$ and $\underline{b}_D(k)$ involved in the discrete-time form (equation 7.20) and (2°) $A_D(k)$ and $\underline{b}_D(k)$ are converted to parameters $A_C(t)$ and $\underline{b}_C(t)$ by fitting the frequency responses of the discrete- and continuous-time models and applying model reduction techniques [337].

7.7.3 Pharmacokinetic Application

Tracking the time-varying parameters for the one-compartment model associated with intravenous route of administration is considered here. To unveil the dependence of the time-varying parameters on the states, the proposed procedure is used:

1. Roughly describe the process. To describe the time profiles of amounts of drug in the compartments, the mass conservation law was applied for the compartment. The elementary model of first-order process was used to describe the drug elimination and a differential equation was obtained. The amount of drug implied in this equation was converted in concentration by dividing amount by volume of distribution. The so-obtained first-order differential equation is

$$\frac{dy_1(t)}{dt} = -k_{10}y_1(t) + \frac{u(t)}{V_1} \qquad \text{and} \qquad y_1(0) = 0. \tag{7.34}$$

The state variable $y_1(t)$ represents the concentrations of the drug in the compartment and the input function $u(t)$ describes the drug administration protocol. The parameters $\underline{x}$ involved in the model are the elimination rate constant k_{10} and the volume of distribution V_1 of compartment. For the intravenous infusion of total dose D over an infusion time τ, the input function is given by

$$u(t) = \frac{D}{\tau}[H(t) - H(t-\tau)]$$

where $H(t)$ is the step or Heaviside function. For the purposes of this simulated application, the elimination rate constant k_{10} is assumed to be dependent on the concentration $y_1(t)$ according to the Michaelis–Menten type relationship

$$k_{10} \sim k_{10}(t) = V_m / [K_m + y_1(t)] .$$

Therefore, $k_{10}(t)$ becomes a time-varying parameter in this application.

2. Reveal the time dependency of parameters. The observed data were generated by using the above Michaelis–Menten relationship with

$$\underline{x} = \left[V_1 \; V_m \; K_m \right]^T = \left[20 \; 1 \; 0.5 \; \mathrm{h}^{-1} \; 1 \; \mathrm{mg/l} \right]^T$$

and administration protocol with $D = 200$ mg and $\tau = 1$ min. The obtained reference profile is illustrated in Figure 7.4 by the solid line. The dashed line represents the time profile of $k_{10}(t)$. Since the model is of first order, $n_a = n_b = 1$, and the polynomial form of the transfer function 7.23 is

$$\mathcal{F}(q, \underline{x}) = \frac{b_1 q^{-1}}{1 + a_1 q^{-1}}$$

and the ARX model 7.26 is

$$y(k) + a_1 y(k-1) + \ldots + a_{n_a} y(k - n_a) = b_1 u(k-1) + \ldots + b_{n_b} u(k - n_b) .$$

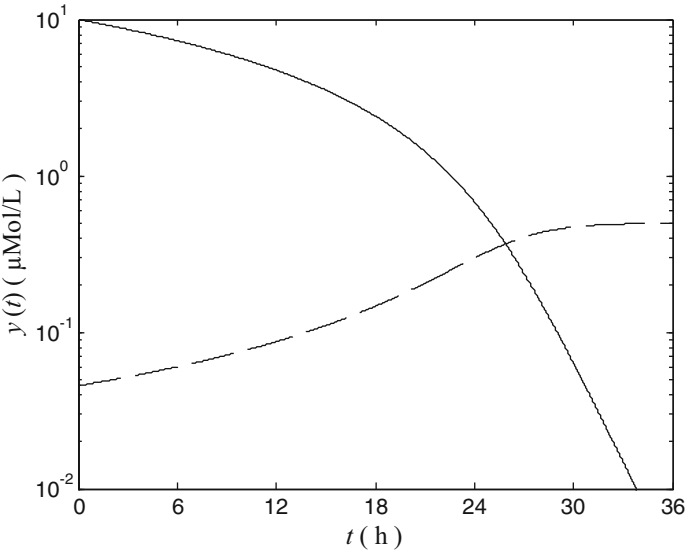


Fig. 7.4 The reference profile of a nonlinear one-compartment model with infusion. The *solid line* represents the kinetic profile and the *dashed line*, the time profile of $k_e(t)$

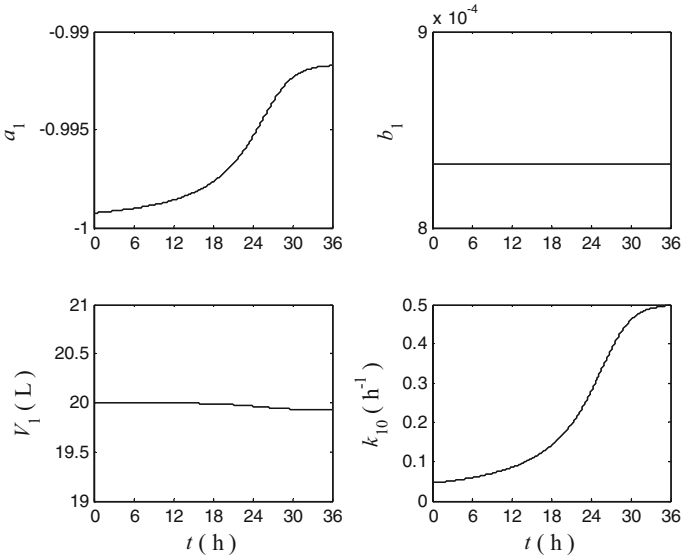


Fig. 7.5 Time-varying coefficients in transfer function (up) and pharmacokinetic parameters (down) vs. time

In a complete reverse way, we attempt to track the $k_{10}(t)$ time variations from the sampled every $T = 1$ min reference profile. The horizon supporting 64% of sampling times and measurements was set at about $h_0 = 40$ min, therefore, $\lambda \approx 0.975$ was used for the on-line estimation of algorithm 7.33. Figure 7.5 presents the time profiles of coefficients in the transfer function polynomials and of pharmacokinetic parameters. The $k_{10}(t)$ is the highest variable parameter and copies the time variations of $\alpha_1(t)$. Inversely, the almost constant $b_1(t)$ ensures negligible variations in $V_1(t)$.

3. Analyze the time profile of the time-varying parameters. Again, Figure 7.6 represents the time-varying pharmacokinetic parameters $V_1(t)$ and $k_{10}(t)$ vs. $y_1(t)$. It is obvious from the down subplot that $k_{10}(t)$ is linked with $y_1(t)$ by a saturable relationship of Michaelis–Menten type. Therefore, by means of the on-line estimation algorithm, the postulated nonlinear relationship was revealed.

In conclusion, we obtain a set of differential equations in continuous time having the form of equations 7.15. This kind of equation is called *bilinear* because of the presence of the $\underline{b}\left[\underline{y}(t)\right]u(t)$ term and it is the general formalism for models in biology, ecology, industrial applications, and socioeconomic processes [338]. Bilinear mathematical models are useful to real-world dynamic behavior because of their variable structure. It has been shown that processes described by bilinear models are generally more controllable and offer better performance in control than linear systems. We emphasize that the unstable inherent character of chaotic systems fits exactly within the complete controllability principle discussed for

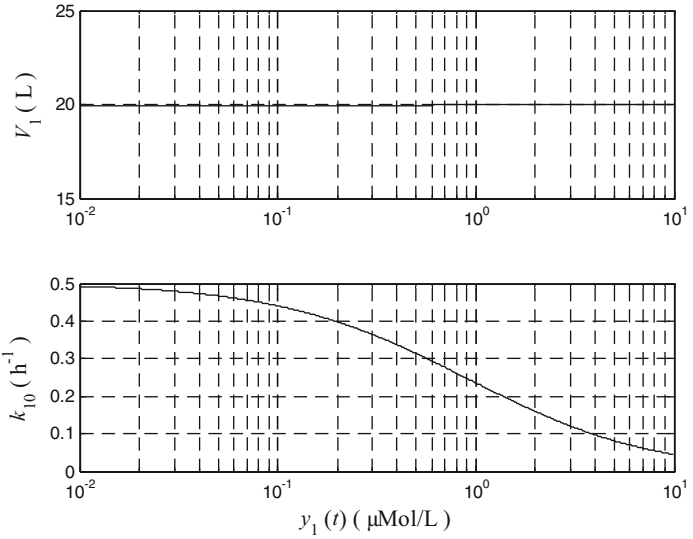


Fig. 7.6 Time-varying pharmacokinetic parameters vs. variations of concentration in the compartment

bilinear mathematical models [338]; additive control may be used to steer the system to new equilibrium points, and multiplicative control, either to stabilize chaotic behavior or to enlarge the attainable space. Then, bilinear systems are of extreme importance in the design and use of optimal control for chaotic behaviors. We can now understand the butterfly effect, i.e., the extreme sensitivity of chaotic systems to tiny perturbations described in Chapter 3.

Chapter 8
Deterministic Compartmental Models

This is Polyfemos the copper Cyclops whose body is full of water and someone has given him one eye, one mouth and one hand to each of which a tube is attached. Water appears to drip from his body and to gush from his mouth, all the tubes have regular flow. When the tube connected to his hand is opened his body will empty within 3 days, while the one from his eye will empty in one day and the one from his mouth in 2/5 of a day. Who can tell me how much time is needed to empty him when all three are opened together?

Metrodorus (331-278 BC)

Compartmental modeling is a broad modeling strategy that has been used in many different fields, though under varying denominations. Virtually all current applications and theoretical research in compartmental analysis are based on deterministic theory. In this chapter deterministic compartmental models will be presented. The concept of compartmental analysis assumes that a process may be divided though it were occurring in homogeneous components, or "compartments." Various characteristics of the process are determined by observing the movement of material. A compartmental system is a system that is made up of a finite number of compartments, each of which is homogeneous and well mixed, and the compartments interact by exchanging material. Compartmental systems have been found useful for the analysis of experiments in many branches of biology.

We assume that compartment i is occupied at time 0 by q_{i0} amount of material and we denote by $q_i(t)$ the amount in the compartment i at time t. We also assume that no material enters the compartments from the outside of the compartmental system and we denote by $R_{i0}(t)$ the rate of elimination from compartment i to the exterior of the system. Let also $R_{ji}(t)$ be the transfer rate of material from the j-th to i-th compartment. Because the material is distributed in each compartment at uniform concentration, we may assume that each compartment occupies a constant volume of distribution $V_{i.}$ The box in Figure 8.1 represents the i-th compartment of a system of m compartments.

Mathematics is now called upon to describe the compartmental configurations and then to simulate their dynamic behavior. To build up mathematical equations

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_8

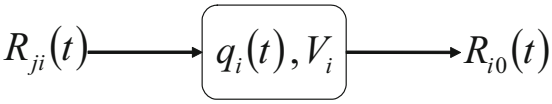


Fig. 8.1 The rates of transfer for the i-th compartment

expressing compartmental systems, one has to express the mass balance equations for each compartment i:

$$\dot{q}_i(t) = -R_{i0}(t) + \sum_{\substack{j=1 \\ j \neq i}}^{m} R_{ji}(t), \tag{8.1}$$

with initial condition $q_i(0) = q_{i0}$. Thus, we obtain m differential equations, one for each compartment i.

8.1 Linear Compartmental Models

Now, some fundamental hypotheses, commonly called laws, were employed to expand the transfer rates appearing in (8.1). Fick's law is largely used in current modeling (cf. Section 2.3 and equation 2.14). It assumes that the transfer rate of material by diffusion between regions l (left) and r (right) with concentrations c_l and c_r, respectively, is

$$R_{lr}(t) = -CL_{lr}(c_r - c_l). \tag{8.2}$$

This law may be applied to the transfer rates $R_{ji}(t)$ of the previous equation for all pairs j and i of compartments corresponding to l and r and for the elimination rate $R_{i0}(t)$, where the concentration is assumed nearly zero in the region outside the compartmental system. One has for the compartment i,

$$\dot{q}_i(t) = -CL_{i0}c_i(t) + \sum_{\substack{j=1 \\ j \neq i}}^{m} CL_{ji}\left[c_j(t) - c_i(t)\right],$$

where CL_{i0} is the *total clearance* from compartment i and CL_{ji} is the *intercompartmental clearance* between i and j. We recall that the clearance has a bidirectional property ($CL_{ji} = CL_{ij}$) and the subscript ij denotes simply the pair of compartments referenced. The initial condition associated with the previous differential equation is denoted by $q_i(0) = q_{i0}$. Using the volumes of distribution V_i and the well-known relationship $q_i(t) = V_i c_i(t)$, we substitute the concentrations with the corresponding amounts of material:

$$\dot{q}_i(t) = -k_{i0}q_i(t) + \sum_{\substack{j=1 \\ j \neq i}}^{m} k_{ji}q_j(t) - \sum_{\substack{j=1 \\ j \neq i}}^{m} k_{ij}q_i(t).$$

The constants k are called the *fractional flow rates*. They have the dimension of time^{-1} and they are defined as follows:

$$\frac{CL_{i0}}{V_i} \triangleq k_{i0}, \qquad \frac{CL_{ij}}{V_i} \triangleq k_{ij}, \qquad \frac{CL_{ij}}{V_j} \triangleq k_{ji}. \tag{8.3}$$

In contrast to the clearance, the fractional flow rates indicate the flow direction, i.e., $k_{ji} \neq k_{ij}$, the first subscript denoting the start compartment, and the second one, the ending compartment. The fractional flow rates and the volumes of distribution are usually called *microconstants*.

When the volume of the compartment being cleared is constant, the assumption that the fractional flow rate is constant is equivalent to assuming that the clearance is constant. But in the general case, in which the volume of distribution cannot be assumed constant, the use of fractional flow rates k is unsuitable, because the magnitude of k depends as much upon the volume of the compartment as it does upon the effectiveness of the removal process. In contrast, the clearance depends only upon the overall effectiveness of removal, and can be used to characterize any removal process whether it be constant or changing, capacity-limited or supply-limited [339].

Through the following procedure the equations for a deterministic model can be obtained:

1. Represent the underlying mechanistic model with the desired physiological structure through a set of phenomenological compartments with their interconnections.
2. For each compartment in the configuration, apply the mass balance law to obtain the differential equation expressing the variation of amount per unit of time. In these expressions, constant or variable fractional flow rates k can be used.
3. Solve the system of differential equations obtained for all the compartments by using classical techniques or numerical integration (e.g., Runge–Kutta) [340].

Therefore, Fick's law, when applied to all elements of the compartmental structure, leads to a system of linear differential equations. There are as many equations as compartments in the configuration. If we set

$$k_{ii} = k_{i0} + \sum_{\substack{j=1 \\ j \neq i}}^{m} k_{ij},$$

the equation for the i-th compartment is

$$\dot{q}_i(t) = -k_{ii} q_i(t) + \sum_{\substack{j=1 \\ j \neq i}}^{m} k_{ji} q_j(t), \tag{8.4}$$

associated with initial conditions q_{i0}. In the previous equation, the $q_i(t)$ and q_{i0} amounts of material can be compiled in vector forms as $\underline{q}(t)$ and $\underline{q}_0$, respectively. In the same way, the fractional flow rates k_{ij} may be considered as the (i,j)-th elements of the $m \times m$ fractional flow rates matrix $\mathbf{K}$. Thus, the set of linear differential equations can be expressed as

$$\dot{\underline{q}}^T(t) = \underline{q}^T(t)\,\mathbf{K},$$

and having the following solution:

$$\underline{q}^T(t) = \underline{q}_0^T \exp(\mathbf{K}t), \tag{8.5}$$

where the initial conditions are *postmultiplied* by $\exp(\mathbf{K}t)$, which is defined by

$$\exp(\mathbf{K}t) \triangleq \mathbf{I} + \sum_{i=1}^{\infty} \frac{\mathbf{K}^i t^i}{i!}.$$

In most pharmacokinetic applications, one can assume that the system is open and at least weakly connected. This is the case of mammillary compartmental models, where the compartment no.1 is referred to as the *central compartment* and the other compartments are referred to as the *distribution compartments*, characterized by $k_{i0} = 0$ and $k_{ij} = 0$ for $i,j = 2 : m$. For open mammillary compartmental configurations, the eigenvalues of $\mathbf{K}$ are distinct, real, and negative, implying that

$$q_i(t) = \sum_{j=1}^{m} B_{ij} \exp\left(-b_j t\right),$$

the so-called formula of *sum of exponentials*, which is common in pharmacokinetics. The B_{ij} and positive b_j are often called *macroconstants*, and they are functions of the microconstants. The equations relating these formulations are given explicitly for the common two- and three-compartment models in many texts [332, 341]. It should be noted, however, that the addition of a few more compartments usually complicates the analysis considerably.

8.2 Routes of Administration

In practice, it is unlikely to have compartmental models with initial conditions unless there are residual concentrations obtained from previous administrations. Drugs are administered either by extravascular, or intravascular in single or repeated experiments. Extravascular routes are oral, or intramuscular routes, and intravascular are the constant rate short- and long-duration infusions.

- For the extravascular route, the rate of administration is

$$u_{ev}(t) = q_0 k_a \exp(-k_a t),$$

where q_0 is the amount of material initially given to the extravascular site of administration and k_a is the fractional flow rate for the passage of material from the site of administration toward the recipient compartment; k_a is the absorption rate constant.

- For the intravascular route with constant rate, we have

$$u_{iv}(t) = \frac{q_0}{T_E - T_S}\left[H(t - T_S) - H(t - T_E)\right],$$

where q_0 is the amount of material given at a constant rate in the venous compartment between the starting time T_S and the ending time T_E. Here, $H(t)$ is the step Heaviside function.

Extravascular and intravascular routes can be conceived as concomitant or repeated, e.g., delayed oral intake with respect to an intramuscular administration, or piecewise constant rate infusions, etc. Applying the superposition principle, the contribution of all administration routes in the same recipient compartment is given by the following input function:

$$u(t) = \sum_{i=1}^{m_{ev}} q_{0i} k_{ai} \exp\left[-k_{ai}(t - T_i)\right] + \sum_{i=1}^{m_{iv}} \frac{q_{0i}}{T_{E_i} - T_{S_i}}\left[H(t - T_{S_i}) - H(t - T_{E_i})\right],$$

where the m_{ev} and m_{iv} administrations preceding the time t are associated with the q_{0i} amounts of material. T_i is the time of the i-th extravascular administration, and T_{S_i} and T_{E_i} are the starting and ending times in the i-th intravascular administration. The contribution of the input function $u(t)$ in the mass balance differential equation for the recipient compartment is represented by an additive term in the right-hand side of (8.1).

8.3 Time–Concentration Profiles

In (8.4), by dividing the amounts $q_i(t)$ by non-time-dependent volumes of distribution V_i, one obtains the differential equations for the concentrations $c_i(t)$:

$$\dot{c}_i(t) = -k_{ii} c_i(t) + \sum_{\substack{j=1 \\ j \neq i}}^{m} \frac{V_j}{V_i} k_{ji} c_j(t). \tag{8.6}$$

Additional assumptions further reduce the complexity of these equations. One such assumption is the incompressibility of the volumes of distribution or, as usually known, the *flow conservation*. This assumption applied to compartment j leads to

$$\sum_{\substack{i=1\\ i\neq j}}^{m} V_i k_{ij} = V_j \sum_{\substack{i=1\\ i\neq j}}^{m} k_{ji}.$$

In the special case of a mammillary compartmental configuration, the above relation allows one to express the volume of distribution in peripheral compartments as functions of the fractional flow rates and the volume of distribution of the central compartment $V_j = \left[k_{1j}/k_{j1}\right] V_1$ for $j = 2 : m$. Substituting this relationship in (8.6), we obtain

$$\dot{c}_i(t) = -k_{ii} c_i(t) + \sum_{\substack{j=1\\ j\neq i}}^{m} k_{ij} c_j(t).$$

This set of linear differential equations can be expressed as $\dot{\underline{c}}(t) = \mathbf{K}\underline{c}(t)$, and it has the following solution:

$$\underline{c}(t) = \exp(\mathbf{K}t)\,\underline{c}_0,$$

where the initial conditions are *premultiplied* by $\exp(\mathbf{K}t)$ (instead of the *postmultiplication* in the case of amounts; cf. equation 8.5).

These equations are widely used to simulate simple or complex compartmental systems and currently to identify pharmacokinetic systems from observed time–concentration data. However, it is not always possible to write the equations in terms of concentrations that represent true physical blood or plasma levels. In practice, it may occur that some, say two, compartments exchange so rapidly on the time scale of an experiment that they are not distinguishable but merge kinetically into one compartment. If the two compartments represent material that exists at different concentrations in two different spaces, or two forms of a compound in one space, the calculated concentration may not correspond to any actual measurable concentration and thus may be misleading. For this reason the development of differential equations in terms of compartment amounts $q_i(t)$ is more general. If these equations are available, it is not difficult to convert to concentrations $c_i(t)$ by assuming that V_i is a proportionality constant, called the *apparent volume of distribution*, and to solve the equations as long as the volumes are constant in time [342]. If the volumes are changing the problem becomes more difficult.

8.4 Random Fractional Flow Rates

The deterministic model with random fractional flow rates may be conceived on the basis of a deterministic transfer mechanism. In this formulation, a given replicate of the experiment is based on a particular realization of the random fractional flow

rates and/or initial amounts $\underline{\Theta}$. Once the realization is determined, the behavior of the system is deterministic. In principle, to obtain from the assumed distribution of $\underline{\Theta}$ the distribution of $q_i(t)$, $i = 1 : m$, the common approach is to use the classical procedures for transformation of variables. When the model is expressed by a system of differential equations, the solution can be obtained through the theory of random differential equations [343–345]. However, in practice, one can find the moments directly using conditional expectations (cf. Appendix D):

$$E[q_i(t)] = E_{\underline{\Theta}}[q_i(t \mid \underline{\Theta})],$$
$$Var[q_i(t)] = Var_{\underline{\Theta}}[q_i(t \mid \underline{\Theta})].$$

Besides the deterministic context, the predicted amount of material is subjected now to a variability expressed by the second equation. This expresses the random character of the fractional flow rate, and it is known as *process uncertainty*. Extensive discussion of these aspects will be given in Chapter 11.

Example 4. One-Compartment Model

As an illustration of the procedure, consider the one-compartment model $q(t) = q_0 \exp(-kt)$. Assuming that k has a gamma distribution $k \sim \text{Gam}(\lambda, \mu)$, one has the solutions

$$E[q(t)] = q_0 E[\exp(-kt)] = q_0 (1 + t/\lambda)^{-\mu},$$
$$Var[q(t)] = q_0^2 Var[\exp(-kt)] = q_0^2 \left[(1 + 2t/\lambda)^{-\mu} - (1 + t/\lambda)^{-2\mu}\right].$$

Figure 8.2 shows $E[q(t)]$ and $E[q(t)] \pm \sqrt{Var[q(t)]}$ with $q_0 = 1$ and $k \sim \text{Gam}(2, 2)$. It is noteworthy that confidence intervals are present due to the variability of the fractional flow elimination rate k. This variability is inherent to the process and completely different from that introduced by the measurement devices. ■

8.5 Nonlinear Compartmental Models

Many systems of interest are actually nonlinear:

- A first formulation considers the transfer rates of material from compartment i to j as functions of the amounts in all compartments $\underline{q}(t)$ and of time t, i.e., $R_{ij}\left[\underline{q}(t), t\right]$. In this case, $R_{ij}(t)$ in (8.1) should be substituted with $R_{ij}\left[\underline{q}(t), t\right]$. If we expand the $R_{ij}\left[\underline{q}(t), t\right]$ in a Taylor series of $\underline{q}(t)$ and retain only the linear terms, the nonlinear transfer rates take the form $k_{ij}(t)\, q_i(t)$ and one obtains a linear time-varying compartmental model.

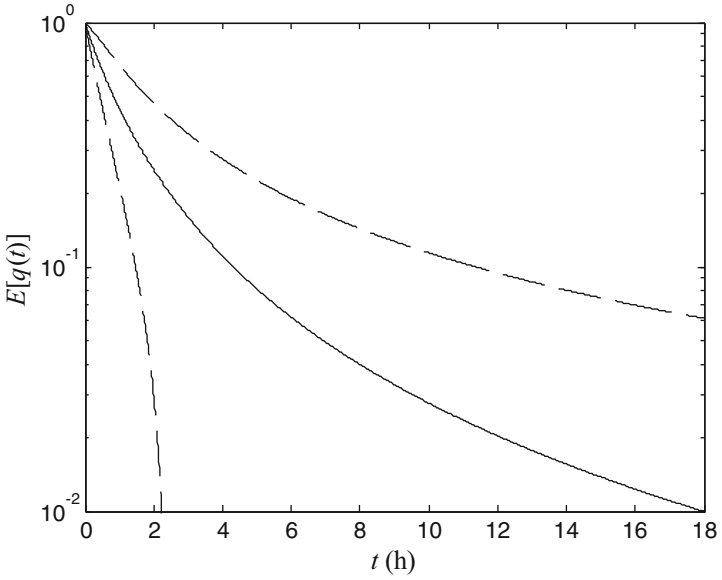


Fig. 8.2 One-compartment model with gamma-distributed elimination flow rate $k \sim \text{Gam}(2, 2)$. The *solid line* represents the expected profile $E[q(t)]$, and *dashed lines*, the confidence intervals $E[q(t)] \pm \sqrt{Var[q(t)]}$

- A second formulation considers the fractional flow rate of material as a function of $\underline{q}(t)$ and t, i.e., $k_{ij}\left[\underline{q}(t), t\right]$. In this case, k_{ij} in (8.4) should be substituted with $k_{ij}\left[\underline{q}(t), t\right]$.

Therefore, the transfer rates and the fractional flow rates are functions of the vector $\underline{q}(t)$ and t. The dependence on t may be considered as the exogenous environmental influence of some fluctuating processes. If no environmental dependence exists, it is more likely that the transfer rates and the fractional flow rates depend only on $\underline{q}(t)$. Nevertheless, since $\underline{q}(t)$ is a function of time, the observed data in the inverse problem can reveal only a time dependency of the transfer rate, i.e., $R_{ij}(t)$, or of the fractional flow rate, i.e., $k_{ij}(t)$. Hence, the dependency of $R_{ij}(t)$ and $k_{ij}(t)$ on $\underline{q}(t)$ is obscured, and a second-level modeling problem now arises, i.e., how to regress the observed dependency on the $\underline{q}(t)$ and t separately. This problem is mentioned in Section 7.7.

Until now, the compartmental model was considered as consisting of compartments associated with several anatomical locations in the living system. The general definition of the compartment allows us to associate in the same location a different chemical form of the original molecule administered into the process. In other words, the compartmental analysis can include not only diffusion phenomena but also chemical reaction kinetics.

One source of nonlinear compartmental models is processes of enzyme-catalyzed reactions that occur in living cells. In such reactions, the reactant combines with

an enzyme to form an enzyme–substrate complex, which can then break down to release the product of the reaction and free enzyme or can release the substrate unchanged as well as free enzyme. Traditional compartmental analysis cannot be applied to model enzymatic reactions, but the law of mass balance allows us to obtain a set of differential equations describing mechanisms implied in such reactions. An important feature of such reactions is that the enzyme is sometimes present in extremely small amounts, the concentration of enzyme being orders of magnitude less than that of substrate.

8.5.1 The Enzymatic Reaction

The mathematical basis for enzymatic reactions stems from work done by Michaelis and Menten in 1913 [346]. They proposed a situation in which a substrate reacts with an enzyme to form a complex, one molecule of the enzyme combining with one molecule of the substrate to form one molecule of complex. The complex can dissociate into one molecule of each of the enzyme and substrate, or it can produce a product and a recycled enzyme. Schematically, this can be represented by

$$\begin{aligned} &[\text{substrate}] + [\text{enzyme}] \underset{k_{-1}}{\overset{k_{+1}}{\rightleftarrows}} [\text{complex}]\,, \\ &[\text{complex}] \overset{k_{+2}}{\rightarrow} [\text{product}] + [\text{enzyme}]\,. \end{aligned} \tag{8.7}$$

In this formulation k_{+1} is the rate parameter for the forward enzyme–substrate reaction, k_{-1} is the rate parameter for the backward reaction, and k_{+2} is the rate parameter for the creation of the product.

Let $s(t)$, $e(t)$, $c(t)$, and $w(t)$ be the amounts of the four species in the reaction (8.7), and s_0 and e_0 the initial amounts for substrate and enzyme, respectively. The differential equations describing the enzymatic reaction,

$$\begin{aligned} \dot{s}(t) &= -k_{+1}s(t)\left[e_0 - c(t)\right] + k_{-1}c(t)\,, & s(0) &= s_0, \\ \dot{c}(t) &= k_{+1}s(t)\left[e_0 - c(t)\right] - (k_{-1} + k_{+2})\,c(t)\,, & c(0) &= 0, \\ \dot{w}(t) &= k_{+2}c(t)\,, & w(0) &= 0, \end{aligned} \tag{8.8}$$

are obtained by applying the law of mass balance for the rates of formation and/or decay, and the conservation law for the enzyme, $e_0 = e(t) + c(t)$.

Relying on a suggestion by Segel [347], we make the variables of the above equations dimensionless

$$x(\tau) = \frac{s(t)}{s_0}, \qquad y(\tau) = \frac{c(t)}{e_0}, \qquad z(\tau) = \frac{w(t)}{s_0},$$

$$\lambda = \frac{k_{+2}}{k_{+1}s_0}, \qquad \kappa = \frac{k_{-1} + k_{+2}}{k_{+1}s_0}, \qquad \varepsilon = \frac{s_0}{e_0},$$

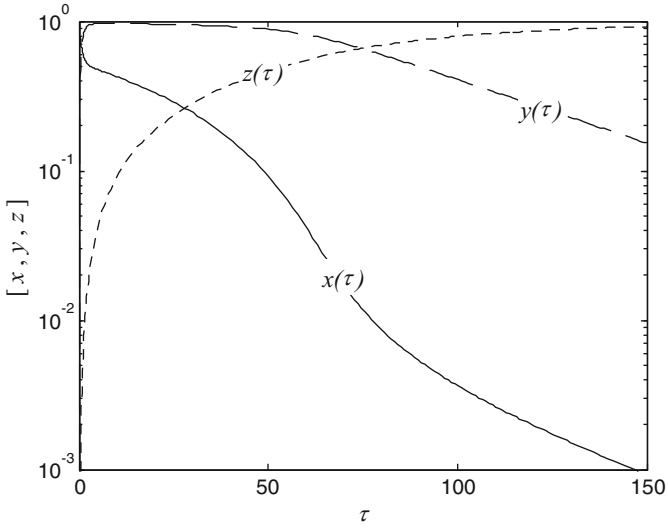


Fig. 8.3 Profiles of dimensionless reactant amounts, substrate $x(\tau)$, complex $y(\tau)$, and product $z(\tau)$

with $\tau = k_{+1}e_0 t$ and $\kappa \geqslant \lambda$. The set of differential equations becomes

$$\begin{aligned}
\dot{x}(\tau) &= -x(\tau)[1 - y(\tau)] + (\kappa - \lambda)\, y(\tau), & x(0) &= 1,\\
\dot{y}(\tau) &= \varepsilon\{x(\tau)[1 - y(\tau)] - \kappa y(\tau)\}, & y(0) &= 0,\\
\dot{z}(\tau) &= \lambda y(\tau), & z(0) &= 0.
\end{aligned}$$

This system cannot be solved exactly, but numerical methods easily generate good solutions. The time courses for all reactant species of reaction (8.7) generated from the previous equations with $(\kappa, \lambda) = (0.015, 0.010)$ and $\varepsilon = 2$ are shown in the semilogarithmic plot of Figure 8.3. We note that:

- The substrate $x(\tau)$ drops from its initial condition value, equal to 1, at a rapid rate, but quickly decelerates. Progressively, and for $\tau > 50$, the substrate decreases rapidly in a first phase and then slowly, in a second phase. This irregular profile of substrate in the semilogarithmic plot is reflected as a *concavity* or *nonlinearity*, as it is usually called.
- The intermediate compound complex $y(\tau)$ reaches a maximum (called *quasi-steady state* in biology) that persists only for a time period and then decreases; this time period corresponds to the period of nonlinearity for the substrate time course. In fact, saturation of the complex form is responsible for the nonlinearity in the substrate time course. During this period, there is no free enzyme to catalyze the substrate conversion toward the product.
- The product $z(\tau)$ reaches the maximum plateau level asymptotically. In contrast to the substrate profile, the nonlinear behavior along the saturation of the complex is not easily defined on the product profile.

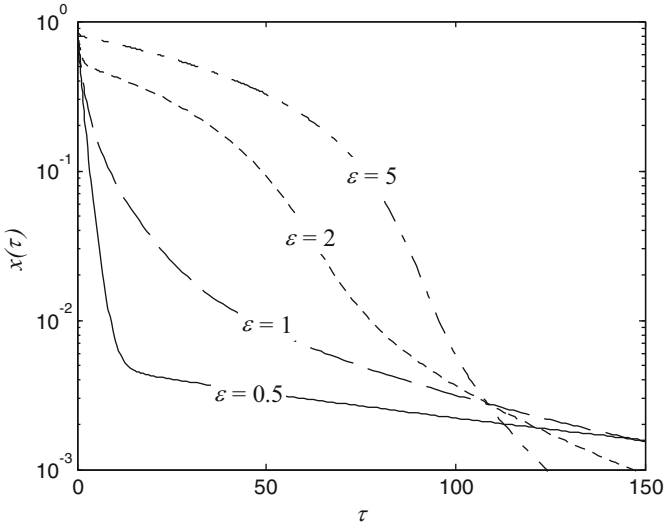


Fig. 8.4 Influence of ε on the substrate $x(\tau)$ profiles with fixed $(\kappa, \lambda) = (0.015, 0.010)$ and $\varepsilon = (0.5, 1, 2, 5)$

Figure 8.4 shows the influence of ε on the $x(\tau)$ shape. For fixed (κ, λ), we simulated the time courses for $\varepsilon = 0.5$, 1, 2, 5. It is noted that the shape of the substrate profiles varies remarkably with the values of ε; thus profiles of biphasic, power-law, and nonlinear type are observed. So, the sensitivity of the kinetic profile regarding the available substrate and enzyme amounts is studied by using several ε values: for low substrate or high enzyme amounts the process behaves according to two decaying convex phases, in the reverse situation the kinetic profile is concave, revealing nonlinear behavior.

Other processes that lead to nonlinear compartmental models are processes dealing with transport of materials across cell membranes that represent the transfers between compartments. The amounts of various metabolites in the extracellular and intracellular spaces separated by membranes may be sufficiently distinct kinetically to act like compartments. It should be mentioned here that Michaelis–Menten kinetics also apply to the transfer of many solutes across cell membranes. This transfer is called *facilitated diffusion* or in some cases *active transport* (cf. Chapter 2). In facilitated diffusion, the substrate combines with a membrane component called a carrier to form a carrier–substrate complex. The carrier–substrate complex undergoes a change in conformation that allows dissociation and release of the unchanged substrate on the opposite side of the membrane. In active transport processes not only is there a carrier to facilitate membrane crossing, the carrier mechanism is somehow coupled to energy dissipation so as to move the transported material up its concentration gradient.

8.6 Complex Deterministic Models

The branching pattern of the vascular system and the blood flow through it has continued to be of interest to anatomists, physiologists, and theoreticians [4, 348, 349]. The studies focusing on the geometric properties such as lengths, diameters, generations, orders of branches in the pulmonary, venular, and arterial tree of mammals have uncovered the principles on which these properties are based. Vascular trees seem to display roughly the same dichotomous branching pattern at different levels of scale, a property found in fractal structures [350–352]. The hydrodynamics of blood flow in individual parts of the dichotomous branching network have been the subject of various studies. Recently, West et al. [353], relying on an elegant combination of the dynamics of energy transport and the mathematics of fractal geometry, developed a hydrodynamic model that describes how essential materials are transported through space-filling fractal networks of branching tubes.

Although these advances provide an analysis of the scaling relations for mammalian circulatory systems, models describing the transport of materials along the entire fractal network of the mammalian species are also needed. Pharmacokinetics and toxicokinetics, the fields in which this kind of modeling is of the greatest importance, are dominated by the concept of homogeneous compartments [354]. Physiologically based pharmacokinetic models have also been developed that define the disposition patterns in terms of physiological principles [282, 354, 355]. The development of models that study the heterogeneity of the flow and the materials distribution inside vascular networks and individual organs has also been fruitful in the past years [294, 356–358]. Herein, we present a simple model for the heterogeneous transport of materials in the circulatory system of mammals, based on a single-tube convection–dispersion system that is equivalent to the fractal network of the branching tubes.

8.6.1 Geometric Considerations

We consider a fractal arterial tree that consists of several branching levels where each level consists of parallel vessels, Figure 8.5A. Each vessel is connected to m vessels of the consequent branching level [353]. We make the assumption that the vessel radii and lengths at each level k follow a distribution around the mean values ρ_k and μ_k, respectively. The variance of the vessel radii and lengths at each level produces heterogeneity in the velocities.

The total flow across a section of the entire tree is constant (conservation of mass). This allows us to replace the tree with a single one-dimensional tube. Since the tree is not area-preserving and the area of the cross section of the tube is equal to the total area of the cross sections of each level of the tree, the total cross-sectional area of subsequent levels increases, i.e., the tube is not cylindrical (Figure 8.5A–C).

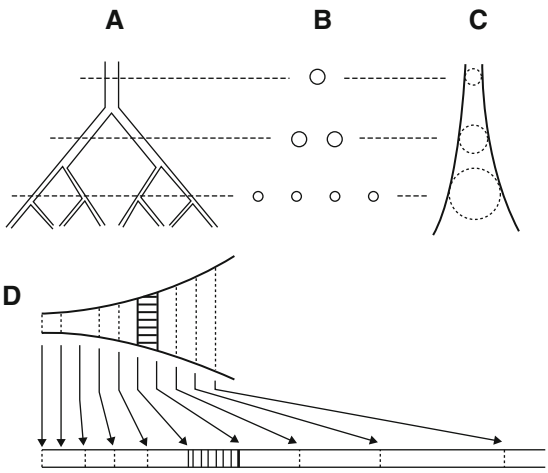


Fig. 8.5 (**A**) Schematic representation of the dichotomous branching network. (**B**) Cross sections at each level. (**C**) Single tube with continuously increasing radius. (**D**) Volume-preserving transformation of the varying radius tube to a fixed radius tube. Reprinted from [359] with permission from Springer

Based on the scaling properties of the fractal tree, the noncylindrical tube is described in terms of a continuous spatial coordinate, z, which replaces the branching levels of the fractal tree from the aorta to the capillaries. As suggested by West [353], both the radii and the vessel lengths scale according to "cubic law" branching, i.e., $\rho_{k+1}/\rho_k = \mu_{k+1}/\mu_k = m^{-1/3}$. These assumptions allow us to obtain the expression for the area $\mathcal{A}(z)$ of the noncylindrical tube (Figure 8.5C) as a function of the coordinate z:

$$\mathcal{A}(z) = \frac{\pi \rho_0^2 \mu_0 \widetilde{m}}{z\,(1-\widetilde{m}) + \mu_0 \widetilde{m}}, \tag{8.9}$$

where ρ_0 and μ_0 are the radius and the length of aorta, respectively, and $\widetilde{m} = m^{1/3}$.

Further, a volume-preserving transformation allows the replacement of the varying radius tube with a tube of fixed radius ρ_0 and fixed area $\mathcal{A}_0 = \pi \rho_0^2$ (Figure 8.5D). This is accomplished by replacing z with a new coordinate z^* with the condition that the constant total flow of the fluid across a section is kept invariant under the transformation:

$$z = \frac{\mu_0 \widetilde{m}}{\widetilde{m} - 1} \left\{ 1 - \exp\left[\frac{z^* \,(1-\widetilde{m})}{\mu_0 \widetilde{m}} \right] \right\}. \tag{8.10}$$

8.6.2 Tracer Washout Curve

The disposition of a solute in the fluid as it flows through the system is governed by convection and dispersion. The convection takes place with velocity

$$v\,(z) = \frac{\mathcal{A}_0}{\mathcal{A}(z)} v_0, \tag{8.11}$$

where v_0 is the velocity in the aorta and $\mathcal{A}(z)$ is given by (8.9). If molecular diffusion is considered negligible, dispersion is exclusively geometric and consists of two components originating from the variance of the path lengths and of the vessel radii. Because the components are independent of each other, the global form of the dispersion coefficient is

$$D(z) = \left[k_1\sigma_{10}^2 + 2k_2\sigma_{20}^2\frac{\mu_0}{\rho_0}\right]\left[\frac{\mathcal{A}_0}{\mathcal{A}(z)}\right]^2 v_0, \tag{8.12}$$

where k_1 and k_2 are proportionality constants, and σ_{10}^2 and σ_{20}^2 are the variances of the radius and the length of aorta, respectively [357, 360, 361]. The equation that describes the concentration $c(z,t)$ of solute inside the tube is a convection–dispersion partial differential equation:

$$\frac{\partial c(z,t)}{\partial t} = \frac{\partial}{\partial z}\left[D(z)\frac{\partial c(z,t)}{\partial z}\right] - v(z)\frac{\partial c(z,t)}{\partial z}$$

with $D(z)$ and $v(z)$ given by (8.12) and (8.11), respectively. Applying the transformation (8.10), the previous equation becomes a simple convection–dispersion equation with constant coefficients:

$$\frac{\partial c(z^*,t)}{\partial t} = D_0^*\frac{\partial^2 c(z^*,t)}{\partial z^{*2}} - v_0^*\frac{\partial c(z^*,t)}{\partial z^*}, \tag{8.13}$$

where

$$D_0^* = k_0 v_0,\ \ v_0^* = \left(\widetilde{m}\frac{k_0}{\mu_0} + 1\right)v_0,\ \ k_0 = k_1\sigma_{10}^2 + 2k_2\sigma_{20}^2\frac{\mu_0}{\rho_0}.$$

These forms relate the dependence on the system characteristics. Equation (8.13) describes the concentration $c(z^*,t)$ of a solute in a tree-like structure that corresponds to the arterial tree of a mammal. Considering also the corresponding venular tree situated next to the arterial tree and appropriate inflow and outflow boundary conditions, we are able to derive an expression for the spatiotemporal distribution of a tracer inside a tree-like transport network. We also make the assumption that the arterial and venular trees are symmetric, that is, have the same volume V; then, the total length is $L = V/\mathcal{A}_0$. The initial condition is $c(z^*,0) = 0$ and the boundary conditions are:

- Inflow at $z^* = 0$:

$$\left[-D_0^*\frac{\partial c(z^*,t)}{\partial z^*} + v_0^* c\left(z^*,t\right)\right]\Big|_{z^*=0} = \frac{q_0}{a_0}\delta(t)$$

  where q_0 is the dose, and $\delta(t)$ is the Dirac delta function.

- Outflow at $z^* = L$:

$$\left.\frac{\partial c\left(z^*, t\right)}{\partial z^*}\right|_{z^*=L} = 0.$$

The outflow concentration $c\left(L, t\right)$ of the above model describes tracer washout curves from organs that have a tree-like network structure, and it is given by an analytic form reported in [359].

8.6.3 Model for the Circulatory System

Based on the above, an elementary pharmacokinetic model considering the entire circulatory system was constructed. Thus, apart from the arterial and venular trees, a second set of arterial and venular trees, corresponding to the pulmonary vasculature, must be considered as well. These trees follow the same principles of (8.10) and (8.13), i.e., tubes of radius ρ_0 are considered with appropriate length to accommodate the correct blood volume in each tree.

8.6.3.1 Structure

An overall tube of appropriate length L is considered and is divided into four sequential parts, characterized as arterial, venular, pulmonary arterial, and pulmonary venular, Figure 8.6.

We assign the first portion of the tube length from $z^* = 0$ to $z^* = z_c^*$ to the arterial tree, the next portion from $z^* = z_c^*$ to $z^* = z_p^*$ to the venular, and the rest from $z^* = z_p^*$ to $z^* = L$ to the two symmetrical trees of the lungs. We consider that the venular tree is a structure similar to the arterial tree, only of greater, but fixed, capacity. Also, the two ends of the tube are connected, to allow recirculation of the fluid. This is implemented by introducing a boundary condition, namely $c\left(0, t\right) = c\left(L, t\right)$, which makes the tube ring-shaped. The "heart" is located at two separate points. The left ventricle-left atrium is situated at $z^* = 0$, and the right ventricle-right atrium is situated at $z^* = z_p^*$, Figure 8.6.

8.6.3.2 Dispersion

Two separate values were used for the dispersion coefficient D_a for the arterial segment and D_p for the pulmonary segment. For the venular segment we consider that the dispersion coefficient has the value $D_a\left(z_p^* - z_c^*\right)/z_c^*$, meaning that it is proportional to the length of the segment. The flux preservation boundary condition,

$$D_p \left.\frac{\partial c\left(z^*, t\right)}{\partial z^*}\right|_{z^*=L} = D_a \left.\frac{\partial c\left(z^*, t\right)}{\partial z^*}\right|_{z^*=0},$$

must also be satisfied.

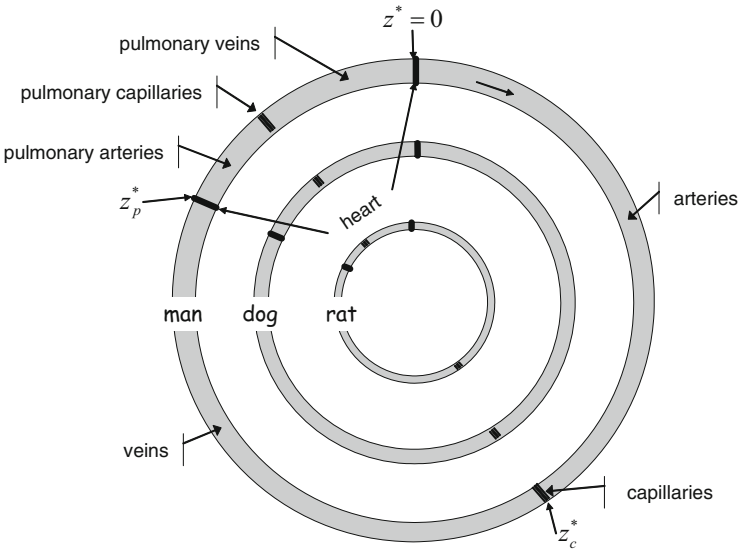


Fig. 8.6 Schematic representation of the ring-shaped tube that models the circulatory system of a mammal. Blood flows clockwise. The tube is divided into segments corresponding to the arterial, venular, pulmonary arterial, and pulmonary venular trees

8.6.3.3 Elimination

The contribution of elimination of drugs is appreciable and is integrated into the model. A segment in the capillary region of the tube ($z^* \approx z_c^*$) is assigned as the elimination site and a first-order elimination term $kc\,(z^*, t)$ is now introduced in (8.13). The length of the elimination segment is arbitrarily set to $0.02L$, which is in the order of magnitude of the capillary length. The position of the elimination site is imprecise in physiological terms, but it is the most reasonable choice in order to avoid further model complexity.

8.6.3.4 Drug Administration and Sampling

The necessary initial condition for the intravenous administration of an exogenous substance, $c\,(z^*, 0)$, which is the spatial profile of c at the time of administration, is determined by the initial dose and the type of administration. This profile may have the shape of a "thin" Gaussian function if an intravenous bolus administration is considered, or the shape of a "rectangular" gate for constant infusion. The reference location z_0^* of this profile for an intravenous administration must be set close to the heart. Similarly, when lung administration is considered, z_0^* should be set in the capillary area of the lungs. Due to the geometric character of the model, a sampling site z_s^* should be either specified, in simulation studies, or calculated when fitting is performed.

The final model can be summarized as follows:

$$\frac{\partial c\left(z^*,t\right)}{\partial t}=\frac{\partial}{\partial z^*}\left[D^*\left(z^*\right)\frac{\partial c\left(z^*,t\right)}{\partial z^*}\right]-v_0^*\frac{\partial c\left(z^*,t\right)}{\partial z^*}-W\left(z^*\right)kc\left(z^*,t\right),$$

where $W\left(z^*\right)$ is a combination of delayed in space Heaviside functions, i.e., $W\left(z^*\right)=H\left(z^*-z_c^*+0.01L\right)-H\left(z^*-z_c^*-0.01L\right)$, and

$$D^*\left(z^*\right)=\begin{cases}D_a & \text{for } 0<z^*\le z_c^*,\\ D_a\left(z_p^*-z_c^*\right)/z_c^* & \text{for } z_c^*<z^*\le z_p^*,\\ D_p & \text{for } z_p^*<z^*\le L.\end{cases}$$

Boundary and initial conditions are considered as discussed above.

Example 5. Indocyanine Green Injection

The model was used to identify indocyanine green profile in man after a $q_0 = 10\,\text{mg}$ intravenous bolus injection. Both injection and sampling sites (z_0^* and z_s^*, respectively) were closely located on the ring-shaped tube. The model of drug administration was a "thin" Gaussian function:

$$c\left(z^*,0\right)=\frac{q_0}{V}\sqrt{\frac{b}{\pi}}\exp\left[-b\left(\frac{z^*}{L}-\frac{z_0^*}{L}\right)^2\right].$$

This administration corresponds to a bolus injection at the cephalic vein. The parameters set in the model were $m = 3$, $\mu_0 = 50\,\text{cm}$, $\mathcal{A}_0 = 3\,\text{cm}^2$, and $b = 10^5$. The estimated model parameters were:

- Structure: $z_c^*/L = 0.28$, $z_0^*/L = 0.83$, $z_p^*/L = 0.85$, and $V = 4.4\,\text{l}$. These values result in $L = 1470\,\text{cm}$.
- Dispersion and elimination: $D_a = 1826\,\text{cm}^2\,\text{s}^{-1}$, $D_p = 1015\,\text{cm}^2\,\text{s}^{-1}$, $v_0 = 44.98\,\text{cm}\,\text{s}^{-1}$, and $k = 1.13\,\text{s}^{-1}$.

Figure 8.7 depicts the fitted concentration profile of indocyanine green at the sampling site along with the experimental data. ■

A one-dimensional linear convection–dispersion equation was developed with constant coefficients that describes the disposition of a substance inside a tree-like fractal network of tubes that emulates the vascular tree. Based on that result, a simple model for the mammalian circulatory system is built in entirely physiological terms consisting of a ring-shaped, one-dimensional tube. The model takes into account dispersion, convection, and uptake, describing the initial mixing of intravascular tracers. This model opens new perspectives for studies dealing with the disposition of intravascular tracers used for various hemodynamic purposes, e.g., cardiac output measurements [362, 363], volume of circulating blood determination [362], and liver function quantification [364]. Most importantly, the model can be expanded and used for the study of xenobiotics that distribute beyond the intravascular space.

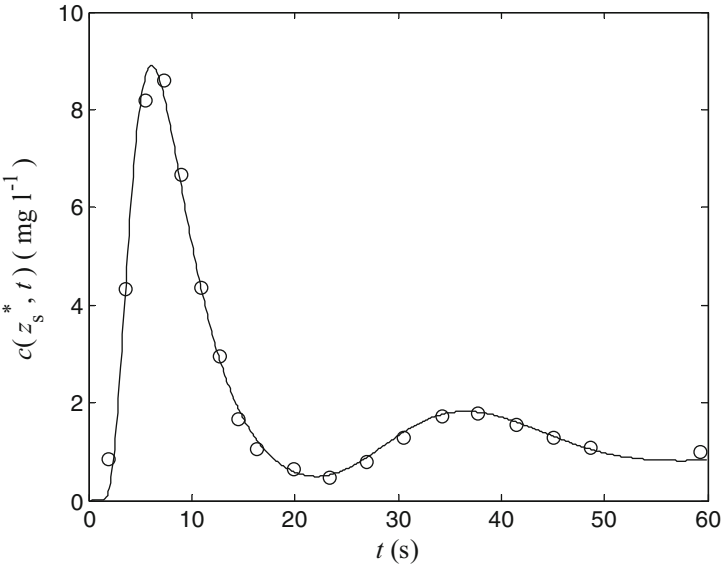


Fig. 8.7 Indocyanine profile at the sampling location $z_s^* = 1220$ cm after intravenous bolus administration of 10 mg. The peaks correspond to successive passes of the drug bolus from the sampling site as a result of recirculation. The dots indicate the experimental data

In future developments of the model, the positioning of organs that play an important role in the disposition of substances can be implemented by adding parallel tubes at physiologically based sites to the present simple ring-shaped model. Consequently, applications can be envisaged in interspecies pharmacokinetic scaling and physiologically based pharmacokinetic-toxicokinetic modeling, since both fields require a realistic geometric substrate for hydrodynamic considerations.

8.7 Compartmental Models and Heterogeneity

Initially, the deterministic theory was applied to describe the movement of a population of tracer molecules. Briefly, a drug administered as a bolus input into an organ modeled by homogeneous compartments results in a time–concentration curve describing the amount of the drug remaining in the organ as a function of the elapsed time of the form of a sum of exponential terms. Possibly because the individual molecules are infinitesimal in size, in most of the literature the implicit assumption is made of deterministic flow patterns. So, compartmental analysis, grounded on deterministic theory, has provided a rich framework for quantitative modeling in the biomedical sciences with many applications to tracer kinetics in general [365, 366] and also to pharmacokinetics [341]. The linear combinations of exponential function forms have provided a very rich class of curves to fit to time–concentration data, and compartmental models turn out to be good approximations for many processes.

Thus, compartmental models have been used extensively in the pharmacokinetic literature for some time, but not without criticism. These criticisms were directed:

- First, at the compartmental approach per se grounded on the assumption of homogeneous compartments. Compartmental models are in fact appropriate when there is an obvious partitioning of the material in the process into discrete portions, the compartments that exchange amounts of materials. From a theoretical standpoint, there has always been a consensus that the notion of a homogeneous compartment is merely a simplified representation for different tissues that are pooled together [367, 368].
- Second, at the fact that the models obtained are not necessarily exact because mixing in a compartment is not instantaneous. How good a compartment model is depends on the relative rates of mixing within a compartment as compared to the transfer rates between the compartments. Mixing may occur by diffusion, various types of convection, and combinations of them, so it is difficult to come up with a uniform theory of mixing. Ideally, we should measure the concentration of material throughout the process and define mixing in terms of the time course of a ratio such as the standard deviation divided by the mean concentration.
- Third, at the ill-conditioning of numerical problems for parameter estimation with models involving a large number of exponential terms. Wise [324] has developed a class of powers of time models as alternatives to the sums of exponentials models and has validated these alternative models on many sets of experimental data. From an empirical standpoint, Wise [269] reported "1000 or more" published time–concentration curves where alternative models fit the data as well or better than the sums-of-exponentials models.

Moreover, it is clear that even the continuous models are often unreliable models. Matter is atomic, and at a fine enough partition, continuity is no longer an acceptable solution. Furthermore, living tissues are made up of cells, units of appreciable size that are the basic structural and functional units of living things. And cells are not uniform in their interiors; they contain smaller units, the cellular organelles. There is inhomogeneity at a level considerably above the molecular. All these facts enhanced the criticism against determinism and the use of homogeneous compartments. More realistic alternatives have aimed at removing the limiting assumption of homogeneity:

- The process was considered as continuous and compartmental models were used to approximate the continuous systems [366]. For such applications, there is no specific compartmental model that is the best; approximation improves as the number of compartments is increased. In order to put compartmental models of continuous processes in perspective it may help to recall that the first step in obtaining the partial differential equation, descriptive of a process continuous in the space variables, is to discretize the space variables so as to give many *microcompartments*, each uniform in properties internally. The differential equation is then obtained as the limit of the equation for a microcompartment as its spatial dimensions go to zero. It is better to approximate the continuous processes with a finite-compartment system rather than go to the limit. In that

case, the partial differential equation is approximated by a set of simultaneous ordinary differential equations. In philosophy, compartmental modeling shares basic ideas with the finite element method, where the structure of the system is also used to define the elements of a partition of the system. But even if a finite-compartment approximation is used, how can we define the approximation error and its dependence on the size of the compartmental model? In addition, many compartmental models approximating continuous processes are so large that it may be difficult to deal with them and it may be useful or necessary to lump some of the compartments into one compartment. This raises a set of questions about the errors incurred in aggregation and about the optimal way of aggregating compartments.

- Noncompartmental models were introduced as models that allow for transport of material through regions of the body that are not necessarily well mixed or of uniform concentration [273]. For substances that are transported relatively slowly to their site of degradation, transformation, or excretion, so that the rate of diffusion limits their rate of removal from the system, the noncompartmental model may involve diffusion or other random-walk processes, leading to the solution in terms of the partial differential equation of diffusion or in terms of probability distributions. A number of noncompartmental models deal with plasma time–concentration curves that are best described by power functions of time.
- Physiological and circulatory models have been developed, and they have provided information of physiological interest that was not available from compartmental analysis. Rapidly, physiological models turned to the modeling of complex compartmental structures. In contrast, circulatory models associated with a statistical framework have proved powerful in describing heterogeneity in the process [271, 369]. Recently, the above presented complex model for the entire circulatory system was built, describing initial mixing following an intravascular administration in a tree-like network by a relatively simple convection–dispersion equation [359, 370].
- Stochastic compartmental analysis assumes probabilistic behavior of the molecules in order to describe the heterogeneous character of the processes. This approach is against the unrealistic notion of the "well-stirred" system, and it is relatively simpler mathematically than homogeneous multicompartment models. At first glance, this seems to be a paradox since the conventional approaches rely on the simpler hypothesis of homogeneity. Plausibly, this paradox arises from the analytical power of stochastic approaches and the unrealistic hypothesis of homogeneity made by compartmental analysis. Nevertheless with only a few exceptions, stochastic modeling has been slow to develop in pharmacokinetics and only recently have some applications also included stochastic behavior in the models.

In conclusion, compartmental models are generally well determined if there is an obvious partitioning of the material into compartments, and if the mixing processes within these compartments are considerably faster than the exchanges between the compartments.

Chapter 9
Fractional Pharmacokinetics

It leads to a paradox, from which one day useful consequences will be drawn.

Gottfried Wilhelm von Leibnitz (1646–1716)
when asked (1695) by Guillaume de L'Hospital
"what would be the result of half-differentiating?"

by Dr. A. DOKOUMETZIDIS
Faculty of Pharmacy National and Kapodistrian
University of Athens

Diffusion is one of the main mechanisms of various processes in living species and as such, plays an important role in the course of a drug in the body. Processes like membrane permeation, dissolution of solids, and dispersion in cellular matrices are considered to take place by diffusion. As mentioned in Chapter 2, diffusion is classically described by Fick's law and is based on the fact that a molecule makes a random walk, where its mean squared displacement is proportional to time. However, in the last few decades, strong experimental evidence has suggested that this is not always true and diffusional processes that deviate from this law have been observed. These are either faster (super-diffusion) or slower (sub-diffusion) than the classic case and the mean square displacement is a power of time, with exponent greater or less than 1, respectively, [371]. This type of diffusion gives rise to kinetics that are referred to as anomalous, to indicate the fact that deviate from the classic description [371]. Moreover, anomalous kinetics can also result from *reaction-limited processes* and *long-time trapping*. It is thought that anomalous kinetics introduces memory effects in the process that need to be accounted for to correctly describe it. As mentioned in Chapter 7, a theory that describes such anomalous kinetics is the so- called *fractal kinetics* where explicit power functions of time, in the form of time-dependent coefficients, are used to account for the memory effects. In pharmacokinetics several data sets have been characterized by power laws [269, 372] which has been justified by the presence of anomalous diffusion. These include mainly pharmacokinetics of drugs that distributed in deeper tissues [369] and bone seeking elements [281, 373].

An alternative theory to describe anomalous kinetics uses *fractional calculus* [30, 374], which introduces derivatives and integrals of fractional order, such as

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_9

half or 3 quarters. Although fractional calculus was introduced by Leibniz more than 300 years ago, it is only within the last couple of decades that real-life applications have been explored [375–377]. It has been shown that differential equations with *fractional derivatives* describe experimental data of anomalous diffusion more accurately. In this chapter the recent applications of fractional calculus in pharmacokinetics are presented, which have formed a new area of research referred to as *fractional pharmacokinetics*.

9.1 Fractional Calculus

9.1.1 The Fractional Derivative

Derivatives of integer order n, $\mathrm{d}^n f(t)/\mathrm{d}t^n$ of a function $f(t)$ are well defined. For a fractional order of differentiation α, where for simplicity we assume that $0 < \alpha < 1$, the α-th derivative is defined through *fractional integration* and successive ordinary differentiation. Fractional integration of order α is defined, according to the *Riemann–Liouville* integral [378]

$$ {}_0D_t^{-\alpha} f(t) \triangleq \frac{1}{\Gamma(\alpha)} \int_0^t \left(t - t'\right)^{\alpha-1} f\left(t'\right) \mathrm{d}t' $$

where $\Gamma(.)$ is the gamma function. Consequently, fractional differentiation is defined as

$$ {}_0D_t^{\alpha} f(t) \triangleq \frac{\mathrm{d}}{\mathrm{d}t} \left[\frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{f(t')}{(t-t')^{\alpha}} \mathrm{d}t' \right]. $$

This is the Riemann–Liouville definition of the fractional derivative. One can notice that the fractional integration is basically a convolution integral between the function and a power law of time, i.e., ${}_0D_t^{-\alpha} f(t) = t^{\alpha-1} * f(t)$, where the star "$*$" denotes convolution, accounting for the memory effects of the studied process. The fractional derivatives have properties that are not intuitive, for example, the half derivative of a constant λ with respect to x does not vanish and instead is $\lambda/\sqrt{\pi x}$. The left-side index "0" of the D operator denotes the lower end of the integration which in this case has been assumed to be zero. However alternative lower bounds can be considered leading to different definitions of the fractional derivative with slightly different properties. An alternative lower bound which has been considered is "$-\infty$" and is referred to as the Wyel definition [375], which accounts for the entire "history" of the studied function, and is considered preferable in some applications. In fact one of the disadvantages of the Riemann–Liouville definition with the "0" lower bound is that when used in differential equations it gives rise to initial conditions that involve the fractional integral of the function and are difficult to interpret physically. This is one of the reasons the Wyel definition has been introduced, but this definition may not be very practical for most applications either, as it involves an initial condition at time $-\infty$.

An alternative definition of the fractional derivative which is referred to as the Caputo definition is preferable for most physical processes as it involves explicitly the initial condition at time zero [374]. The definition is

$$ {}_0^C D_t^\alpha f(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{\dot{f}(t')}{(t-t')^\alpha} \mathrm{d}t' $$

where the upper-left index "C" stands for Caputo and the "dot" in $\dot{f}(t)$ denotes classic differentiation. This definition for the fractional derivative, apart from the more familiar initial conditions, gives rise to more familiar properties, one of them being that the *Caputo derivative* of a constant is in fact zero as usual. The different definitions of the fractional derivative give different results but these are not contradicting, since they apply for different conditions and it is a matter of choosing the appropriate one for each specific application. All the definitions collapse to the usual derivative for integer values of the order of differentiation.

9.1.2 Fractional Differential Equations

The most common type of kinetics encountered in pharmaceutical literature are zero- and first-order kinetics. The fractional versions of these types of kinetics are presented below and take the form of fractional order ordinary differential equations (FDE). Throughout this presentation the Caputo version of the fractional derivative is considered for the reasons already explained.

9.1.2.1 Zero-Order Kinetics

The classical zero-order kinetics equation, where the rate of change of quantity $q(t)$, expressed in mass units, is considered to be constant and equal to k_0, expressed in (mass) / (time) units, is given by $\dot{q}(t) = k_0$. Its solution is a linear function of time and when the initial condition is zero, it has the form $q(t) = k_0 t$. The fractional expression for the zero-order kinetics equation can be obtained by replacing the derivative of order 1 by a derivative of fractional order α

$$ {}_0^C D_t^\alpha q(t) = k_{0f} $$

where k_{0f} is a constant with dimension (mass) / (time)$^\alpha$. The solution of this equation for initial condition $q(t) = 0$ is a power law [374]

$$ q(t) = \frac{k_{0f}}{\Gamma(\alpha+1)} t^\alpha. $$

9.1.2.2 First-Order Kinetics

The first-order differential equation, where the rate of change of quantity $q(t)$ is proportional to its current value, is given by $\dot{q}(t) = k_1 q(t)$. Its solution by considering an initial condition of $q(t) = q_0$ is given by the classical equation of exponential decay $q(t) = q_0 \exp(-k_1 t)$. In fractional terms however, the first-order equation can be written by replacing the derivative of order 1 by a fractional one

$$ {}_0^C D_t^{\alpha} q(t) = -k_{1f} q(t) \tag{9.1} $$

where k_{1f} is a constant with $(\text{time})^{-\alpha}$ dimension. The solution of this equation can be found in most books or papers of the fast growing literature on fractional calculus [374] and for initial condition $q(t) = q_0$ it has the form

$$ q(t) = q_0 E_{\alpha}\left(-k_{1f} t^{\alpha}\right) $$

where $E_{\alpha}(.)$ is a *Mittag–Leffler function* [374] which is defined as

$$ E_{\alpha}(x) \triangleq \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(\alpha k + 1)}. $$

The function $E_{\alpha}(x)$ is a generalization of the exponential function and it collapses to the exponential when $\alpha = 1$, i.e., $E_1(x) = \exp(x)$. The general form of the Mittag–Leffler function with two parameters α and β is also defined as

$$ E_{\alpha,\beta}(x) \triangleq \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(\alpha k + \beta)}. $$

The solution of equation 9.1 basically means that the fractional derivative of order α of the function $E_{\alpha}(t^{\alpha})$ is itself a function of the same form, exactly like the classic derivative of an exponential is also an exponential. It also makes sense to restrict α to values $0 < \alpha < 1$, since for values of α larger than 1 the solution of equation 9.1 is non-monotonous and negative values for $q(t)$ appear.

From these elementary equations the basic relations for the time evolution in drug disposition can be formulated, with the assumption of diffusion of drug molecules taking place in heterogeneous space.

9.1.3 Solving FDE by the Laplace Transform

FDE can easily be written in the Laplace domain since each of the fractional derivatives can be transformed similarly to the ordinary derivatives, as follows, for order $\alpha \leq 1$

$$L\left\{{}_0^C D_t^\alpha f(t)\right\} = s^\alpha \tilde{f}(s) - s^{\alpha-1} f(0) \tag{9.2}$$

where $\tilde{f}(s)$ is the Laplace transform of $f(t)$ [374]. For $\alpha = 1$, the Laplace transform 9.2 collapses to the classic expression for ordinary derivatives, i.e., $L\{f(t)\} = s\tilde{f}(s) - f(0)$.

For example, the following simple FDE

$${}_0^C D_t^{1/2} q(t) = -q(t)$$

with initial value $q(0) = 1$ can be written in the Laplace domain, by applying the transform 9.2, as follows

$$s^{1/2}\tilde{q}(s) - s^{-1/2} q(0) = -\tilde{q}(s).$$

By substituting the initial value, the above can be rearranged as

$$\tilde{q}(s) = \frac{1}{s + \sqrt{s}}$$

By applying the inverse Laplace transform to the previous expression, the analytical solution of the FDE can be obtained; it is a Mittag–Leffler function of order one-half, and

$$q(t) = E_{1/2}\left(-t^{1/2}\right) = \exp(t)\left[1 + \operatorname{erf}\left(-\sqrt{t}\right)\right].$$

Although it is always easy to transform an FDE in the Laplace domain and in most cases feasible to rearrange it, solving it explicitly for the system variables, it is more difficult to apply the inverse Laplace transform, such that an analytical solution in the time domain is obtained. However, it is possible to perform that step numerically using a Numerical Inverse Laplace Transform (NILT) algorithm [379].

9.2 Fractional Calculus in Pharmacokinetics

9.2.1 A Basic Fractional Model

In the simplest pharmacokinetic relationship, the intravenous bolus injection with first-order elimination in a one-compartment model, the drug concentration $c(t)$ follows the common expression

$$\dot{c}(t) = -k_{10} c(t)$$

where k_{10} is the elimination rate constant. The fractional version of this equation [380] can be written as

$$ {}_0^C D_t^{\alpha} c(t) = -k_{10f} c(t) $$

where k_{10f} is a constant with dimensions $(\text{time})^{-\alpha}$. As already mentioned, the solution of this equation can be written as

$$ c(t) = c_0 E_{\alpha}\left(-k_{10f} t^{\alpha}\right) \qquad \text{for} \qquad \alpha < 1 \tag{9.3} $$

where c_0 is the ratio (dose) / (volume of distribution). By introducing $t_0 = k_{10f}^{-\frac{1}{\alpha}}$, the above equation becomes

$$ c(t)/c_0 = E_{\alpha}\left[-(t/t_0)^{\alpha}\right]. \tag{9.4} $$

This equation for small times behaves as a *stretched exponential*, i.e.,

$$ \exp\left[-\frac{(t/t_0)^{\alpha}}{\Gamma(1+\alpha)}\right], $$

while for large values of time as a power law, i.e.,

$$ \frac{(t/t_0)^{-\alpha}}{\Gamma(1-\alpha)} $$

(Figure 9.1). It is therefore a good candidate to describe various data sets exhibiting power-law-like kinetics due to the slow diffusion of the drug in deeper tissues.

The relationship 9.3 applies for the simplest case of fractional pharmacokinetics. It accounts for the anomalous diffusion process, which may be considered to be the limiting step of the entire kinetics. Classic clearance may be considered not to be the limiting process here and is absent from the equation.

9.2.2 *Fractionalizing Linear Multicompartment Models*

A single ordinary differential equation is easily fractionalized by changing the derivative on the left-hand side to a fractional order as in the previous paragraph. However, in pharmacokinetics and other fields where compartmental models are used, two or more ordinary differential equations are often necessary and it is not as straightforward to *fractionalize systems of differential equations*, especially when certain properties such as mass balance need to be preserved.

When a compartmental model with two or more compartments is being built, typically an outgoing mass flux becomes an incoming flux to the next compartment.

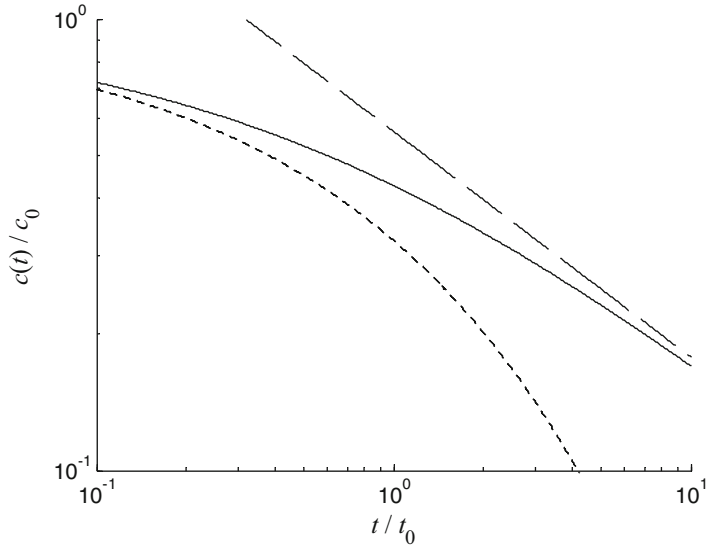


Fig. 9.1 Log–log plot of relationship 9.4 (*solid line*), the stretched exponential function (*dotted line*), and the power-law function (*dashed line*) for $\alpha = 0.5$. The relationship 9.4 starts close to the stretched exponential and finishes close to the power function

Thus, an outgoing mass flux that is defined as a rate of fractional order cannot appear as an incoming flux into another compartment, as a rate of a different fractional order, without violating mass balance [381]. It is therefore, in general, impossible to fractionalize multicompartment systems simply by changing the order of the derivatives on the left-hand side of the ordinary differential equations. The latter is possible only in the special case where a common fractional order is considered for all ordinary differential equations, what is referred of commensurate order. In the general case, of the non-commensurate orders, a different approach for fractionalizing systems of ordinary differential equations needs to be applied.

In the following part, a general form of a fractional two-compartment system is considered and then generalized to a system of an arbitrary number of compartments, which first appeared in [382]. A general ordinary linear two-compartment model is defined by the following system of linear ordinary differential equations

$$\begin{aligned} \dot{q}_1(t) &= -k_{12}q_1(t) + k_{21}q_2(t) - k_{10}q_1(t) + u_1(t) \\ \dot{q}_2(t) &= k_{12}q_1(t) - k_{21}q_2(t) - k_{20}q_2(t) + u_2(t) \end{aligned} \tag{9.5}$$

where $q_1(t)$ and $q_2(t)$ are the mass or molar amounts of material in the respective compartments and the k_{ij} constants control the mass transfer between the two compartments and elimination from each of them. The notation convention used for the indices of the rate constants is that the first corresponds to the source

compartment and the second to the target one, e.g., k_{12} corresponds to the transfer from compartment 1 to 2, k_{10} corresponds to the elimination from compartment 1, etc. The dimensions of all the k_{ij} rate constants are $(\text{time})^{-1}$. Input rates $u_i(t)$ in each compartment may be zero, constant, or time dependent. Initial values for $q_1(t)$ and $q_2(t)$ have to be considered also, $q_1(0)$ and $q_2(0)$, respectively.

In order to fractionalize this system, first the ordinary system is integrated, obtaining a system of integral equations and then the integrals are fractionalized as shown in [382]. Finally, the fractional integral equations are differentiated in an ordinary way. The resulting fractional system contains ordinary derivatives on the left-hand side and with Riemann–Liouville derivatives on the right-hand side

$$\dot{q}_1(t) = -k_{12f}\,{}_0D_t^{1-\alpha_{12}} q_1(t) + k_{21f}\,{}_0D_t^{1-\alpha_{21}} q_2(t) - k_{10f}\,{}_0D_t^{1-\alpha_{10}} q_1(t) + u_1(t)$$

$$\dot{q}_2(t) = k_{12f}\,{}_0D_t^{1-\alpha_{12}} q_1(t) - k_{21f}\,{}_0D_t^{1-\alpha_{21}} q_2(t) - k_{20f}\,{}_0D_t^{1-\alpha_{20}} q_2(t) + u_2(t)$$

where $0 < \alpha_{ij} < 1$ is a constant representing the order of the specific process. Different values for the orders of different processes may be considered, but the order of the corresponding terms of a specific process is kept the same when these appear in different equations, e.g., there can be an order α_{12} for the transfer from compartment 1 to 2 and a different order α_{21} for the transfer from compartment 2 to 1, but the order for the corresponding terms of the transfer, from compartment 1 to 2, α_{12}, is the same in both equations. Also the index "f" in the rate constant was added to emphasize the fact that these are different to the ones of 9.5 and carry dimensions $(\text{time})^{-\alpha}$.

Now, it is convenient to rewrite the above FDE system with Caputo derivatives. An FDE with Caputo derivatives accepts the usual type of initial conditions involving the variable itself, as opposed to Riemann–Liouville derivatives which involve an initial condition on the derivative of the variable, which is not practical. When the initial values are zero then the respective Riemann–Liouville and Caputo derivatives are the same. This is convenient since a zero initial value is very common in compartmental analysis. When the initial value is not zero, converting to a Caputo derivative is possible, for the particular term with a non-zero initial value. The conversion from a Riemann–Liouville to a Caputo derivative is done with the following expression

$${}_0D_t^{1-\alpha_{ij}} q_i(t) = {}_0^C D_t^{1-\alpha_{ij}} q_i(t) + \frac{q_i(0)\, t^{\alpha_{ij}-1}}{\Gamma(\alpha_{ij})}. \tag{9.6}$$

Summarizing about initial conditions, we can identify three cases (1°) the initial condition is zero and then the derivative becomes a Caputo by definition, (2°) the initial condition is non-zero but it is involved in a term with an ordinary derivative so it is treated as usual, (3°) the initial condition is non-zero and is involved in a fractional derivative which means that in order to present a Caputo derivative, an additional term, involving the initial value appears, by substituting the

relationship 9.6. Alternatively, a zero initial value for that variable can be assumed, with a Dirac delta input to account for the initial quantity for that variable.

So, the previous system of FDE with two compartments can be reformulated, by using as fractional derivatives the Caputo derivatives. Also, it is easy to generalize the above approach to a system with an arbitrary number of n compartments as follows

$$\dot{q}_i(t) = -k_{i0}\, {}_0^C D_t^{1-\alpha_{i0}} q_i(t) - \sum_{j\neq i}^{n} k_{ij}\, {}_0^C D_t^{1-\alpha_{ij}} q_i(t) + \sum_{j\neq i}^{n} k_{ji}\, {}_0^C D_t^{1-\alpha_{ji}} q_j(t) + u_i(t) \tag{9.7}$$

for $i = 1 : n$. Here Caputo derivatives have been considered throughout since, as explained above, this is feasible. The system of equations 9.7 is too general for most purposes as it allows every compartment to be connected with every other. Typically the connection matrix would be much sparser than that, with most compartments being connected to just one neighboring compartment while only a few "hub" compartments would have more than one connection.

The advantage of the described approach of fractionalization is that each transport process is fractionalized separately, rather than fractionalizing each compartment or each equation. Thus, processes of different fractional orders can coexist since they have consistent orders when the corresponding terms appear in different equations. Also, it is important to note that equation 9.7 does not have problems, such as violation of mass balance or inconsistencies with the units of the rate constants.

As mentioned, FDE can easily be written in the Laplace domain. In the case of FDE of the form of equation 9.7 where the fractional orders are $1 - \alpha_{ij}$, the relationship 9.2 becomes

$$L\left\{{}_0^C D_t^{1-\alpha_{ij}} q_i(t)\right\} = s^{1-\alpha_{ij}} \tilde{q}_i(s) - s^{-\alpha_{ij}} q_i(0)\,.$$

9.3 Examples of Fractional Models

9.3.1 One-compartment Model with Constant Rate Input

A one-compartment model with a fractional elimination and a constant rate input is considered [382]. Even in this simple one-compartment model, it is necessary to employ the approach of fractionalizing each process separately, described above, since the constant rate of infusion is not in fractional order. That would have been difficult if one followed the approach of changing the order of the derivative of the left-hand side of the ordinary differential equation, however here it is straightforward.

The system can be described by the following equation

$$\dot{q}(t) = k_{01} - k_{10f}\, {}_0^C D_t^{1-\alpha} q(t) \tag{9.8}$$

with $q(0) = 0$ and where k_{01} is a zero-order input rate constant, with dimensions (mass)/(time), k_{10f} is a rate constant with units (time)$^{-\alpha}$, and α is a fractional order less than 1. The previous equation can be written in the Laplace domain as

$$s\tilde{q}(s) - q(0) = \frac{k_{01}}{s} - k_{10f}\left[s^{1-\alpha}\tilde{q}(s) - s^{-\alpha}q(0)\right].$$

Since $q(0) = 0$, the above equation can be solved to obtain

$$\tilde{q}(s) = \frac{k_{01}s^{\alpha-2}}{s^{\alpha} + k_{10f}}.$$

By applying the following inverse Laplace transform formula (equation 1.80 in ref. [374], page 21)

$$L^{-1}\left\{\frac{s^{\alpha-\beta}}{s^{\alpha}+k}\right\} = t^{\beta-1}E_{\alpha,\beta}(-kt^{\alpha})$$

where $E_{\alpha,\beta}$ is the Mittag–Leffler function with two parameters. For $\beta = 2$ the following is obtained

$$q(t) = k_{01}tE_{\alpha,2}\left(-k_{10f}t^{\alpha}\right). \tag{9.9}$$

In theorem 1.4 of ref. [374], the following expansion for the Mittag–Leffler function is proven to hold for $|z| \to \infty$

$$E_{\alpha,\beta}(z) = -\sum_{i=1}^{p}\frac{z^{-i}}{\Gamma(\beta-\alpha)} + O\left(|z|^{-1-p}\right).$$

Applying this formula for relationship 9.9 and keeping only the first term of the sum since the rest are of higher order, the limit of 9.9 for t going to infinity can be calculated [382]

$$\lim_{t\to\infty}\{q(t)\} = \lim_{t\to\infty}\left\{k_{01}tE_{\alpha,2}\left(-k_{10f}t^{\alpha}\right)\right\} \approx \lim_{t\to\infty}\left\{\frac{k_{01}}{k_{10f}}\frac{t^{1-\alpha}}{\Gamma(2-\alpha)}\right\} = \infty,$$

for $\alpha < 1$. The fact that the limit of the relationship 9.9 diverges when t goes to infinity, for $\alpha < 1$, means that unlike the classic case, for $\alpha = 1$, where the relationship 9.9 approaches exponentially the steady state k_{01}/k_{10f}, for $\alpha < 1$, there is infinite accumulation. In Figure 9.2 left panel, a plot of relationship 9.9

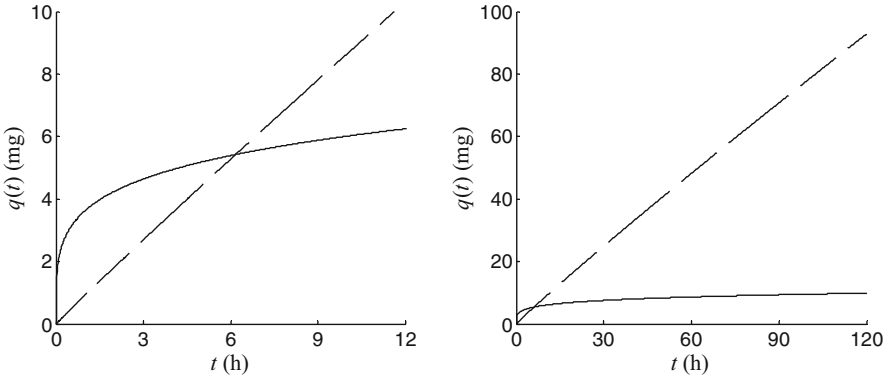


Fig. 9.2 For $\alpha = 0.25$, $k_{10f} = 0.1\ \mathrm{h}^{-\alpha}$, and $k_{01} = 1\ \mathrm{mg/l}$: Plot of relationship 9.9 with constant infusion (*dashed line*) which exhibits infinite accumulation and of relationship 9.10 with power-law infusion (*solid line*) which approaches a steady state. In the right panel, the profiles of left panel are for 10-times longer time span

is shown for $\alpha < 1$ showing that unlike the classic case where a steady state is approached, in the fractional case the amount keeps rising. In Figure 9.2 right panel, the same profiles are shown for 10-times larger time span, demonstrating the effect of continuous accumulation.

The lack of a steady state under constant rate administration which results to infinite accumulation is one of the most important clinical implications of the presence of fractional pharmacokinetics. It is clear that this implication extends to repeated doses as well as constant infusion, which is the most common dosing regimen, and can be important in chronic treatments. In order to avoid accumulation the constant rate administration must be adjusted to replaced by a decreasing with time rate. Indeed, in equation 9.8, if the constant rate of infusion k_{01} is replaced by the term $f(t) = k_{01} t^{-(1-\alpha)}$ [378], then the solution of the resulting FDE is, instead of relationship 9.9, the following

$$q(t) = k_{01} \Gamma(\alpha)\, t^{\alpha} E_{\alpha,\alpha+1}\left(-k_{10f} t^{\alpha}\right). \tag{9.10}$$

The previous solution converges to the steady state $\Gamma(\alpha)\, k_{01}/k_{10f}$ as time goes to infinity, while for the special case of $\alpha = 1$, the steady state is the usual k_{01}/k_{10f}. Similarly, for the case of repeated doses, if a steady state is intended to be achieved, in the presence of fractional elimination of order α, then the usual constant rate of administration, e.g., a constant daily dose, needs to be replaced by an appropriately decreasing rate of administration. As shown in [378], the decreasing rate of administration can be achieved by the following two approaches. The first approach uses the same dose given at increasing inter-dose intervals, i.e.,

$$T_i = \left(T_{i-1}^{\alpha} + \alpha \Delta \tau^{\alpha}\right)^{1/\alpha},$$

where T_i is the time of the i-th dose and $\Delta\tau$ is the inter-dose interval of the corresponding kinetics of order $\alpha = 1$. The second approach is based on a decreasing dose $q_{0,i}$ described by the following equation

$$q_{0,i} = \frac{q_0}{\alpha}\left[(i+1)^{\alpha} - i^{\alpha}\right].$$

given at constant time intervals. In this way an ever decreasing administration rate is implemented which compensates the decreasing elimination rate due to the fractional kinetics.

9.3.2 Two-Compartment Intravenous Model

Based on the generalized approach for the fractionalization of compartmental models, which allows mixing different fractional orders, developed in the previous section, a two-compartment fractional pharmacokinetic model is considered, shown schematically in Figure 9.3. Compartment 1 (central) represents general circulation and well-perfused tissues while compartment 2 (peripheral) represents deeper tissues. Three transfer processes (fluxes) are considered: elimination from the central compartment and a mass flux from the central to the peripheral compartment, which are both assumed to follow classic kinetics (order 1), while a flux from the peripheral to the central compartment is assumed to follow slower fractional kinetics accounting for tissue trapping.

The system is formulated mathematically as follows

$$\begin{aligned} \dot{q}_1(t) &= -(k_{10} + k_{12})\, q_1(t) + k_{21f}\, {}_0^C D_t^{1-\alpha} q_2(t) \\ \dot{q}_2(t) &= k_{12} q_1(t) - k_{21f}\, {}_0^C D_t^{1-\alpha} q_2(t) \end{aligned} \tag{9.11}$$

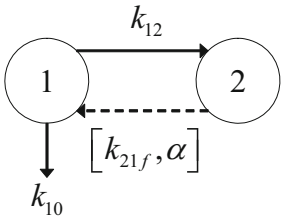


Fig. 9.3 A fractional 2-compartment pharmacokinetic model with an intravenous bolus. Elimination from the central compartment and a mass flux from the central to the peripheral compartment, which are both assumed to follow classic kinetics (order 1), while a flux from the peripheral to the central compartment is assumed to follow slower fractional kinetics, accounting for tissue trapping (*dashed arrow*)

with $\alpha < 1$ and initial conditions are $q_1(0) = q_0$, $q_2(0) = 0$ which account for a bolus dose injection and no initial amount in the peripheral compartment. Note that it is allowed to use Caputo derivatives here since the fractional derivatives involve only terms with $q_2(t)$ for which there is no initial amount, which means that Caputo and Riemann–Liouville derivatives are identical (relationship 9.6).

Applying the Laplace transform to the above system, the following algebraic equations are obtained

$$s\tilde{q}_1(s) - q_1(0) = -(k_{10} + k_{12})\tilde{q}_1(s) + k_{21f}\left[s^{1-\alpha}\tilde{q}_2(s) - s^{-\alpha}q_2(0)\right]$$
$$s\tilde{q}_2(s) - q_2(0) = k_{12}\tilde{q}_1(s) - k_{21f}\left[s^{1-\alpha}\tilde{q}_2(s) - s^{-\alpha}q_2(0)\right]$$

Solving for $\tilde{q}_1(s)$ and $\tilde{q}_2(s)$ and substituting the initial conditions

$$\tilde{q}_1(s) = \frac{q_0\left(s^\alpha + k_{21f}\right)}{(s + k_{12} + k_{10})\left(s^\alpha + k_{21f}\right) - k_{12}k_{21f}} \tag{9.12}$$

for the central compartment and

$$\tilde{q}_2(s) = \frac{q_0 s^{\alpha-1}k_{12}}{(s + k_{12} + k_{10})\left(s^\alpha + k_{21f}\right) - k_{12}k_{21f}} \tag{9.13}$$

for the peripheral compartment. The above expressions can be used by a NILT algorithm [379] to simulate values of $q_1(t)$ and $q_2(t)$ in the time domain.

Kilbas et al. [383] pointed out that the inverse Laplace transform of such relationships could lead to closed-form solutions. These solutions were derived in [384] to be

$$q_1(t) = q_0\sum_{n=0}^{\infty}(-1)^n k_{21f}^n \sum_{\lambda=0}^{n}\Omega E_{1,\lambda+\alpha n+1}^{n+1}\left[-(k_{10} + k_{12})t\right] +$$
$$q_0\sum_{n=0}^{\infty}(-1)^n k_{21f}^{n+1} \sum_{\lambda=0}^{n}\Omega t^\alpha E_{1,\lambda+\alpha(n+1)+1}^{n+1}\left[-(k_{10} + k_{12})t\right] \tag{9.14}$$

for relationship 9.12 and

$$q_2(t) = q_0 k_{12}\sum_{n=0}^{\infty}(-1)^n k_{21f}^n \sum_{\lambda=0}^{n}\Omega t E_{1,\lambda+\alpha n+2}^{n+1}\left[-(k_{10} + k_{12})t\right].$$

for relationship 9.13, with the common term

$$\Omega = \frac{k_{10}^\lambda n!}{(n-\lambda)!\lambda!}t^{\lambda+\alpha n}.$$

Nevertheless, these closed-form solutions involve terms with an infinite series of Mittag–Leffler functions and they are hard to implement and apply in practice.

Note, that primarily, $q_1(t)$ is of interest, since in practice, we only have data from this compartment. The time–amount profile $q_1(t)$ may be used to obtain the drug concentration in the blood $c(t)$ according to the elementary definition $c(t) \triangleq q_1(t)/V_1$ involving the distribution volume of the central compartment V_1. The so obtained model $c(t)$ can be fitted to the pharmacokinetic data in order to estimate parameters V_1, k_{10}, k_{12}, k_{21f}, and α.

9.4 Applications of Fractional Models

9.4.1 Amiodarone Pharmacokinetics

An application of the fractional two-compartment model, the system of 9.11 to amiodarone, was presented in [382]. Amiodarone is an antiarrhythmic drug known for its anomalous, nonexponential pharmacokinetics, which have important clinical implications due to the accumulation pattern of the drug in long-term administration. The fractional two-compartment model of the previous section was used to analyze an amiodarone intravenous data set which first appeared in [385] and estimates of the model parameters were obtained. The values for $q_1(t)$ were obtained from the expression of $\tilde{q}_1(s)$ in the Laplace domain (relationship 9.12) using a NILT algorithm [379]. In Figure 9.4 the model predicted values are plotted together with the data demonstrating good agreement for the 60 day period of this study. The estimated fractional order was $\alpha = 0.587$ and nonexponential character of the curve is evident, while the model follows well the data both for long and for short times, unlike empirical power laws which explode at $t = 0$.

9.4.2 Other Pharmacokinetic Applications

Apart from the amiodarone example, other applications of fractional pharmacokinetics have appeared in literature. From the authors Popovic et al. various applications of fractional pharmacokinetics to model drugs have appeared, namely for diclofenac [386], valproic acid [387], bumetanide [388], and methotrexate [389]. Also, Copot et al. have used a fractional pharmacokinetics model for propofol [390]. In most of these cases the fractional model has been compared with an equivalent ordinary pharmacokinetic model and has been found superior.

FDE have been proposed to describe drug response too, apart from their kinetics. Verotta has proposed several alternative fractional pharmacokinetic-dynamic models that are capable to describe pharmacodynamic times series with favorable properties [391]. Although these models are empirical, i.e., they have no mechanistic

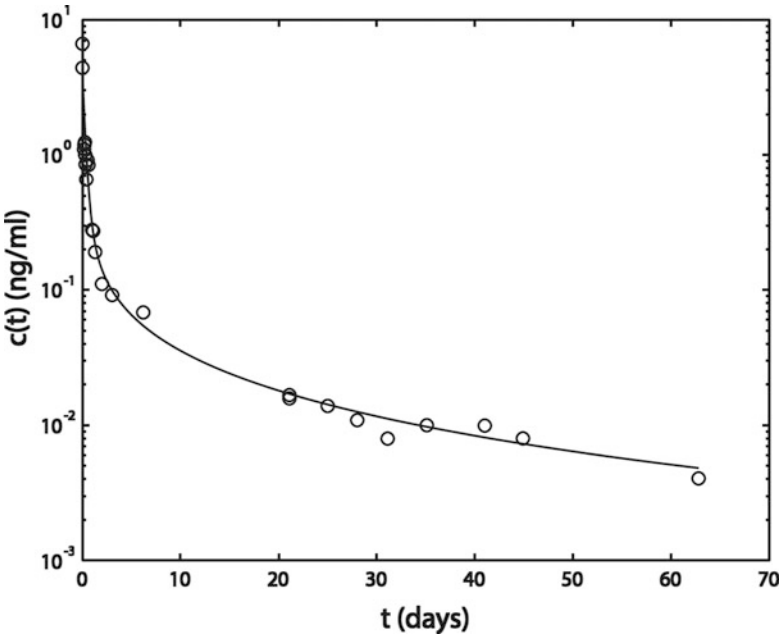


Fig. 9.4 The fractional two-compartment model of relationships 9.11. The fitted concentrations (line) to the amiodarone data (circles) were obtained from the expression $\tilde{q}_1(s)$ in the Laplace domain (relationship 9.12) using a NILT algorithm. Parameter estimates were: $k_{10} = 1.49\ \mathrm{d}^{-1}$, $k_{12} = 2.95\ \mathrm{d}^{-1}$, $k_{21f} = 0.48\ \mathrm{d}^{-\alpha}$, $\alpha = 0.58$, $q_0/V_1 = 4.72\ \mathrm{ng\ m}l^{-1}$

basis, they are attractive since the memory effects of the FDE can link smoothly the concentration to the response with a variable degree of influence, while the shape of the responses generated by fractional pharmacokinetic-dynamic models can be very flexible, with very few parameters.

Overall, fractional kinetics offers an elegant description of anomalous kinetics, i.e. nonexponential terminal phases, the presence of which has been acknowledged in pharmaceutical literature extensively. The approach offers simplicity and a valid scientific basis, since it has been applied in problems of diffusion in physics and biology. It introduces the Mittag–Leffler function which describes power-law behaved data well, in all time scales, unlike the empirical power laws which describe the data only for large times. Despite the mathematical difficulties, fractional pharmacokinetics is an interesting approach for the toolbox of the pharmaceutical scientist.

Chapter 10
Modeling and Simulation in Bioequivalence

There is nothing so unequal, as the equal treatment of unequals.
Aristotle (384–322 BC)

by Dr. V. KARALIS
Faculty of Pharmacy
National and Kapodistrian University of Athens

The purpose of bioequivalence testing is to assess the in vivo "equivalence" between two drug products of the same active moiety; namely, the test *T* and the innovator's formulation of the same active substance called as reference *R* product [237, 392]. In turn, proving bioequivalence is necessary in order to ensure therapeutic equivalence between the two products under comparison. Assessment of bioequivalence relies on the assumption that a product's therapeutic profile is a function of the concentration of the active substance in the effect site, which is dependent on drug's concentration in the general circulation. Thus, two drug products (*T* vs. *R*) are considered to be bioequivalent if their time–concentration profiles are sufficiently similar to ensure comparable clinical performance [393]. In other words, the main principle in bioequivalence testing is the proof of equivalence in pharmacokinetics which in turn is extrapolated to equivalence in therapeutics.

Formally, a *T* product is considered bioequivalent to the originator's product *R* if it contains the same active substance and when it is administered at the same molar dose, no significant differences are observed in the rate and extent of its absorption compared to the *R* formulation [237, 392]. In order to standardize and elucidate the bioequivalence assessment procedure, regulatory health agencies worldwide have issued dozens of guidelines in bioequivalence testing where special emphasis is placed on pharmacokinetic issues such as the type of pharmacokinetic parameters to be estimated, the statistical analysis of these pharmacokinetic measures, the selection of the appropriate clinical design which will allow the reliable estimation of the pharmacokinetic parameters, handling of missing drug concentrations, the use of parent drug or metabolite data, etc.

It becomes obvious that bioequivalence is a wide field having as cornerstones pharmacokinetics and regulatory aspects. However, an in depth analysis of bioequivalence cannot be accomplished without the acting role of statistics and the

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_10

use of computational approaches. For this reason, pharmacokinetic modeling and simulation have been intimately related to bioequivalence. It is quite common in the bioequivalence field that modeling and simulation studies which have been reported in the literature, actually triggered or served as the basis for clarifying or setting many significant issues of the regulatory guidelines. The significant contribution of modeling and simulation was greatly acknowledged by the European Medicines Agency (EMA), and a "Modeling and Simulation Working Group" has been set in the EMA. The aim of this group is to provide support to other committees such as the committee for medicinal products for human use, the pediatric committee, etc.

The use of modeling and simulation in bioequivalence assessment has been used in several aspects of bioequivalence assessment such as the: choice of the suitable pharmacokinetic metrics for expressing extent and rate of drug absorption, statistical framework, setting of the appropriate clinical design, sample size estimation, evaluation of pharmacokinetic equivalence in terms of the pharmacodynamic behavior, bioequivalence assessment in case of highly variable drugs, role of metabolites in the determination of bioequivalence, medicines interchangeability, and more recently modeling and simulation approaches for two-stage designs. In addition, modeling and simulation techniques are now officially being used for reasons of justifying biowaivers, the pharmacokinetic extrapolation of drugs' performance to pediatric populations, elucidating the methodology for the assessment of similar biopharmaceuticals (i.e., biosimilars), as well as to incorporate the concepts of phenotyping and genotyping in the bioequivalence field. Finally, some "typical" modeling and simulation methods, as the case of population pharmacokinetic (see Appendix C) modeling, have been explored for their applicability in bioequivalence. In this chapter, many of the above-mentioned cases will be presented.

10.1 General Methodology

This section starts with a rough description of the general concept of the modeling and simulation techniques used in bioequivalence [394–399]. The main idea is schematically shown in Figure 10.1 which actually reveals that the main elements required are the following:

- Random generation of a *virtual population sample* of subjects with specific pharmacokinetic (and in some cases pharmacodynamic) properties using Monte Carlo simulations.
- *All parameters of the virtual population of the volunteers are completely determined by the modeler*. In other words, one can adjust the pharmacokinetic properties of the drug (i.e., the pharmacokinetic model), the level of any type of variability (e.g., between-subject, within-subject, etc.), the absolute magnitude of the bioequivalence measure, the administered dose of the drug, the number of subjects, the sampling scheme, etc.
- More specifically for bioequivalence purposes, at least two virtual population samples are generated where each one refers to the different drug administration, namely, the T or R product.

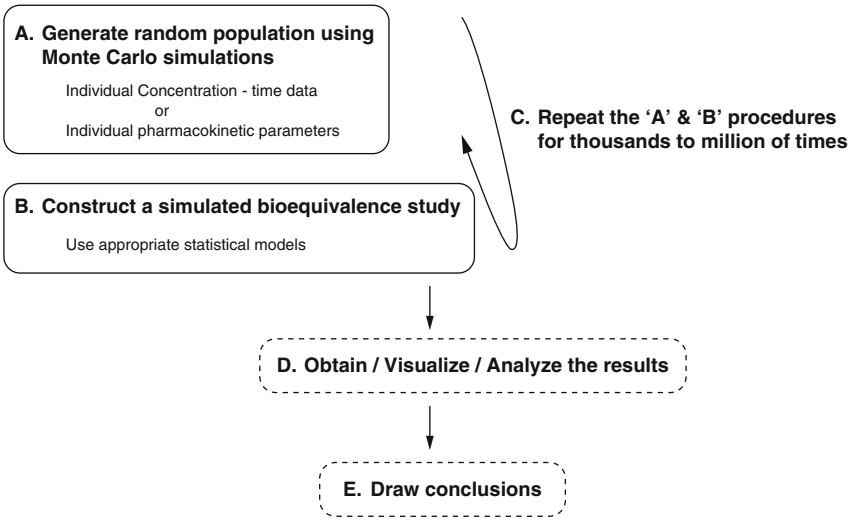


Fig. 10.1 A schematic representation of the classical modeling and simulation methodology used for bioequivalence purposes

- In a subsequent step, these subjects are appropriately classified into groups in order to *construct a certain clinical design* (e.g., the classic two-period, two-sequence, crossover design or more complex designs of any type).
- Pharmacokinetic estimations can be made by using the *typical pharmacokinetic methods* (e.g., noncompartmental analysis, Napierian ln transformation, etc.) or any other newly introduced method by the modeler.
- *Statistical analysis* is applied, using the appropriate bioequivalence criteria, and a decision regarding bioequivalence or not is made.
- The above-mentioned procedure is *repeated for many times* (in nowadays, usually more than $100,000$ to millions iterations).
- The value of a *factor* under study (e.g., the absorption rate constant) is changed and the entire process is repeated again.
- Finally, one can get information about the performance of bioequivalence under many different situations (scenarios). This allows drawing conclusions on many different pharmacokinetic conditions.

10.2 Examples of Modeling and Simulation

The purpose of this section is to present examples and illustrate how modeling and simulation techniques have been used to elucidate some unresolved issues or introduce new ideas in bioequivalence assessment.

10.2.1 Bioequivalence Measures

In the past, one of the first fields, where modeling and simulation methods were applied to bioequivalence testing, was the choice of the pharmacokinetic parameters which should be used in statistical evaluation. In order to answer this query one should recall the basic aims of the bioequivalence appraisal. The scientific basis for the bioequivalence assessment is based on the pharmacokinetic comparison, namely, the *rate* and *extent* of absorption between *T* and *R* [394, 395]. Currently, it is widely accepted that in case of immediate release formulations the rate of absorption should be assessed by the peak plasma concentration c_{max}, whereas the extent of absorption is expressed by the area under the time–concentration curve *AUC* from time zero to the last sampling point or the last measurable concentration, whichever occurs earlier [394, 395, 400, 401]. In addition, other pharmacokinetic measures are further used to provide supplementary information. These include the area under the plasma time–concentration curve extrapolated to infinity AUC_{∞}, the time t_{max} at which c_{max} occurs, and the terminal slope of the time–concentration curve. For modified release products, other pharmacokinetic parameters are used like the maximum and the minimum drug concentration at the steady state, the *AUC* between each interval of administration, the peak–trough fluctuation [402, 403]. In all cases, the areas are calculated using noncompartmental approaches, namely using the linear- or log-trapezoidal rule (Figure 10.2).

However, before concluding on the above-mentioned bioequivalence measures several other metrics had been proposed in the literature and evaluated. In all cases, the investigation of the potential of a pharmacokinetic metric to be established as a bioequivalence measure was based on computational approaches. Most of the research in this field took place during the last two decades of the 20th century.

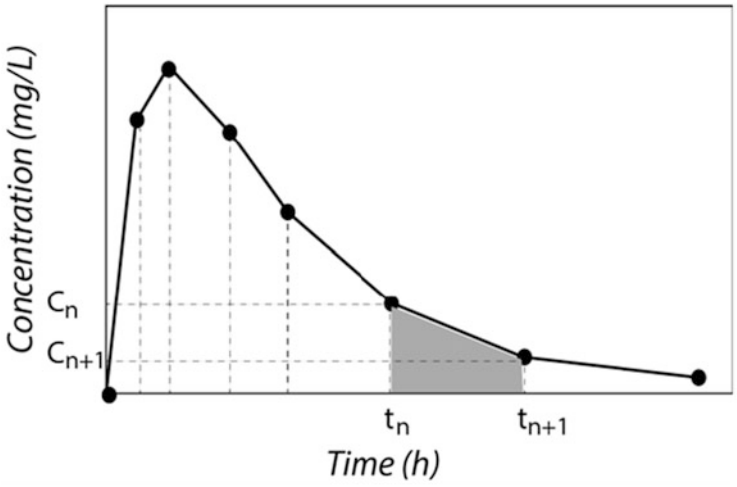


Fig. 10.2 A non-compartmental approach to calculate the area under the time–concentration curve: the application of the linear trapezoidal method

Even though, at that time the computational methodology was at its first stages of development, the modeling and simulation techniques were in essence the only available tool for investigation.

Several articles have appeared in the literature adopting naïve or more advanced computational methodologies. In this vein, c_{max} was criticized as a metric which (apart from the rate of absorption) also express the extent of absorption [395, 404, 405]. Other studies had also raised similar concerns about the use of c_{max} as a metric of the rate of absorption [406–412]. Several pharmacokinetic metrics were proposed in order to resolve some drawbacks of the pre-referred bioequivalence measures. For example, the performance of the c_{max}/AUC ratio was analyzed using modeling and simulation techniques. The background rationale of this normalization was rather simple; since c_{max} is influenced by the extent of absorption, normalization by AUC would create a metric which is independent from the degree of absorption.

Meanwhile, scientists used modeling and simulation techniques to introduce a completely different approach for comparing the time–concentration profiles of the T and R products. This procedure, which named as "direct curve comparison" metrics, aimed at quantifying the (dis)similarity of the two time–concentration profiles based on a point-by-point comparison [413–416]. The proposed direct curve comparison metrics are the following:

1. Rescigno indices (ξ_1 and ξ_2)

$$\xi_i = \left[\frac{\int_0^\infty |R_i - T_i|^i \, dt}{\int_0^\infty |R_i + T_i|^i \, dt} \right]^{1/i}$$

   where i is an integer positive number (1 and 2 in case of ξ_1 and ξ_2, respectively), whereas R_i and T_i refer to the plasma concentrations of the *test* and *reference* product at the i-th time point, respectively.
2. Difference factor f_1 defined by the relationship 5.29. This metric was originally proposed for the comparison of two in vitro dissolution curves [413]. The f_1 metric reflects the mean difference by taking into account all data from $i = 1$ to n.
3. Similarity factor f_2 defined by the relationship 5.30.

It should be emphasized here that the modeling and simulation techniques, for the investigation of the suitable bioequivalence measure, are centered around the concept of drug *exposure* [397, 400, 406]. According to the official EMA regulatory guideline [237], AUC is used to express the *extent of exposure*, whereas the peak exposure is reflected on the maximum observed plasma concentration (i.e., c_{max}).

10.2.2 Statistical Framework

Bioequivalence assessment is, in essence, a comparison between the pharmacokinetic performances of two drug products. Thus, at the initial setting of bioequivalence, there was a need to establish a specific and complete statistical framework for the comparison.

10.2.2.1 Seeking for the Appropriate Hypothesis

The first issue that was necessary to be clarified was the choice of the appropriate statistical hypothesis. In this context, several statistical approaches have been proposed [417–420].

Firstly, the so-called *power approach* was proposed which assumes a nominal hypothesis of bioequivalence or equivalently of no difference between the two drug products

$$\mathcal{H}_0^{\bullet} : \mu_T - \mu_R = 0 \text{ vs. } \mathcal{H}_1^{\bullet} : \mu_T - \mu_R \neq 0$$

where μ_T and μ_R refer to the average bioequivalence measures (in the ln-domain) of T and R product, respectively.

The power approach is an ad hoc method of testing the interval $\mathcal{H}_0^{\bullet}$ according to which the statistical analysis can be carried out by using the standard two-sided t-test methodology. If $\mathcal{H}_0^{\bullet}$ is rejected, then it cannot be concluded that the T and R products are bioequivalent. But, when $\mathcal{H}_0^{\bullet}$ is not rejected, then the assessment continues by assessing if the power is higher than 80%. A graphical illustration of the rejection region of the power approach is depicted in Figure 10.3A.

An alternative approach was proposed by Hauck and Anderson [417]. In this case, two-interval hypotheses are defined

$$\mathcal{H}_0 : \mu_T - \mu_R \leq -\delta \text{ or } \mathcal{H}_0 : \mu_T - \mu_R \geq +\delta$$

vs.

$$\mathcal{H}_1 : -\delta < \mu_T - \mu_R < +\delta$$

where δ is the predefined acceptance limit (a positive number); usually set at ln (1.25).

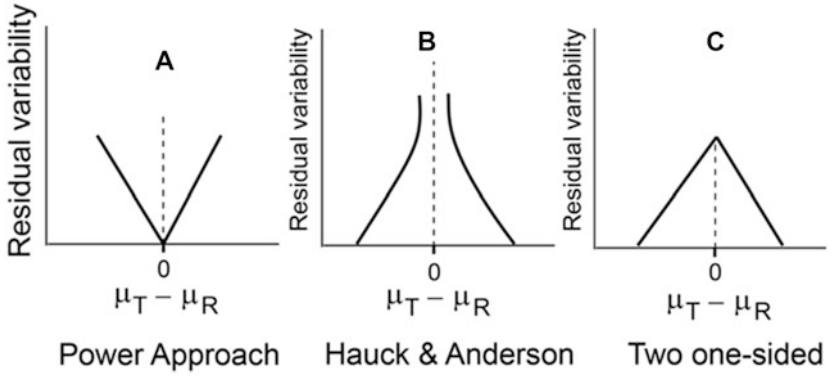


Fig. 10.3 An oversimplified graphical illustration of the rejection regions of the three different bioequivalence hypotheses: (**A**) the power approach, (**B**) the Hauck and Anderson method, and (**C**) the two one-sided test procedure. In all cases, the difference (in the ln-domain) between the mean measures of bioequivalence of the two products under comparison is plotted versus residual variability

In other words, according to the Hauck and Anderson approach, the nominal hypothesis refers to "no bioequivalence,"while the alternative hypothesis assumes bioequivalence. This setting allowed the specification of an upper limit of the type I error (i.e., false positive). This implies that the erroneous determination of bioequivalence or equivalently the patient risk cannot be higher than the nominal level of 5%. Figure 10.3B shows the rejection region of the "Hauck and Anderson approach.".

Three years later, a modified approach for the hypothesis was proposed by D. Schuirmann who introduced the *two one-sided test procedure* (TOST) [418]. The nominal and the alternative hypothesis, of the two sets of one-sided hypotheses, are the following:

$$\mathcal{H}_{01} : \mu_T - \mu_R \leq -\delta \quad \mathcal{H}_{11} : \mu_T - \mu_R > -\delta$$

and

$$\mathcal{H}_{02} : \mu_T - \mu_R \geq \delta \quad \mathcal{H}_{12} : \mu_T - \mu_R < \delta$$

The TOST procedure relies on the decomposition of the interval hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ into two sets of one-sided hypotheses. The TOST procedure actually states that the rejection of the interval hypothesis $\mathcal{H}_0$, and thus concluding equivalence of μ_T and μ_R, can take place if and only if both $\mathcal{H}_{01}$ and $\mathcal{H}_{02}$ are rejected at a pre-specified nominal level of significance α. In other words, bioequivalence can be denoted when both $\mathcal{H}_{11}$ and $\mathcal{H}_{12}$ hold.

Assuming normal distribution, the two sets of one-sided hypotheses can be tested with ordinary one-sided t-tests. Thus, the TOST method is actually identical to the procedure of declaring equivalence based on the $(1 - 2\alpha)$ % confidence interval for $\mu_T - \mu_R$; namely, if it lies completely within some pre-specified limits (an upper and a lower limit of acceptance). Figure 10.3C depicts schematically the rejection region of the TOST procedure.

It is worth mentioning that the approaches presented above which were published $25 - 30$ years ago, were published under the title "Pharmacometrics,"which is now a popular field of application of the modeling and simulation techniques [417, 418].

10.2.2.2 Seeking for the "Limits" and the "Methodology"

Even though the construction of an appropriate hypothesis setting is necessary, it is not sufficient for the purposes of bioequivalence assessment. A complete and detailed description of the statistical methodology is required, as well as, the choice of the acceptance limits.

As far as the limits of acceptance, a fixed $\pm 20\%$ boundary has been set as a universal guideline. A strict justification for this value has not appeared in the literature, but this value was adopted since a 20% difference was considered as being non-significant in terms of mean clinical effect. Therefore, it seemed reasonable as a limit for product quality [421].

Also, all calculations take place in the Napierian logarithm domain of the pharmacokinetic data (e.g., AUC, $c_{\max}$). The reason for the ln-transformation is two-fold (1°) a clinical rationale which suggests that the comparison should focus on the ratio rather than the difference between the mean values of the pharmacokinetic parameters and (2°) a pharmacokinetic/statistical rationale due to the fact that a multiplicative model is used [422–424]. The latter, properly coupled with the $\pm 20\%$ difference, leads to the 80–125% rule.

Thus, according to the general modeling and simulation procedure, depicted in Figure 10.1, the pharmacokinetic parameters are calculated for each subject and then they are transformed into the ln-domain. The statistical analysis takes place afterwards, where the mean sample values of the T and R pharmacokinetic parameters are calculated, as well as, the residual variability. Based on the TOST procedure, the $(1-2\alpha)$ % confidence interval is estimated in the normal domain and compared to the pre-set acceptance limits. Typically, a 90% two-sided confidence interval is used, since the nominal level of significance is set at 5%. Thus, if the 90% confidence interval of the mean difference (of the pharmacokinetic parameters of T and R) is within the 80–125% limits, then bioequivalence is declared.

It will be described later on in this chapter, that in some cases (e.g., highly variable drugs) the constant 80–125% limits, outlined above, do not suffice, and other limits are specified.

Average Bioequivalence The assessment procedure using the $(1-2\alpha)$ % confidence interval, which is described above, actually implies the assumption of *average bioequivalence* [392, 422]. Indeed, average bioequivalence refers to equivalence in averages of the marginal distributions of the pharmacokinetic parameters of the T and R products. As it is already mentioned, this approach is equivalent to the two one-sided t-tests [418].

In order to explain the encountered variability, the statistical evaluation with the average bioequivalence approach includes the following "effects": Sequence, Formulation, Period, and the nested effect Subject-within-Sequence. The variability that cannot be described by the pre-referred factors is characterized as "residual." In case of the classic 2×2 design, the following general linear model is used in the average bioequivalence procedure

$$Y_{ijk} : \mu + S_{ik} + T_j + P_{(j,k)} + e_{ijk} \tag{10.1}$$

where

- Y_{ijk} is the pharmacokinetic response (ln-transformed) of the i-th subject in the k-th sequence for the j-th formulation,
- μ is the overall mean,
- T_j refers to the effect for the j-th treatment (i.e., T or R),
- $P_{(j,k)}$ is the period effect at which the j-th formulation is administered at the k-th sequence,
- S_{ik} is the effect of the i-th subject in the k-th sequence,

- e_{ijk} refers to the residual variability in observing Y_{ijk}. The latter is thought to express the within-subject variability.

Based on the regulatory guidelines, the T_j and $P_{(j,k)}$ effects are set as fixed. In the past, the S_{ik} effect was considered as random, but currently the S_{ik} is also considered as fixed [237]. It should be mentioned that in the model described by the relationship 10.1, no carryover effects are assumed, since an adequate washout period is supposed to be included between the two treatment periods.

The average bioequivalence approach relies on the use of fixed bioequivalence limits which are defined by the regulatory authorities [237, 392]. The criterion applied to the determination of bioequivalence is mathematically expressed by

$$-\ln\delta \leq \mu_T - \mu_R \leq \ln\delta$$

where $\ln\delta$ is usually set at $\ln(1.25)$.

Assuming the classic two-treatment, two-period, two-sequence (2×2) crossover design, with equal number subjects in each sequence, the upper and lower limits of the 90% confidence interval are given by [418, 425]

$$Up, Low\ 90\%\ CI = \exp\left[(\mu_T - \mu_R) \pm t_{0.05,N-2}\sqrt{MSE\frac{2}{N}}\right] \tag{10.2}$$

where MSE is the mean square error derived from the linear model described by the relationship 10.1, t is the t-student statistic with $N-2$ degrees of freedom, and N is the total number of subjects enrolled in the bioequivalence study. It is reminded that the terms μ_T and μ_R refer to the ln-transformed mean pharmacokinetic metrics of T and R products, respectively.

Overall, it can be concluded that the average bioequivalence approach is a conventional method which focuses only on the sample average pharmacokinetic parameters and does not make any comparisons of the products' variances. Also, average bioequivalence does not assess subject-by-formulation interaction and ignores the statistical distribution of the metric. For these reasons, it has been characterized and concomitantly criticized as an "one size fit all bioequivalence criterion."

Other Approaches It should be highlighted that even though the application of average bioequivalence can ensure efficacy and safety of the two products under comparison, it cannot address questions like *prescribability* and *switchability*. In order to deal with these issues, the population and individual bioequivalence approaches were introduced [143, 426–435]. The term prescribability refers to the physician's selection for the appropriate drug product for a new patient. This choice is made among the R product and a number of T formulations that have already been proved to be bioequivalent to the R. Another important term is switchability. The latter refers to situation when a *switch* is made from the R product to a T formulation within the same patient.

In order to address the issues of prescribability and switchability two different methodological procedures have been proposed (1°) the population bioequivalence and (2°) the individual bioequivalence.

Population Bioequivalence Approach The concept of population bioequivalence was introduced in order to offer an answer to the issue of prescribability [426–428]. Population bioequivalence requires the knowledge of the total variability of the sample of volunteers, that is, the sum of the between- and the within-subject variabilities. In order to address population bioequivalence two methodological options have been proposed: a reference-scaled and constant-scaled criterion. The reference-scaled approach is mathematically expressed by

$$\frac{(\mu_T - \mu_R)^2 + \left(S_{TT}^2 - S_{TR}^2\right)}{S_{TR}^2} \leq \theta_P$$

while the constant-scaled approach by

$$\frac{(\mu_T - \mu_R)^2 + \left(S_{TT}^2 - S_{TR}^2\right)}{S_{PO}^2} \leq \theta_P$$

where S_{TT}^2 is the total (i.e., within- and between-subject) variability of the T formulation; S_{TR}^2 is the total (i.e., within- and between-subject) variability of the R formulation; S_{PO}^2 is a constant variance term; θ_P is the limit of acceptance for population bioequivalence. The choice between one or the other method depends on the magnitudes of S_{TR} and S_{PO}. The reference-scaled method should be used when $S_{TR} > S_{PO}$, while the constant-scaled criterion is applied if $S_{TR} \leq S_{PO}$ [422].

Individual Bioequivalence Approach As it is mentioned above, individual bioequivalence was introduced to address switchability. The application of individual bioequivalence presupposes the estimation of the within-subject variabilities of the T and R formulation, as well as an additional variance term called as "subject-by-formulation" interaction. As in case of population bioequivalence, two different methodologies have been proposed: a reference- and a constant-scaled criterion. The individual bioequivalence can be expressed by the inequality

$$\frac{(\mu_T - \mu_R)^2 + \left(S_{wT}^2 - S_{wR}^2\right) + S_D^2}{S_{wR}^2} \leq \theta_I \tag{10.3}$$

for the reference-scaled method and

$$\frac{(\mu_T - \mu_R)^2 + \left(S_{wT}^2 - S_{wR}^2\right) + S_D^2}{S_{IO}^2} \leq \theta_I \tag{10.4}$$

for the constant-scaled method. In the above expressions, S_{wT}^2 is the within-subject variability of the T formulation; S_{wR}^2 is the within-subject variability of the R

formulation; S^2_{IO} is a constant within-subject variance term used in individual bioequivalence; S^2_D is the variance component of subject-by-formulation interaction; θ_I is the limit of acceptance for individual bioequivalence. The choice between the use of the reference- or the constant-scaled criterion is based on the relative values of S_{wR} and S_{IO}. Thus, the reference-scaled criterion should be applied if $S_{wR} > S_{IO}$, whereas the constant-scaled method should be used when $S_{wR} \leq S_{IO}$ [422, 429]. Plausibly, the individual bioequivalence approach requires the conduct of a replicate study, which will allow the estimation of the within-subject variability of both *T* and *R* products.

Despite the fact that the individual bioequivalence approach appeared rather attractive, it gained only limited application due to practical problems [434]. More specifically, the FDA proposed the use of non-parametric bootstrap 95% upper confidence interval, namely, a method that requires computational skills. The same method also proposed for population bioequivalence. In order to resolve this computational difficulty an alternative confidence interval procedure was proposed. The latter was based on the Howe's approximation I to a Cornish–Fisher expansion [436]. Based on this method, inequalities 10.3 and 10.4 are re-written in terms of a linear combination of parameters, namely

$$\eta_1 = (\mu_T - \mu_R)^2 + S^2_D + \left(S^2_{wT} - S^2_{wR}\right) - \theta_I S^2_{wR} < 0 \qquad \text{for} \qquad S_{wR} > S_{IO}$$

$$\eta_2 = (\mu_T - \mu_R)^2 + S^2_D + \left(S^2_{wT} - S^2_{wR}\right) - \theta_I S^2_{IO} < 0 \qquad \text{for} \qquad S_{wR} \leq S_{IO}$$

Thus, if one wants a testing methodology for individual bioequivalence, he can estimate the upper 95% confidence interval for the η_1 and η_2 in the above inequalities. This approach can be applied to balanced or unbalanced crossover designs, and can be computed very easily.

10.2.3 Highly Variable Drugs

Even though the above presented classic average bioequivalence approach is widely used, it does not suffice for the case of highly variable drugs or drug products. For the purposes of this chapter the term "highly variable drugs" will be used to refer to cases where the observed coefficient of variation of the within-subject is greater than or equal to 30%, regardless if it is due to the drug substance itself or derives from the product properties [237, 437, 438]. Many causes such as the drug-related issues of the active moiety and physiological/pathological conditions of the human body may contribute significantly to variability [438]. Irrespective of the underlying reason, as within-subject variability increases, it becomes harder to prove bioequivalence even though it exists. From the relationship 10.2, it becomes apparent that as variability (i.e., *MSE*) increases, then it becomes more difficult to declare bioequivalence, unless a larger sample size is used. In order to resolve

this problem, several approaches were proposed like the conduct of multiple-dose studies, the planning of bioequivalence studies in the fed state, as well as the use of individual or population bioequivalence criteria [427, 429, 432–434, 439].

Closely to the end of the 20th century, the idea of *scaled bioequivalence approaches* was invented [440]. Based on this approach, the bioequivalence limits are expanded as a function of the within-subject variability of the reference product

$$Up, Low \text{ bioequiv. } limits = \exp\left(\pm k_{sc} S_{wR}\right)$$

where k_{sc} is a scaling factor and S_{wR} is the within-subject variability. Since then, several other approaches have been introduced which ameliorate the properties of the scaling rationale [262, 264, 425, 441–443]. Thus, the word "scaled" characterizes the bioequivalence acceptance limits which are gradually becoming wider based on estimates of the within-subject variability of the study. Figure 10.4 illustrates several scaling methods either proposed in the literature, or currently accepted by the regulatory agencies.

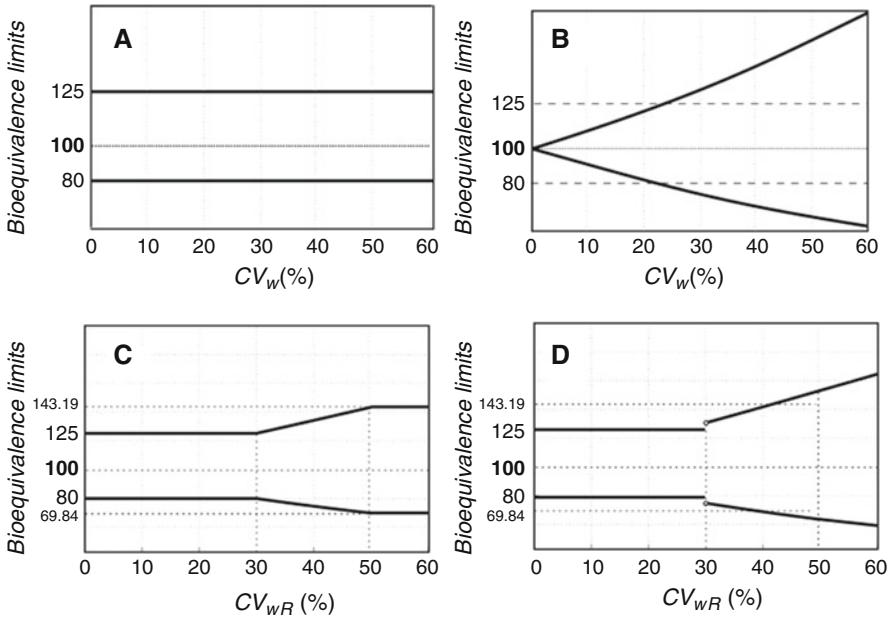


Fig. 10.4 Bioequivalence acceptance limits as a function of the "unexplained" variability. (**A**) and (**B**): x—axis refers to the "pooled" Test-Reference residual variability (actually, its coefficient of variation CV_w), (**C**) and (**D**): x—axis is the calculated within-subject variability of the reference product (the coefficient of variation CV_{wR}). (**A**) uses the classic 80–125% limits, (**B**) is the first scaled approach proposed in 1995, (**C**) refers to the mixed scaled criterion adopted by the EMA, and (**D**) is the mixed scaled proposed by the US FDA

In Europe, bearing in mind all scaled procedures, a novel methodology was adopted by the EMA [237, 444]. Currently, the EMA suggests the application of a *mixed scaled approach* which relies on S_{wR}. According to this method, it is recommended that if the coefficient of variation of S_{wR} (i.e., CV_{wR}) is less than (or equal to) 30% or higher (or equal to) than 50%, then the classic 80–125% limits or the expanded 69.84–143.19% limits should be applied, respectively. For CV_{wR} values between 30 and 50%, the scaled limits should be used. The EMA scaled limits can be expressed mathematically by

$$Up, Low \text{ bioequiv. } limits = \begin{cases} 0.80, 1.25 & \text{for} \quad CV_{wR} \leq 30\% \\ \exp\left(\pm k_{sc}^{\cdot} S_{wR}\right) & \text{for } 30\% < CV_{wR} < 50\% \\ 0.6984, 1.4319 & \text{for} \quad 50\% \leq CV_{wR} \end{cases}$$

where $k_{sc}^{\cdot}$ is a scaling factor ($k_{sc}^{\cdot} = 0.760$) set by the EMA for the scaled approach. In addition, a secondary criterion is used in all cases. The latter defines that the point estimate of the geometric mean ratio of the pharmacokinetic parameters should be between the 80–125% acceptance interval. This constraint is used to avoid the risk of accepting two drug products which in fact may differ significantly in their geometric-mean-ratio values. It should be stated that the scaling procedure applies only to $c_{\max}$, whereas the acceptance limits of AUC are kept equal to the typical levels (80–125%), regardless the level of variability.

A similar criterion, but not completely identical to that proposed by the EMA, was introduced by the US FDA [443, 445]. Based on the FDA method, a *reference-scaled approach* is used for drugs with $CV_{wR} \geq 30\%$. These scaled limits were mathematically expressed by

$$Up, Low \text{ bioequiv. } limits = \begin{cases} 0.80, 1.25 & \text{for } CV_{wR} < 30\% \\ \exp\left[\pm \ln(1.25) \frac{S_{wR}}{S_{w0}}\right] & \text{for } 30\% \leq CV_{wR} \end{cases}$$

where S_{w0} is a constant (set equal to 0.25) referring to the regulatory standardized variation of the FDA approach. Therefore, the actual "constant" value next to S_{wR} is $\ln(1.25)/0.25 \approx 0.893$. In any case, for drugs with CV_{wR} less than 30%, the typical 80–125% limits are applied. Finally, the point geometric mean of T/R ratio estimate of the pharmacokinetic parameter should lie in the region of 80–125%.

It is worth mentioning that the reference-scaled procedures recommended either by the EMA or the FDA, imply the use of full-replicate or a semi-replicate design [443, 445], where at least the R product is administered twice to each subject. The use of replicate designs is essential in order to estimate the within-subject variability of the R product. A more detailed discussion on the replicate designs is provided later in the chapter.

Apart from the similarities between the EMA and the FDA scaling methods, several differences also exist [446, 447]. Firstly, the FDA uses a different (i.e., higher) scaled constant than the EMA. In addition, there is no upper cut-off limit for FDA, as in case of the EMA method (Figure 10.4C). The latter leads to

continuous expansion of the bioequivalence limits (Figure 10.4D). Plausibly, these two differences imply that the FDA method is more permissive than EMA. This attitude becomes even more evident as CV_{wR} increases. Besides, a finding from the modeling and simulation methods revealed that a sample size increase affects more the EMA limits than the FDA [446]. In essence, these two approaches can only lead to identical results either when CV_{wR} is less than 30% or when a large number of subjects are recruited.

It should be underlined that the development of all scaled bioequivalence approaches was based entirely on modeling and simulation techniques. All the above-mentioned articles, dealing with scaling bioequivalence approaches, relied solely on modeling and simulation techniques. The application of modeling and simulation techniques allowed the study of many different scenarios; thus, the properties as well as the pros and cons of each scaled procedure were evaluated. Also, modeling and simulation allowed the comparison of the properties of the EMA's and the FDA's approaches [446, 447]. The common modeling and simulation methodology for this type of analysis is based on the following combination of factors: simulation of bioequivalence studies (see Figure 10.1) under several different conditions of variability, sample size, type of clinical design, similarity in the average pharmacokinetic parameter values, and any setting for certain parameters of the scaling method. The main "tool" for the assessment of each method was the construction of plots called as "power curves."The power curves plots show the percentage of bioequivalence acceptance as a function of a variable, usually the geometric mean ratio of T/R ratio for a certain pharmacokinetic parameter. A typical example of a power curve plot is shown in Figure 10.5.

The main advantage of power curves relies on the fact that they depict the overall performance of the utilized method for several different conditions which are controlled by the modeler. Every point in the "power curves" plot refers to the average percent acceptance value derived from thousands to millions of iterations of a bioequivalence study. Thus, allowing the extraction of robust estimates. Each bioequivalence study is generated using Monte Carlo methods for the random

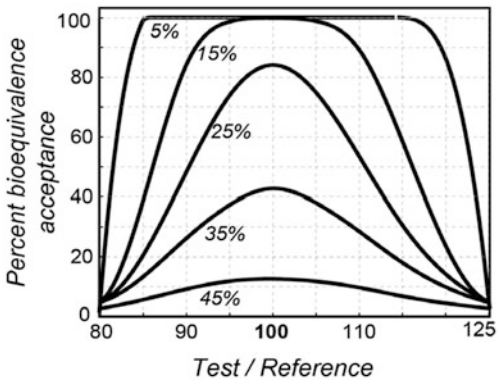


Fig. 10.5 Percent of bioequivalence acceptance (y—axis) of the test vs. the reference product for several values of their geometric mean ratios (x—axis). The % numbers, close to the curves refer to the coefficient of variation of the residual variability. In this case, the geometric mean T/R ratio ranges from 80 to 125%

generation of the pharmacokinetic parameters. Finally, the statistical methodology presented above is applied to each one of the bioequivalence studies.

10.2.4 Clinical Design

The main purpose of the clinical design of a bioequivalence study is to distinguish the formulation effect from other factors that can also affect the outcome of the study. Thus, the study design is primarily dependent on the pharmacokinetic properties of the active substance. A valuable tool for the choice of the appropriate clinical design is the use of modeling and simulation techniques which allow the investigation of the performance of each design under the anticipated conditions of the planned bioequivalence study. The modeling and simulation techniques offer the opportunity to test the candidate designs and choose the one which offers the optimum performance of the study, e.g., higher statistical power, estimation of within-subject variability, avoidance of inflation of type I error, etc.

10.2.4.1 Standard Design

Traditionally, the randomized, two-treatment, two-sequence, two-period (2×2) design (Figure 10.6) is used in bioequivalence assessment [237, 392]. Based on the 2×2 design, each subject receives consecutively the two drug products (i.e., *T* and *R*) in two different time phases, which are separated by an adequate washout period. The duration of the washout period is calculated using the elimination half-life of the drug. The latter should be adequately large (e.g., more than five times the elimination half-life) to ensure that no drug can be detected at the initiation of the second period of the study.

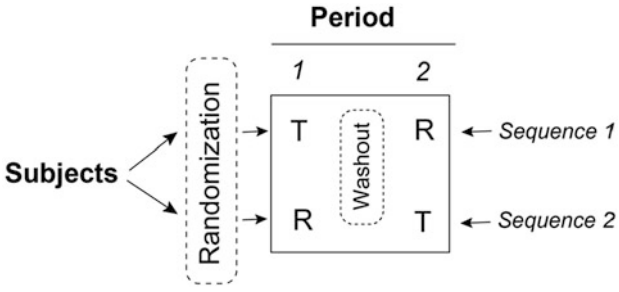


Fig. 10.6 The classic two-treatment, two-period, two-sequence, crossover design. Between the two dosing administrations an adequate washout period should exist in order to avoid carry-over effects. The terms *T* and *R* refer to the Test and the Reference products, respectively

It should be underlined that due to the nature of this design, the within-subject variability of any of the two products (i.e., T or R) cannot be estimated. After applying the linear model (expression 10.1), the estimated residual error is actually an estimate of the pooled variability from the two drug products, which is generally considered (even though it is not always true) to reflect the within-subject variability of the active substance.

10.2.4.2 Alternative Designs

Even though the 2×2 design has been widely used, other possibilities for clinical setting are also accepted such as replicate and two-stage designs.

Replicate Designs In replicate designs, each participant receives the same product more than once. The replicate administration may refer to both formulations or to only one of them. Some examples of replicate designs are shown in Figure 10.7. One advantage of replicate designs is the fact that they allow the calculation of the within-subject variability of a drug product, if this product is administered more than once to the same subject. This implies a very useful application to the case of highly variable drugs, where the knowledge of S_{wR} is needed. Thus, in case of highly variable drugs, it is proposed the use of three- or four-period studies (Figure 10.7) where, at least, the R product is administered twice [237, 443, 445]. In case of Figure 10.7C, each subject receives both T and R products twice, which leads to four periods of administration. Plausibly, this type of design allows the estimation of the within-subject variabilities of the two treatments T and R. On the contrary, replication may only refer to one treatment (i.e., R), which in turn allows the estimation of S_{wR} (e.g., Figure 10.7A and 10.7B).

A common advantage of replicate studies relies on their properties to reduce the required number of subjects [392, 439, 448]. For example, a four-period design requires almost 50% of the subjects of a typical 2×2 study. Plausibly, replicate designs should also be used in case of the individual bioequivalence approach, but they are not mandatory for either the average or the population bioequivalence approach [422]. As mentioned above, the assessment of several replicate designs in bioequivalence assessment has been made using modeling and simulation techniques. Nevertheless, the regulatory authorities do not promote or suggest a specific replicate design.

Two-Stage Designs Two-stage designs are a recent possibility for clinical design in bioequivalence studies [237, 449–451]. The general rationale of two-stage designs methodology is the following: initially (stage 1), a group of subjects is recruited and their pharmacokinetic data are analyzed using the statistical principles discussed earlier in this chapter. If some predefined criteria are satisfied in stage 1 (e.g., bioequivalence is proved), the study stops since there is no need to proceed into stage 2. Otherwise, additional subjects can be included and the study proceeds into stage 2 [452]. Finally, data from both stages are combined and appropriately analyzed. A general schematic presentation of two-stage designs is shown in Figure 10.8.

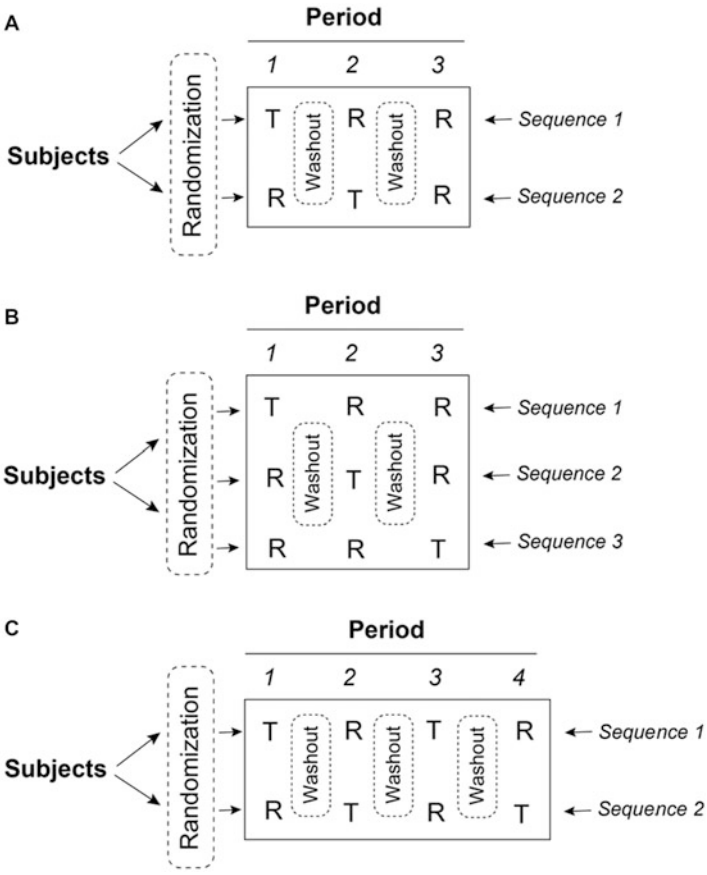


Fig. 10.7 Representative examples of replicate crossover designs used in bioequivalence assessment (**A**) a two-sequence, three-period design, (**B**) a three-sequence, three-period design, and (**C**) a two-sequence, four-period design. Between the consecutive drug administrations, an adequately long washout period should be used

In case of two-stage designs appropriate steps are taken in order to preserve the overall type I error of the study [453]. A type I error is observed when the null hypothesis is rejected even though it is actually true. It is reminded that in bioequivalence assessment, the null hypothesis is that of no equivalence. Thus, the two-stage clinical designs that can be used in bioequivalence assessment should be firstly verified to not lead to inflation of the type I error, namely, to avoid false positive results. In this context, several two-stage designs and their modifications have been proposed in the literature [454–462]. Initially, two studies were published toward the validation of methods on two-stage crossover bioequivalence studies [454, 455]. Figure 10.9 is flowchart of two of the first approaches.

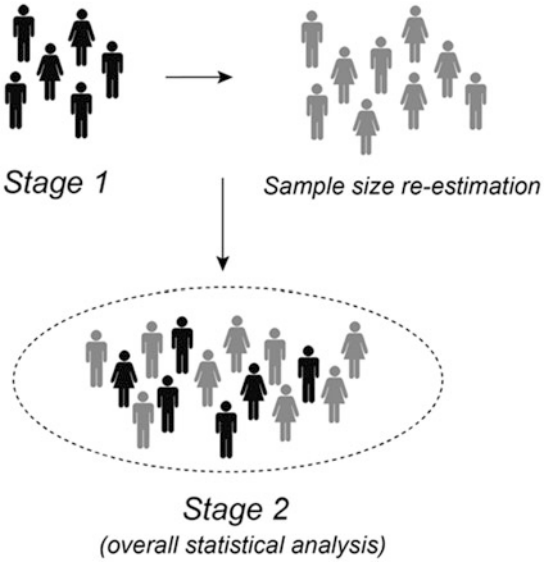


Fig. 10.8 A simplified schematic representation of the two-stage designs used in bioequivalence studies. An initial group of subjects (stage 1) is recruited based on some specific and predefined criteria in the study protocol. If for a certain reason (explicitly pre-specified in the protocol), the study needs to move on to the second stage, then sample-re-estimation takes place. Finally, data from both stages are appropriately combined and bioequivalence assessment is made using specific statistical procedures. Special concern is made to avoid inflation of type I error

It becomes evident that evaluation of a candidate two-stage designs and finally the choice of a two-stage designs is based on modeling and simulation techniques. Some basic steps of the modeling procedure are outlined below:

- The algorithm of the modeling procedure is based on the general principles of the two-stage designs as these are depicted in the relevant flowchart (e.g., as in case of Figure 10.9).
- A total number of two-stages are considered and no provision is made for additional stages.
- Each stage consists of a single 2×2 or a parallel bioequivalence study.
- An initial number of subjects are set for stage 1.
- Monte Carlo methods are used for the random generation of the pharmacokinetic profiles for the number of subjects set before.
- All estimated pharmacokinetic parameters are transformed in the ln-domain where the statistical analysis is further applied.
- For average bioequivalence purposes, the linear statistical model (expression 10.1) is used, as discussed above in this chapter. In particular, for stage 1 the effects considered are: Sequence, Period, Treatment, and Subject(Sequence). After sample size re-estimation in stage 2, the analysis of the entire set of data uses the effects: Sequence, Treatment, Stage, Period(Stage),

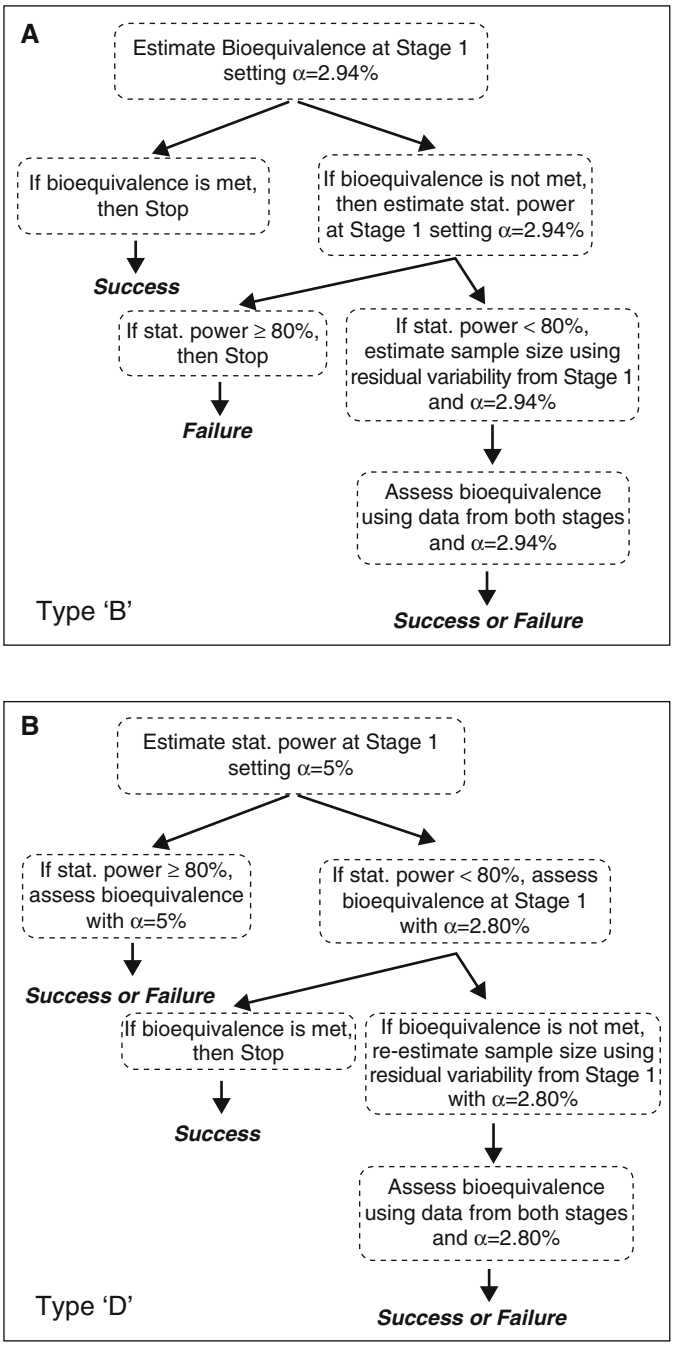


Fig. 10.9 A schematic illustration of two typical adaptive sequential designs proposed for bioequivalence purposes (**A**) the so-called type "B" and (**B**) the type "D" design

Subject(Sequence×Stage). It is worth mentioning that the additional factor "sequence×stage" proposed by the EMA was found to be unnecessary [462]. All these effects are usually considered as "fixed,"even though the choice of "fixed" or "random" effect does not alter the % bioequivalence acceptance [462].

- After finishing stage 1, power calculation is made in order to decide on the next step of the procedure.
- Power estimation relies on several factors like: the observed difference in the average pharmacokinetic parameters between the two drug products, the observed variability, the choice of type I and II errors, etc.
- For stage 2, a minimum sample size is considered to be 2. In any case, the sample size used in stage 2 should be the minimum even number.
- An upper limit for the total sample size can be defined or not, based on the individual two-stage designs examined. Based on this choice, futility criteria can be further used in the algorithm or not.
- The levels of the assumed type I error for stage 1 and 2 are a priori defined. This is necessary in order to control the overall type I error of the two-stage designs. This step is important since it affects the performance of the two-stage designs. An "ideal" two-stage designs should not lead to inflation of the type I error and concomitantly exhibit the highest possible statistical power. To this point, it should be mentioned that in case of the EMA, it is now generally proposed that a 94.12% confidence interval can be used for the analysis of both the first stage and the entire set of data.
- Usually, under each condition, millions of bioequivalence studies are simulated in order to obtain robust estimates.

10.2.5 Sample Size Estimation

Sample size estimation is an issue inextricably related with clinical design and very often is an exhaustive point when planning a bioequivalence study. It is necessary that a bioequivalence study should include an adequate number of subjects in order to have the ability to prove equivalence. Besides, bioequivalence studies should also avoid unnecessary human exposure to drugs. Usually, a statistical power value of at least 80% is required in case of bioequivalence studies. The sample size should also be adequate enough to tolerate for possible subjects dropouts. In any case, the US FDA and the EMA set the lowest limit of a bioequivalence study to be equal to 12 [237].

Setting the required number of subjects of a bioequivalence study (and generally any clinical study) is in essence a calculation step which among others depends on the type of the design (e.g., 2×2, parallel, type of higher-order, etc.), the levels of type I and II errors, the assumed difference between the two products under comparison, the encountered residual variability, the limits of bioequivalence acceptance, etc. This estimation can be made using mathematical formulas, asymptotic methods, or Monte Carlo methods. The latter require computational skills to simulate the

conditions of the bioequivalence study. In addition, they are time-consuming since they are based on iterations; however, they can offer precise estimates of sample size.

At first sight, the use of mathematical formulas appears to be an easy task. In case of the classic 2 × 2 design with ln-transformation of the data several articles have appeared in the literature suggesting mathematical formulas [463–469]. For example, the required sample size can be estimated by using the following relationships:

$$N = \frac{2\sigma_w^2 \left(Z_{1-\beta} + Z_{1-\alpha}\right)^2}{\left[(\mu_T - \mu_R) - \delta\right]^2} \qquad \text{when} \qquad \mu_T > \mu_R$$

or

$$N = \frac{2\sigma_w^2 \left(Z_{1-\beta/2} + Z_{1-\alpha}\right)^2}{\delta^2} \qquad \text{when} \qquad \mu_T = \mu_R$$

where N is the total sample size, σ_w^2 is the residual variance, and Z the normalized statistic associated with the relevant probability. The above relationships assume large sample sizes (or normal distribution) of the data and known variance. In this vein, these relationships can only lead to approximate sample size estimates. More accurate results can be derived using *t*-student statistics and applying computational approaches.

It should be underlined that sample size estimation is a crucial step in the modeling and simulation techniques for two-stage designs. This due to the fact that sample size re-estimation might take place after completion of stage 1. The latter should not only be computed accurately, but also for millions of times. However, in the routine practice of planning a bioequivalence study, the situation has now become quite easy; instead of using the above-mentioned approaches (either the Monte Carlo, or the iterative procedure), tables with sample size estimates can be found in the literature in case of several different clinical designs. Alternatively, commercially or publicly available software exist which can lead to an ease estimate of sample. size

10.2.6 Drug Metabolites

Another field in bioequivalence, where modeling and simulation methods have been extensively applied, is the assessment of the role of metabolites. Traditionally, bioequivalence studies are planned focusing on the measurement of the parent drug. This is justified since the time–concentration profile of the parent compound can detect the potential formulation differences with higher sensitivity compared to the metabolite [470, 471]. However, the contribution of metabolites in bioequivalence assessment has been a controversial issue for several years [470–473].

The methods which have been applied to investigate the role of metabolites in bioequivalence assessment can be divided into two major categories. According to the first approach, data from actual bioequivalence studies are analyzed in order to provide information for the relative performance of the parent drug and the metabolite [474, 475]. The major drawback of this method is the fact that the true condition is unknown, i.e., whether the two drug products are truly bioequivalent or not. The second methodology relies on modeling and simulation techniques, namely, the generation and assessment of simulated data [476–483]. The value of this approach relies on the fact we are able to know the correct answer of the analyzed study. Besides, modeling and simulation techniques allow us to investigate many different conditions and thus to make inferences about different types of parent drug and metabolite kinetics.

The modeling and simulation technique for the assessment of metabolites can roughly be described with the following points:

- Simulate the conditions of a bioequivalence study. The modeler can define the number of subjects and generate the relevant pharmacokinetic profile for each one.
- Generate drug time–concentration data for several pharmacokinetic models and scenarios, in order to allow inferences to as many as possible different conditions. In this vein, pharmacokinetic simulations may refer to one- and two-compartment disposition models for both or any of the parent drug or the metabolite(s), different absorption and elimination kinetics (e.g., first-order, Michaelis–Menten, etc.), place (e.g., central compartment) of formation of the metabolite, other situations like first-pass effect, and so on. For example, a schematic representation of some popular pharmacokinetic models is depicted in Figure 10.10.
- Set the desired types and levels of variability (e.g., between-subject, within-subject, analytical error, etc.), as well as the random error model (e.g., additive, proportional, combined).
- Based on mass balance principles, each pharmacokinetic model is described with a system of ordinary differential equations.
- Apply numerical solution to the above-mentioned ordinary differential equations.
- Set the desired sampling scheme (e.g., sparse, typical, dense).
- Estimate the randomly generated individual time–concentration profiles and the so-derived pharmacokinetic parameters (AUC, c_{max}, etc.).
- Divide the individual pharmacokinetic parameter data into appropriate groups in order to simulate the conditions of the desired clinical design (e.g., 2×2, parallel).
- Apply statistical analysis and assess bioequivalence for each study using the appropriate statistical framework.
- Repeat the entire procedure for thousands of times in order to obtain robust inferences.

For example, based on the mass balance principle, the pharmacokinetic model shown in Figure 10.10 F is mathematically described by the following differential equations and initial conditions:

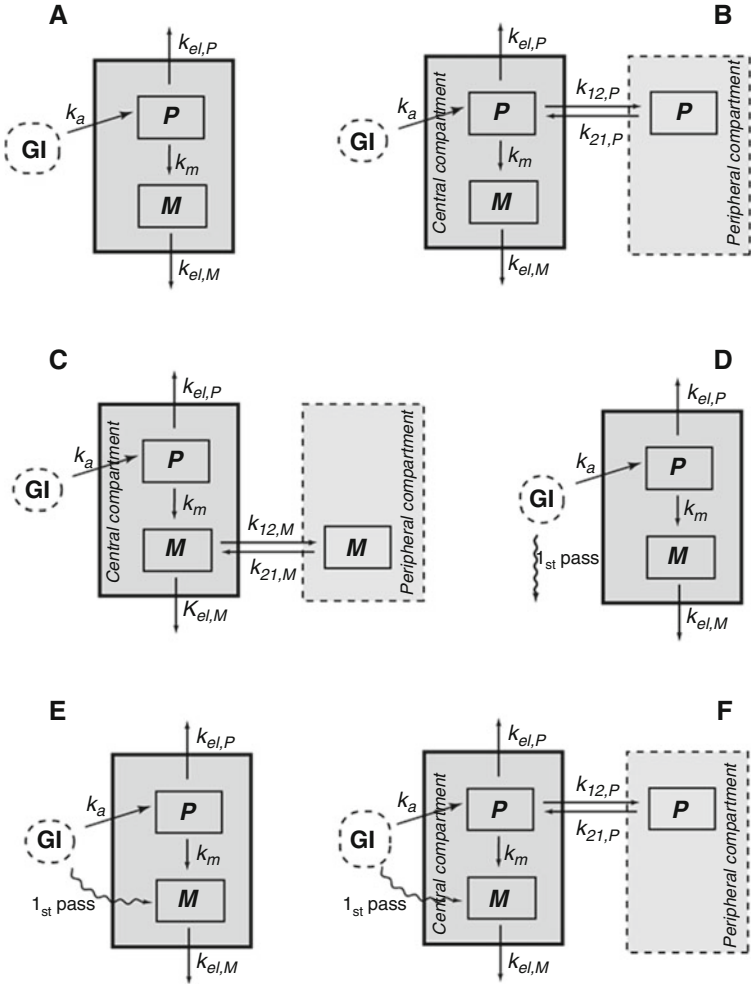


Fig. 10.10 Pharmacokinetic compartmental models used for the simulation of the parent drug (*P*) and metabolite (*M*) kinetics. One- and two-compartment models are shown for *P* and/or *M*. First-pass metabolism (assuming Michaelis–Menten kinetics) may also take place leading either to the metabolite under study (**E** and **F**), or to another compound (**D**). Arrows indicate first-order transfer from one-compartment to another. Key: *GI*, the gastrointestinal tract; k_a, first-order absorption rate constant; k_m, first-order formation rate constant of the metabolite; $k_{el,M}$, elimination rate constant of *M*; $k_{el,P}$, elimination rate constant of *P*; k_{12} central compartment to tissue rate constant (either for *M* or *P*); k_{21} tissue to central compartment rate constant (either for *M* or *P*)

- amount of parent drug in the gastrointestinal system

$$\dot{q}_{GI}(t) = -k_a q_{GI}(t) - \frac{R_{\max} q_{GI}(t)}{k_M + q_{GI}(t)},$$

- concentration of parent drug in the central compartment

$$\dot{c}_1(t) = k_a \frac{q_{GI}(t)}{V_1} - (k_{1E} + k_m)\, c_1(t) - k_{12} c_1(t) + k_{21} \frac{q_2(t)}{V_1},$$

- amount of parent drug in the peripheral compartment

$$\dot{q}_2(t) = k_{12} V_1 c_1(t) - k_{21} q_2(t),$$

- concentration of metabolite in the central compartment

$$\dot{c}_m(t) = k_m \frac{V_1}{V_m} c_1(t) - k_{mE} c_m(t) + \frac{1}{V_m} \frac{R_{\max} q_{GI}(t)}{k_M + q_{GI}(t)},$$

- initial conditions

$$c_1(0) = c_2(0) = c_m(0) = 0 \quad \text{and} \quad q_{GI}(0) = F_a q_0.$$

In the above expressions, $q_{GI}(t)$ is the amount of drug absorbed; $q_i(t)$ and $c_i(t)$ refer to the amount and the concentration, respectively, of the parent drug in the i-th compartment at time t; F_a is the absorbed fraction of the initial dose q_0; V_1 is the volume of distribution of the parent drug in the central compartment; k_a is the first-order absorption rate constant; k_{1E} refers to the elimination rate constant of the moiety from the central compartment; k_m is the first-order formation rate constant of the metabolite; $R_{\max}$ is the maximum biotransformation rate and k_M is the Michaelis–Menten constant implied in the first-pass effect; V_m is the volume of distribution of the metabolite in the central compartment and k_{mE} refers to the elimination rate constant of the metabolite from the central compartment; k_{12} and k_{21} represent the transfer rate constants between the central and peripheral compartments.

10.2.7 Drugs Interchangeability

An issue of special importance in bioequivalence and every day health practice is to investigate how similar are, one to another, two generic products (i.e., T_1 and T_2) which have been found to be bioequivalent to the same R formulation. In bioequivalence assessment, generic drug products are not compared between themselves, but only against the R product. Ideally, a generic can ideally exhibit pharmacokinetic performance identical to the innovator's, but in other cases the pharmacokinetic behavior can be more or less different. Thus, there is chance that bioequivalence between T_1 and T_2 is not ascertained, even though both of them were found to be bioequivalent with the R product. This matter is very important for the reasons of drugs' interchangeability described above in section 10.2.2.2. In order

to address this issue, namely, the comparison of two generics, several approaches have been proposed in the literature which can be divided into two main categories: indirect and direct methods.

10.2.7.1 Indirect Methods

The indirect methods use data from previous bioequivalence studies and apply statistical approaches in order to make inferences about the relationship between T_1 and T_2. In this case, the T_1 and T_2 are not compared directly one to another, since no such kind of data are available. Several indirect approaches have appeared in the literature where each one relies on one or more assumptions [484, 485]. These indirect methods, as well as their main characteristics, are outlined below:

- Application of meta-analysis assuming that (1°) all studies exhibit the same (2×2) design, (2°) the between- and within-subject variability in all bioequivalence studies between T_1, T_2, and R is the same, (3°) homogeneity of distribution of R in all studies [486].
- Use meta-analysis which assumes only that the data from the individual studies derived from the same population [487].
- Apply adjusted indirect comparison of the T_1 and T_2 treatments ignoring the randomization within the individual trials. In this case, the T_1 vs. T_2 comparison is adjusted by the results of their direct comparison with R.
- Studies comparing the properties of the above-mentioned indirect methods using several approaches to calculate the width of the 90% confidence interval such as z-distribution model, homo- or heteroscedastic method, and pragmatic method [484, 485].

10.2.7.2 Direct Methods

The main drawback of the indirect methods is the fact that they rely on several assumptions which are not completely fulfilled in practice. In order to avoid this problem, direct methods have appeared in the literature. The direct methods are based on modeling and simulation techniques and allow a straightforward comparison of the two generics (T_1 and T_2) using several hypothetical conditions and pharmacokinetic scenarios.

The direct comparison methods incorporate Monte Carlo simulations in order to examine all possible relationships between the tested products [488, 489]. The main purpose of direct methods is to make multiple comparisons of $T_1 - T_2 - R$ at the same time, namely, to make comparisons for three products in pairs of two according to a specific bioequivalence framework. In order to achieve this task, the modeling and simulation technique investigates concomitantly the performance of T_1 and T_2 when each one is compared to the R formulation. Obviously, not only the same design is used for all possible combinations ($T_1 - R$, $T_2 - R$, and $T_2 - T_1$), but

also all comparisons are made at the same time. In other words, the T_1, T_2, and R estimates, which are initially used for the $T_1 - R$ and $T_2 - R$ comparisons, are further applied to the $T_2 - T_1$ comparisons.

Several different conditions can be studied such as (1°) different true ratios of T_1/R, T_2/R and thus, T_2/T_1 fractions, (2°) levels of variability, (3°) sample size, etc. The reason for including all these combinations of values is to obtain an insight into the different situations that might be encountered in practice. The modeling and simulation technique follows the general principles presented in Figure 10.1, as well as, in other parts of this chapter. The values of the linear statistical model effects, i.e., Sequence, Period, and Subject(Sequence), remain unaltered for all comparisons; the only exception is the period effect in case of T_2 vs. T_1 comparison which is unavoidably determined by the initial comparisons (T_1 vs. R, T_2 vs. R).

Overall, the modeling and simulation techniques used in the direct methods allow to investigate the relative similarity of the generics in terms of percent bioequivalence acceptance of T_1 when compared to the R product, of T_2 vs. R, and T_2 vs. T_1. In other words, direct methods can provide the necessary evidence to draw conclusions on generics' switchability. In addition, modeling and simulation techniques can introduce new measures like the conditional probability of bioequivalence acceptance which reflects the % acceptance of generic T_2 vs. another generic T_1 given that both T are declared bioequivalent to the same R formulation.

10.3 Rising Applications

Apart from the modeling and simulation techniques which are applied following the classic bioequivalence principles, alternative applications of modeling and simulation have been proposed recently. These approaches include the use of nonlinear mixed effect modeling in bioequivalence studies, the biowaivers justification using modeling and simulation techniques with special emphasis on pediatric marketing authorizations, and the assessment of biosimilars.

10.3.1 Population Approaches

The field of population pharmacokinetics has shown an extensive growth during the last three decades. Population pharmacokinetics is a valuable tool for obtaining insight into the pharmacokinetic behavior of the mean population performance and of each individual. It can also allow the identification and quantification of several sources of variability (e.g., between- and within-subject, inter-occasion variability, etc.) in drug concentrations in the population under study. Also, population modeling has numerous advantages such as incorporating unbalanced designs, modeling

sparse data (e.g., only two or three samples per subject), and examining the role of patient-specific covariates (such as gender, age, body weight). All these topics are presented in the Appendix C.

In the area of bioequivalence, the population pharmacokinetic methodology has also been proposed and several articles have appeared in the literature investigating this issue [490–496]. Population pharmacokinetic modeling approaches have been proposed and evaluated in bioequivalence trials in order to compare the two treatments, namely, the T vs. R product [490, 491]. Several possibilities have introduced such as the global analysis of the time–concentration data, with or without the calculation of the treatment effect. Alternatively, each treatment group can be analyzed separately and individual pharmacokinetic parameters can be calculated. Modeling and simulation techniques have been used for the exploration of these techniques. It can be generally stated that the population pharmacokinetic approach leads to more reliable estimates (than the conventional bioequivalence method) when the number of samples per subject is limited [490]. However, this is a case which is not encountered in actual practice, since special concern is made a priori for the sampling scheme. The latter is imposed by the fact that for bioequivalence purposes, pharmacokinetic methods are not only applied based on the scientist's point of view, but they are also guided by regulations. For example, the regulatory guidelines highlight on the issue of the number of samples per subject, the missing and/or the inferior than the "lower limit of detection" concentration values. Thus, pharmacokinetics and regulations behave like communicating vessels in bioequivalence assessment.

Besides, population pharmacokinetic modeling approaches can be used as a tool for carrying additional information like the genotype and/or phenotype data. Identification of single nucleotide polymorphism can easily be implemented by the nonlinear mixed effects modeling and thus can reduce between-subject variability of a pharmacokinetic parameter.

10.3.2 Justification of Biowaivers

Classically, bioequivalence assessment relies on the use of clinical data from appropriately designed studies. Supplementary to these studies, other type of evidence can further be used such as in vitro data. Nevertheless, the alternate type of evidence can totally replace in vivo data in order to reduce human exposure to drugs, time, and cost of development. These situations are called biowaivers. Traditionally, in vitro data (e.g., the BCS, in vitro–in vivo correlations, etc.) were the only tool for biowaivers. However, more recently modeling and simulation techniques have been introduced as alternatives of in vivo evidence and are now accepted by the health agencies. In fact, the importance of modeling and simulation is now recognized by the EMA, and for this reason a "Modeling and Simulation Working Group" is established in the structure of the EMA.

In this context, a common area for application of modeling and simulation techniques is the case of pediatric marketing authorizations. In this case, the main focus is on the extrapolation of safety and efficacy evidence from adults to children and/or between different age groups of the children. In order to enforce this task, the EMA has published in June 2012 a concept paper toward the extrapolation of safety and efficacy in medicine development.

10.3.3 Biosimilars

During the last years, biopharmaceuticals (i.e., biological drugs) have been successfully used for the management of many diseases. When the patent of an innovator's biopharmaceutical expires, biosimilars may be placed on the market. Assessment of equivalence between biopharmaceuticals is different from the small molecule drugs; in particular, additional conditions should be specified and stricter criteria have to be used. These are mainly due to the fact that (1°) the structure of biopharmaceuticals is complex and cannot be completely characterized, (2°) even small changes in the manufacturing process can lead to important differences in the final product, and (3°) there is always a risk for immune response reactions due to their active protein structure [497–499].

For the evaluation of biosimilarity several methodological approaches have been proposed [500–503]. It is acknowledged that the determination of biosimilarity requires a thorough consideration of the evidence regarding the properties of the two biopharmaceutical products [504]. This evidence may include pharmacokinetic and pharmacodynamic data, as well as clinical safety and efficacy. For this reason, it has been proposed the formation of a fundamental biosimilarity assumption in order to bridge all obtained information. It is suggested that the statistical methodology for the assessment of biosimilarity should include both a step-wise approach and methods to assess the totality of evidence. Recently, mathematical models were proposed for the quantitative assessment of immunogenicity of therapeutic proteins [505, 506]. These models rely on the concurrent modeling of the protein pharmacokinetics and the in vivo anti-drug-antibody responses. Also, statistical methods were evaluated in order to define a depletion criterion used for immunogenicity assessment [507].

Chapter 11
Stochastic Compartmental Models

Résumons nos conclusions… C'est donc en termes probabilistes que les lois de la dynamique doivent être formulées lorsqu'elles concernent des systèmes chaotiques.

Ilya Prigogine (1917–2003)
1977 Nobel Laureate in Chemistry
La fin des certitudes

The "real world" of compartmental systems has a strong stochastic component, so we will present a stochastic approach to compartmental modeling. In deterministic theory developed in Chapter 8, each compartment is treated as being both homogeneous and a continuum. But:

- Biological media are inhomogeneous, and the simplest way to capture structural and functional heterogeneity is to operate at a molecular level. First, one has to model the time spent by each particle in the process and second, to statistically compile molecular behaviors. As will be shown in Section 11.3.4, this compilation generates a process uncertainty that did not exist in the deterministic model, and this uncertainty is the expression of process heterogeneity.
- Matter is atomic, not continuous, and cells and molecules come in discrete units. Thus, in compartmental models of chemical reactions and physiological processes, a compartment contains an integral number of units and in any transfer only an integral number of units can be transferred. Consequently, it is important to develop the theory for such systems in which transfers occur in discrete numbers of units, and that is done in terms of the probabilities of transfer of one unit from one compartment to another or to the outside.

As concluded in Chapter 7, the observed time-varying features of a process are expressions of structural and functional heterogeneity. Observations gathered from such processes were fitted by power-law and gamma-type functions. Marcus was the first to suggest stochastic modeling as an alternative working hypothesis to the empirical power-law or gamma-type functions [325]. At the same time,

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_11

stochastic modeling began to provide applications in compartmental analysis either as multivariate Markov immigration–emigration models [508–510], or as random-walk models [323, 324], or as semi-Markov (Markov renewal) models [511, 512].

In deterministic theory we started with the definition of a compartment as a kinetically homogeneous amount of material. The equivalent definition in stochastic theory is that the probability of a unit participating in a particular transfer out of a compartment, at any time, is the same for all units in the compartment.

11.1 Probabilistic Transfer Models

The present stochastic model is the so-called *particle model*, where the substance of interest is viewed as a set of particles.[1] We begin consideration of stochastic modeling by describing Markov-process models, which loosely means that the probability of reaching a future state depends only on the present and not the past states. We assume that the material is composed of particles distributed in an m-compartment system and that the stochastic nature of material transfer relies on the independent random movement of particles according to a continuous-time *Markov process*.

11.1.1 Definitions

The development of probabilistic transfer models is based on two probabilities, a conditional probability and a marginal one, commonly stated as *transfer* and *state probabilities*, respectively.

- The transfer probability $p_{ij}(t_\circ, t)$ gives the conditional probability that "a given particle resident in compartment i at time $t_\circ$ will be in compartment j at time t." Because the particles move independently, the transfer probabilities do not depend on the number of other particles in the compartments. In this way, the $p_{ij}(t_\circ, t)$ serve to express the Markovian process. Indeed, the Markov process can be expressed in terms of the $m \times m$ transfer-intensity matrix $\mathbf{H}(t)$ with (i,j)-th element $h_{ij}(t)$ given by

$$h_{ij}(t) = \lim_{\Delta t \to 0} \frac{p_{ij}(t, t+\Delta t)}{\Delta t} \qquad \text{and} \qquad h_{ii}(t) = h_{i0}(t) + \sum_{\substack{j=1 \\ j \neq i}}^{m} h_{ij}(t). \tag{11.1}$$

[1] The terms "drug molecule" and "particle" will be used in this chapter interchangeably.

The so-defined elements $h_{ij}(t)$ of the transfer-intensity matrix are called the *hazard rates*, and define the conditional probability

$$\Pr[\text{transfer to } j \text{ by } t + \Delta t \mid \text{present in } i \text{ at } t] \triangleq h_{ij}(t)\,\Delta t + o(\Delta t)$$

that "a given particle resident in i at time t leaves by $t + \Delta t$ to go in j," where Δt is small and $o(\Delta t)$ denotes all possible higher-order terms of Δt.

- The state probability $p_{ij}(t)$ is the special case of the transfer probability where $t_\circ$ is the starting time, i.e., $t_\circ \equiv 0$. The state probability gives the probability that "a given particle starting in i at time 0 is resident in compartment j at time t." All these probabilities may be considered as the (i, j)-th elements of the $m \times m$ state probabilities matrix $\mathbf{P}(t)$, with $p_{ij}(0) = 0$ when $i \neq j$. Also, to allow all possible movements, of particles starting from any initial position, the initial conditions $p_{ii}(0)$ are set to 1, i.e., $\mathbf{P}(0) = \mathbf{I}$.

In the above expressions, indices i and j may vary between 1 and m with $i \neq j$. Moreover, j may be set to 0, denoting the exterior space of the compartmental configuration.

To obtain equations for the state probabilities, write the equation for the state probability at $t + \Delta t$ as the sum of joint probabilities for all the mutually exclusive events that enumerate all the possible ways in which "a particle starting in i at 0 could pass through the various compartments at time t" AND "end up in j at $t + \Delta t$." These joint probabilities can be expressed as the product of a marginal by a conditional probability. The state probability $p_{ij}(t)$ that "a given particle starting in i at time 0 is resident in compartment s at time t" plays the role of the marginal and the transfer probability $h_{sj}(t)\,\Delta t$ that "a given particle resident in compartment s at time t will next transfer to compartment j, i.e., at time $t + \Delta t$" plays the role of the conditional probability

$$p_{ij}(t + \Delta t) = p_{ij}(t)\left[1 - h_{jj}(t)\,\Delta t\right] + \sum_{\substack{s=1 \\ s \neq j}}^{m} p_{is}(t)\,h_{sj}(t)\,\Delta t + o(\Delta t).$$

The first term in the right-hand side expresses the joint probability that "a particle starting in i at 0 is present in j at time t" AND "remains in j at $t + \Delta t$," and the second term expresses the sum of joint probabilities that "a particle starting in i at 0 is present in each compartment s, except j, at time t" AND "ends up in j at $t + \Delta t$."

Rearranging, taking the limit $\Delta t \to 0$ in the above difference equations, and neglecting the higher-order terms of Δt, one obtains m^2 differential equations, namely the *probabilistic transfer model*

$$\dot{p}_{ij}(t) = -h_{jj}(t)\,p_{ij}(t) + \sum_{\substack{s=1 \\ s \neq j}}^{m} h_{sj}(t)\,p_{is}(t).$$

These equations are linear differential equations with time-varying coefficients since the hazard rates are time-dependent and may be presented in matrix form as

$$\dot{\mathbf{P}}(t) = \mathbf{P}(t)\,\mathbf{H}(t)$$

with initial conditions $\mathbf{P}(0) = \mathbf{I}$. These models are referred to as *generalized compartmental models* and can be studied using the time-dependent Markov theory [513, 514] but are not of present interest.

In what follows, we will rather restrict ourselves mainly to the standard Markov process in the probabilistic transfer model with time-independent hazard rates. This is equivalent to assuming that the transfer probabilities do not depend on either the time the particle has been in the compartment or the previous history of the process, and

$$\mathbf{H}(t) \equiv \mathbf{H}. \tag{11.2}$$

These equations lead to the matrix solution

$$\mathbf{P}(t) = \exp(\mathbf{H}t). \tag{11.3}$$

In most pharmacokinetic applications, the system is open and the eigenvalues of $\mathbf{H}$ are real and negative. This implies that the solution has the form of a sum of negative exponentials.

11.1.2 The Basic Steps

To illustrate the successive steps in this procedure, we present the case of a simple two-compartment model, Figure 11.1. There will be four differential equations, one for each combination of the i and j previously introduced indices. For example, to obtain the differential equation for $j = 1$, one has to advocate the necessary events for a particle starting in i to pass through the two compartments at time t and end up in 1 at $(t + \Delta t)$:

- "the particle is present in 1 at time t," associated with the state probability $p_{i1}(t)$ AND "it remains in the compartment during the interval from t to $(t + \Delta t)$," associated with the transfer probability $[1 - (h_{10} + h_{12})\,\Delta t]$,

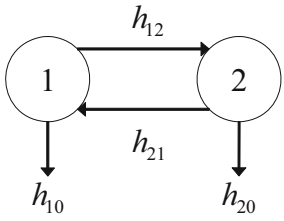


Fig. 11.1 Two-compartment configuration

OR

- "the particle is present in 2 at time t," associated with the state probability $p_{i2}(t)$ AND "it goes to 1 during the interval from t to $(t + \Delta t)$," associated with the transfer probability $h_{21}\Delta t$.

Therefore, the probability of the desired joint event may be written as

$$p_{i1}(t + \Delta t) = p_{i1}(t)[1 - (h_{10} + h_{12})\Delta t] + p_{i2}(t)[h_{21}\Delta t].$$

To obtain the differential equation for $j = 2$, one has to advocate the necessary events for a particle starting in i to pass through the two compartments at time t and end up in 2 at $(t + \Delta t)$:

- "the particle is present in 2 at time t," associated with the state probability $p_{i2}(t)$ AND "it remains in the compartment during the interval from t to $(t + \Delta t)$," associated with the transfer probability $[1 - (h_{20} + h_{21})\Delta t]$,

OR

- "the particle is present in 1 at time t," associated with the state probability $p_{i1}(t)$ AND "it goes to 2 during the interval from t to $(t + \Delta t)$," associated with the transfer probability $h_{12}\Delta t$.

Therefore, the probability of the desired joint event may be written as

$$p_{i2}(t + \Delta t) = p_{i2}(t)[1 - (h_{20} + h_{21})\Delta t] + p_{i1}(t)[h_{12}\Delta t].$$

Rearranging and taking the limit $\Delta t \to 0$ for the above difference equations, one has

$$\begin{aligned} \dot{p}_{i1}(t) &= -(h_{10} + h_{12})p_{i1}(t) + h_{21}p_{i2}(t), \\ \dot{p}_{i2}(t) &= h_{12}p_{i1}(t) - (h_{20} + h_{21})p_{i2}(t), \end{aligned} \tag{11.4}$$

with initial conditions $p_{ii}(0) = 1$ and $p_{ij}(0) = 0$ for $i \neq j$, where $i = 1, 2$. The above differential equations have as solution

$$\begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{21}(t) & p_{22}(t) \end{bmatrix} = \exp \begin{bmatrix} -(h_{10} + h_{12})t & h_{12}t \\ h_{21}t & -(h_{20} + h_{21})t \end{bmatrix}. \tag{11.5}$$

Markov processes have **H** matrices with real negative eigenvalues, which lead to models that are linear combinations of decaying exponentials, which are analogous to the deterministic models. In the presence of distinct multiple eigenvalues, the probability profiles are mixtures of exponentials multiplied by integer powers of time. The integer powers are the numbers of distinct eigenvalues [366]. Nevertheless, in practice, functions that include parts with noninteger powers of time have been needed to provide a satisfactory fit to data [323, 509]. In the face of these "impossible" experimental results, alternative working hypotheses should be created, e.g., retention-time distribution models.

11.2 Retention-Time Distribution Models

A stochastic model may also be defined on the basis of its retention-time distributions. In some ways, this conceptualization of the inherent chance mechanism is more satisfactory since it relies on a continuous-time probability distribution rather than on a conditional transfer probability in discretized units of size Δt.

One first needs the basic notions associated with a continuous probability distribution. Consider the age or the retention time of a molecule in the compartment as a random variable, A. Let:

- $f(a) = \mathrm{d}F(a)/\mathrm{d}a$ be the density function of ages A of the molecules in the compartment,
- $F(a) = \Pr[A < a]$ be the distribution function of A, i.e., the probability that "the molecule will leave the compartment prior to attaining age a," and
- $\mathcal{S}(a) = \Pr[A \geq a] = 1 - F(a)$ be the *survival function*, i.e., the probability that "the molecule survives in the compartment to age a."

From the above relations, the *hazard function* $h(a)$ is defined as

$$h(a) \triangleq f(a)/\mathcal{S}(a). \tag{11.6}$$

Also from this definition, the simple relationship

$$\frac{\mathrm{d}\log \mathcal{S}(a)}{\mathrm{d}a} = -h(a) \tag{11.7}$$

links the survival and the hazard functions.

11.2.1 Probabilistic vs. Retention-Time Models

We look now for the evaluation of the state probability $p(t)$ that "the molecule is in the compartment at time t" in the case of a one-compartment model. To this end, consider the partition $0 = a_1 < a_2 < \ldots < a_{n-1} < a_n = t$ and the $n-1$ mutually exclusive events that "the molecule leaves the compartment between its age instants a_{i-1} and a_i." The state probability $p(t)$ equals the probability of the complement of the above $n-1$ mutually exclusive events,

$$p(t) = 1 - \sum_{i=2}^{n} \Pr[\text{leaves by } a_{i-1} \text{ to } a_i] = 1 - \sum_{i=2}^{n} [F(a_i) - F(a_{i-1})]$$

$$= 1 - \sum_{i=2}^{n} [\mathcal{S}(a_{i-1}) - \mathcal{S}(a_i)] = 1 - [\mathcal{S}(a_1) - \mathcal{S}(a_n)] = \mathcal{S}(t).$$

Therefore, the survival function $\mathcal{S}(a)$ plays the same role as the state probability $p(t)$. But the former independent variable a is defined as the endogenous or within-compartment measure of time after the particle introduction to the compartment, whereas the independent variable t denotes some exterior, exogenous time measure in a system. Only for the one-compartment model do a and t have the same meaning.

The link between the probabilistic transfer model and retention-time distribution model may be explicitly demonstrated by deriving the conditional probability implied in the one-compartment probabilistic transfer model. We look for the probability, $\mathcal{S}(a + \Delta a)$, that "a particle survives to age $(a + \Delta a)$." Clearly, the necessary events are that "the particle survives to age a," associated with the state probability $\mathcal{S}(a)$ AND that "it remains in the compartment during the interval from a to $(a + \Delta a)$," associated with the conditional probability $[1 - h\Delta a]$, where h is the probabilistic hazard rate. Therefore, the probability of the desired joint event may be written as

$$\mathcal{S}(a + \Delta a) = \mathcal{S}(a)[1 - h\Delta a].$$

Then, we can write

$$h\Delta a = 1 - \frac{\mathcal{S}(a + \Delta a)}{\mathcal{S}(a)} = 1 - \frac{1 - F(a + \Delta a)}{\mathcal{S}(a)} = \frac{\Delta F(a)}{\mathcal{S}(a)} \approx \frac{f(a)\,\Delta a}{\mathcal{S}(a)}.$$

Then, the probabilistic hazard rate h is the particular hazard function value $h(a)$ evaluated at a specified age a. For the retention-time distribution models, $h(a)\,\Delta a$ gives the conditional probability "that a molecule that has remained in the compartment for age a leaves by $a + \Delta a$." In other words, the probabilistic hazard rate is the instantaneous speed of transfer.

It is noteworthy that only for the exponential distribution is the hazard rate $h(a) = f(a)/\mathcal{S}(a) = \kappa$ not a function of the age a, i.e., the molecule "has no memory" and this is the main characteristic of Markovian processes. In other words, the assumption of an exponential retention time is equivalent to the assumption of an age-independent hazard rate. One practical restriction of this model is that the transfer mechanism must not discriminate on the basis of the accrued age of a molecule in the compartment. In summary, it is clear that the formulations in the probabilistic transfer model and in the retention-time distribution model are equivalent. In the probabilistic transfer model we assume an age-independent hazard rate and derive the exponential distribution, whereas in the retention-time distribution model we assume an exponential distribution and derive an age-independent hazard rate.

For multicompartment models, in addition to the retention-time distributions within each compartment, we require the specification of the *transition probabilities* ω_{ij} of transfer among compartments. These ω_{ij}'s, assumed age-invariant, give the probabilities of transfer from a donor compartment i to each possible recipient compartment j. From (11.1), it follows that $\omega_{ij} = h_{ij}/h_{ii}$ is the probability that a particle in i will transfer to j on the next departure.

11.2.2 Markov vs. Semi-Markov Models

Consider now a multicompartment structure aiming not only at describing the observed data but also at providing a rough mechanistic description of how the data were generated. This mechanistic system of compartments is envisaged with the drug flowing between the compartments. The stochastic elements describing these flows are the ω_{ij} transition probabilities as previously defined. In addition, with each compartment in this mechanistic structure, one can associate a retention-time distribution $f_i(a)$. The so-obtained multicompartment model is referred to as the *semi-Markov* formulation. The semi-Markov model has two properties, namely that:

- the transition probabilities ω_{ij} are time-invariant; this implies that the sequence of compartment visitations for a particle may be described by a Markov chain and
- the retention-time distributions are arbitrary.

The semi-Markov formulation in the compartmental context was originally proposed by Purdue [512] and Mehata and Selvan [515]. The present approach attempts to characterize fully the mechanistic flow pattern between compartments and to use nonmechanistic models with the smallest number of parameters to describe the within-compartment processes. The experimenter might first divide the system into compartments based on known theory. The retention-time distributions within each compartment are specified either through expert knowledge from hazard rates or by fitting alternative models to data. The ω_{ij} transition probabilities are then determined. One advantage of using these nonmechanistic retention times is the incorporation of inhomogeneous compartments and consequential particle age discrimination with a minimum number of additional parameters.

The Markov model is a special case of the semi-Markov model in which all the retention-time variables are exponentially distributed, $A_i \sim \text{Exp}(\kappa_i)$, and κ_i is the parameter of the exponential. In this case, the semi-Markov model parameters are $\kappa_i = h_{ii}$ and $\omega_{ij} = h_{ij}/\kappa_i$ for $j \neq i$ and $i, j = 1 : m$. This results from the assumption of the Markov model given in (11.1), which implies that the conditional transfer probability from i to j in a time increment Δt is time-invariant, or in other words is independent of the "age" of the particle in the compartment. Particles with such a constant flow rate, or hazard rate, are said to "lack memory" of their past retention time in the compartment.

Figure 11.2 shows a two-compartment Markov model with parameters h_{10}, h_{12}, and h_{21} and the semi-Markov model with exponentially distributed retention times with parameters κ_1, κ_2, and ω. The conversion relationships are

$$h_{10} = \kappa_1 (1 - \omega) \qquad \text{and} \qquad h_{12} = \kappa_1 \omega \qquad \text{and} \qquad h_{21} = \kappa_2$$

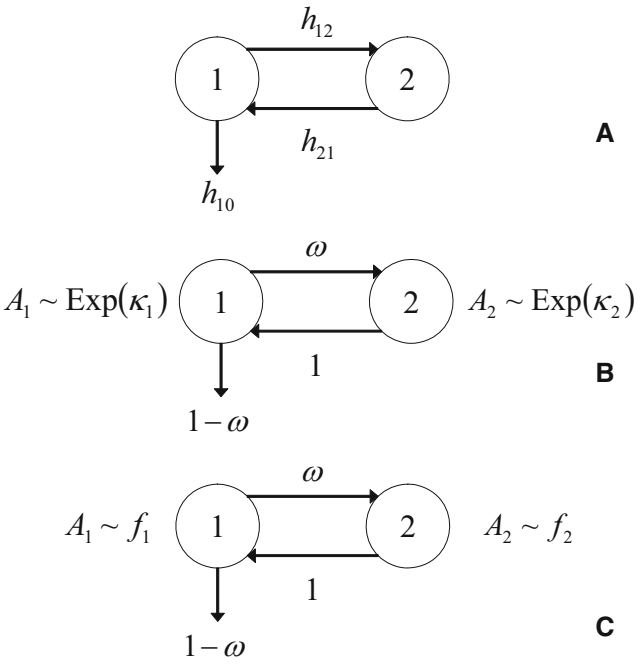


Fig. 11.2 The two-compartment Markov model (**A**) vs. the semi-Markov model with exponential retention times (**B**) vs. the general semi-Markov model (**C**)

from semi-Markov to Markov, and

$$\kappa_1 = h_{10} + h_{12} \qquad \text{and} \qquad \kappa_2 = h_{21} \qquad \text{and} \qquad \omega = \frac{h_{12}}{h_{10} + h_{12}}$$

from Markov to semi-Markov with exponential retention-time distribution. We note that the compartmental structure of the semi-Markov model is simply determined by means of ω. The figure also shows the general semi-Markov model having the same structure determined by ω but allowing several distribution models f_1 and f_2 for the retention times A_1 and A_2, respectively, in the compartments.

The causes of nonexponential retention times, and hence age-varying hazard rates, may be numerous. Two general reasons for such retention times in pharmacokinetic applications are noninstant mixing and compartmental heterogeneity. Noninstant mixing, for example, is likely to occur in compartmental models with oral dosing. Inhomogeneous compartments, on the other hand, are a natural consequence of the lumping inherent in dividing the body into two or three compartments, for example, into a central and a peripheral compartment. Conversely but less likely, if all the drug particles in a compartment were homogeneous and also well stirred, then the transfer processes that determine how the drug particles leave the compartment could not discriminate on the basis of the accumulated age of a particle in the compartment. Hence such homogeneous, well-stirred compartments could be modeled using exponential transit times.

Table 11.1 Density, survival, and hazard functions.

	$f(a)$	$\mathcal{S}(a)$	$h(a)$
Exp(κ)	$\kappa \exp(-\kappa a)$	$\exp(-\kappa a)$	κ
Erl(λ, ν)	$\frac{\lambda^\nu a^{\nu-1}}{(\nu-1)!} \exp(-\lambda a)$	$\exp(-\lambda a) \sum_{i=0}^{\nu-1} \frac{(\lambda a)^i}{i!}$	$\frac{\frac{\lambda^\nu a^{\nu-1}}{(\nu-1)!}}{\sum_{i=0}^{\nu-1} \frac{(\lambda a)^i}{i!}}$
Ray(λ)	$\lambda^2 a \exp\left[-\frac{1}{2}(\lambda a)^2\right]$	$\exp\left[-\frac{1}{2}(\lambda a)^2\right]$	$\lambda^2 a$
Wei(λ, μ)	$\mu\lambda^\mu a^{\mu-1} \exp[-(\lambda a)^\mu]$	$\exp[-(\lambda a)^\mu]$	$\mu\lambda^\mu a^{\mu-1}$

11.2.3 Irreversible Models

11.2.3.1 One-Compartment Model

The one-compartment model is the typical simple irreversible model. For the 1-compartment model and only when initial conditions are given, the exterior time t and the molecule ages a are the same. The state probability $p(t)$ that a molecule is in the compartment is $\mathcal{S}(t)$:

$$p(t) \equiv \mathcal{S}(t). \tag{11.8}$$

One has simply to assume a particular probability distribution for A with the survival function available in a closed form, namely the exponential, Erlang, Rayleigh, and Weibull. Table 11.1 summarizes the probability density functions, survival functions, and hazard rates for the above-mentioned distributions. In these expressions, λ is the scale parameter and μ and ν are shape parameters with $\kappa, \lambda, \mu > 0$ and $\nu = 1, 2, \ldots$.

- Exponential distribution. The survival function is a single exponential $p(t) = \exp(-\kappa t)$. A deterministic one-compartment model produces the same profile, so one can say that this model is the single-exponential distribution of residence times. However, following instantaneous administration of drugs, the time–concentration observed profiles sometimes present two decreasing phases on the semilogarithmic plot. This may be described using the one-compartment model and assuming a mixed distribution consisting of two exponential survival functions $p(t) = \gamma \exp(-\kappa_1 t) + (1-\gamma) \exp(-\kappa_2 t)$, where γ ($\gamma < 1$) represents the relative contribution of the first exponential term. These biphasic profiles are usually attributed to the two-compartment models. However, there is no rigorous conjunction between the two-exponential and two-compartment models since more complex compartmental models may also give biphasic-like profiles with certain combinations of the microconstants. In the same way, one can use the single compartment model and conceive mixed survival functions containing three or more exponential forms leading to three- or more-phasic profiles. It follows therefore that one cannot discriminate on the basis of observed

data alone between the situation in which the survival function in the single compartment is the sum of two exponentials and the situation in which a single exponential survival function is associated with each of the two compartments present in the configuration.

- Erlang distribution. We assume that $A \sim \mathrm{Erl}(\lambda, \nu)$. The state probability is

$$p(t) = \exp(-\lambda t) \sum_{i=0}^{\nu-1} \frac{(\lambda t)^i}{i!}.$$

- Weibull and Rayleigh distributions. From Table 11.1, we have

$$p(t) = \exp[-(\lambda t)^{\mu}] \tag{11.9}$$

for the Weibull distribution and as a special case with $\mu = 2$, the Rayleigh distribution.

Figure 11.3 depicts state probability curves for the Erlang and the Weibull distributions. The hazard rates as functions of time are also illustrated. For $\nu = 1$ and $\mu = 1$, we obtain the behavior corresponding to an exponential retention-time distribution and to the one-compartment deterministic profile.

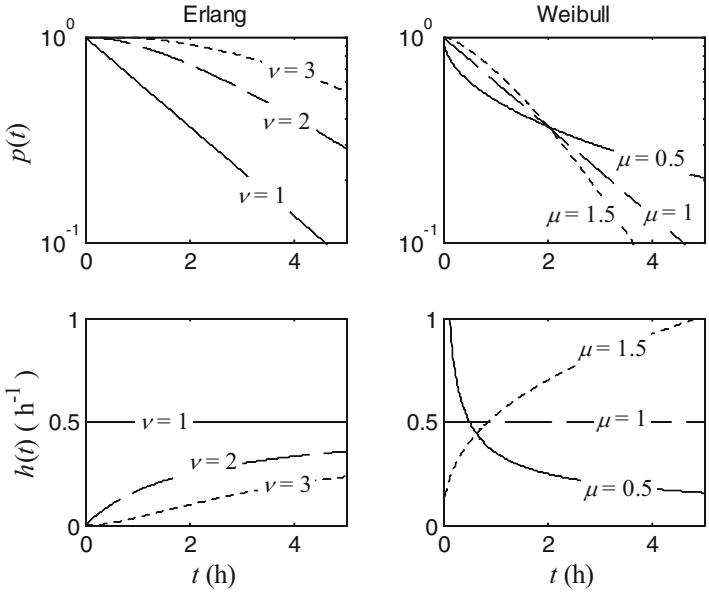


Fig. 11.3 State probabilities and hazard functions with $\lambda = 0.5\,\mathrm{h}^{-1}$, and $\nu = 1, 2, 3$ and $\mu = 0.5, 1, 1.5$ for Erlang and Weibull distributions, respectively

For $\nu > 1$, in case of an Erlang distribution, the rate function at age 0 is $h(0) = 0$, after which the rate increases and the kinetic profile has a log-concave form. This provides an initial dampening of the retention probability of newly introduced particles. Then, the rate is asymptotic to λ as the age increases. This implies that the age discrimination within the compartment diminishes, either rapidly or slowly depending on ν, as the retention time increases.

The Weibull distribution allows noninteger shape parameter values, and the kinetic profile is similar to that obtained by the Erlang distribution for $\mu > 1$. When $0 < \mu < 1$, the kinetic profile presents a log-convex form and the hazard rate decreases monotonically. This may be the consequence of some saturated clearance mechanisms that have limited capacity to eliminate the molecules from the compartment. Whatever the value of μ, all profiles have common ordinates, $p(1/\lambda) = \exp(-1)$.

These qualitative features are typical of data from inhomogeneous compartments and/or compartments with noninstant initial mixing. Reciprocally, many compartments that are not well stirred have these properties. For these reasons, the Erlang and Weibull retention-time distributions have been very useful in practice to fit to data. In a theoretical context, Weiss classifies the retention-time distributions according to the log-convexity or concavity of the corresponding time–concentration profiles. Moreover, that author attempts to explain these profiles by assuming time-varying mechanisms as time-varying volume of distribution or clearance capacities [516–518]. As in Section 7.5 for the empirical models, these investigations again reveal the strong link between time dependence and process heterogeneity.

11.2.3.2 Multicompartment Models

Consider the irreversible two-compartment model with survival, distribution, and density functions $\mathcal{S}_1(a)$, $F_1(a)$, $f_1(a)$ and $\mathcal{S}_2(a)$, $F_2(a)$, $f_2(a)$ for "ages" a of molecules in compartments 1 and 2, respectively. We will assume that at the starting time, the molecules are present only in the first compartment. The state probability $p_1(t)$ that "a molecule is in compartment 1 at time t" is $\mathcal{S}_1(a)$ with $t = a$; the external time t is the same with the age of the molecule in the compartment 1, i.e., $p_1(t) = \mathcal{S}_1(t)$. The state probability $p_2(t)$ that "a molecule survives in compartment 2 after time t" depends on the length of the time interval a between entry and the 1 to 2 transition, and the interval $t - a$ between this event and departure from the system. To evaluate this probability, consider the partition $0 = a_1 < a_2 < \ldots < a_{n-1} < a_n = t$ and the $n - 1$ mutually exclusive events that "the molecule leaves the compartment 1 between the time instants a_{i-1} and a_i." By applying the *total probability theorem* (cf. Appendix D), $p_2(t)$ is expressed as

$$\sum_{i=2}^{n} \Pr[\text{survive in 2 to } t \mid \text{leave 1 by } a_{i-1} \text{ to } a_i] \Pr[\text{leave 1 by } a_{i-1} \text{ to } a_i].$$

If $\max(a_{i-1} - a_i) \to 0$:

- $\Pr[\text{survive in 2 to } t \mid \text{leave 1 by } a_{i-1} \text{ to } a_i] = \mathcal{S}_2(t-a)$ and
- $\Pr[\text{leave 1 by } a_{i-1} \text{ to } a_i] = F_1(a_i) - F_1(a_{i-1}) = \mathrm{d}F_1(a)$

with $a \in [a_{i-1}, a_i]$. It follows that $p_2(t)$ is the Stieltjes integral of $\mathcal{S}_2$ with respect to F_1 (cf. Appendix E):

$$p_2(t) = \int_0^t \mathcal{S}_2(t-a)\,\mathrm{d}F_1(a) = \int_0^t \mathcal{S}_2(t-a) f_1(a)\,\mathrm{d}a = f_1 * \mathcal{S}_2(t), \tag{11.10}$$

that is the convolution of the density function f_1 in the input site with the survival function $\mathcal{S}_2$ in the sampling site. Similarly, the probability that "the molecule will leave the compartment 2 prior to time t" is

$$\int_0^t F_2(t-a)\,\mathrm{d}F_1(a) = f_1 * F_2(t),$$

that is, the convolution of the density function f_1 in the input site with the distribution function F_2 in the sampling site.

This result can be generalized in the case of a catenary irreversible m-compartment model [515]; the state probability in the compartment i ($i = 2 : m$) at t is given by

$$p_i(t) = f_1 * \ldots * f_{i-1} * \mathcal{S}_i(t).$$

An elegant form of the previous expressions is obtained in the frequency domain. The convolution becomes the product of the Laplace transform of the survival and the density functions:

$$\tilde{p}_i(s) = \tilde{f}_1(s) \ldots \tilde{f}_{i-1}(s)\,\tilde{\mathcal{S}}_i(s), \tag{11.11}$$

where $\tilde{f}(s)$ is the Laplace transform of $f(t)$, $\tilde{f}(s) = L\{f(t)\}$ (cf. Appendix F).

For analyzing general irreversible compartmental configurations, Agrafiotis [519] developed a semi-Markov technique on the basis of conditional distributions on the retention time of the particles in the compartments before transfer into the next compartment. This approach uses the so-called *forces of separation*, and it is quite different from the one introduced at the beginning of this section, where the distribution of the retention time in each compartment is independent of the compartment that the particle is transferring to.

11.2.4 Reversible Models

Consider the reversible two-compartment model that is explained by way of the semi-Markov formulation as illustrated in Figure 11.2C. We will assume that at the starting time all molecules are present in compartment 1. A single molecule that is present at the initial time in compartment 1 stays there for a length of time that has a single-passage density function $f_1(a)$. Then, it has the possibility to leave the system definitively with probability $1-\omega$ or reach compartment 2 with probability ω. The retention time in this compartment is governed by the single-passage density function $f_2(a)$. At the end of its stay in compartment 2, the molecule reenters compartment 1. Our goal is to evaluate the probability $p_1(t)$ that "a molecule survives in 1 after time t." This event is the compilation of the following mutually exclusive events:

- "survive in 1 without visit in 2" with probability $\mathcal{S}_1(t)$,
- "survive in 1 with 1 visit in 2" with probability $\omega \mathcal{S}_1 * f_1 * f_2(t)$,
- "survive in 1 with 2 visits in 2" with probability $\omega^2 \mathcal{S}_1 * f_1 * f_2 * f_1 * f_2(t)$,

etc. for an infinite number of visits. The probability of a composite event is then

$$p_1(t) = \mathcal{S}_1(t) * \sum_{i=0}^{\infty} \omega^i [f_1 * f_2(t)]^{*i},$$

where f^{*m} denotes the m-fold convolution of f with itself. This last expression has the structure of the renewal density for positive random variables, which is studied in probability theory [514]. Taking the Laplace transform of this expression, the probability has a simpler form in the frequency domain:

$$\tilde{p}_1(s) = \tilde{\mathcal{S}}_1(s) \sum_{i=0}^{\infty} \omega^i \tilde{f}_1^i(s) \tilde{f}_2^i(s).$$

Because of the boundness of density functions $0 < \int_0^t f(a)\mathrm{d}a < 1$, the infinite sum has a closed-form expression:

$$\tilde{p}_1(s) = \frac{\tilde{\mathcal{S}}_1(s)}{1 - \omega \tilde{f}_1(s) \tilde{f}_2(s)}. \tag{11.12}$$

The probability $p_2(t)$ that "a molecule survives in 2 after time t" will be given by using the inverse Laplace transform of

$$\tilde{p}_2(s) = \frac{\omega \tilde{f}_1(s) \tilde{\mathcal{S}}_2(s)}{1 - \omega \tilde{f}_1(s) \tilde{f}_2(s)}.$$

If at the starting time the molecules are present only in compartment 2, the Laplace transform of the state probabilities in the compartments are

$$\tilde{p}_1(s) = \frac{\tilde{f}_2(s)\,\tilde{\mathcal{S}}_1(s)}{1-\omega\tilde{f}_1(s)\tilde{f}_2(s)} \qquad \text{and} \qquad \tilde{p}_2(s) = \frac{\tilde{\mathcal{S}}_2(s)}{1-\omega\tilde{f}_1(s)\tilde{f}_2(s)}.$$

To deal with more complex compartmental configurations, block diagrams and transfer functions are now introduced. Block diagrams are extensively used in the automatic control field [334] to represent the functionality of a process. These are diagrams involving a set of elements each of them representing a given function. When the process is described by a mathematical model, each element of the block diagram represents a mathematical operation such as scaling, integration, addition, and multiplication. Here block diagrams are used to represent a compartmental configuration by specifying the pathways and the retention sites that a molecule can encounter when it is administered in the system. Associated with the block diagrams, the concept of a transfer function is of fundamental importance in the analysis of control systems and feedback problems in general. After specifying in the block diagram a site of external action, i.e., the input or administration site, and a site of observation, i.e., the output or sampling site, the transfer function is defined as the ratio of the Laplace transform of the output of a system to its input.

For instance, in the simple one-compartment model associated with gamma retention-time distribution $A \sim \text{Gam}(\lambda, \mu)$,

$$f(a) = \frac{\lambda}{\Gamma(\mu)}(\lambda a)^{\mu-1}\exp(-\lambda a),$$

the transfer function is

$$\tilde{f}(s) = \frac{\lambda^{\mu}}{(s+\lambda)^{\mu}}.$$

The usefulness of the transfer functions lies in the fact that:

- The problem of obtaining the transfer function of a complex system, composed of two or more simple elements, consists in combining the transfer functions of its elements following some elementary rules [334].
- Once the transfer function is known for a particular complex system, then the response of the system to any known input is readily found by multiplying the transfer function by the Laplace transform of the input.

The same problem arises when the pharmacokinetic system is decomposed into subsystems that can be characterized by transfer functions, but where no closed-form solution emerges for the complete system. This holds for the evaluation of the time–amount curve after oral administration when the models of the input and disposition process are known. In these cases, numerical techniques may be of substantial help in performing inverse Laplace transforms. These methods fall into two classes,

(1°) approximations by Fourier series expansion and (2°) numerical integration in the complex plane. The validation of these methods and the performance of the available software have been tested directly with pharmacokinetic models [520]. These evaluations showed that simulations as well as parameter estimations from functions defined only in the Laplace domain can be associated with the same degree of reliability as in the conventional case, in which the models are directly given as functions of time. These techniques are commonly used in pharmacokinetics for recirculation drug modeling [521, 522] or physiological modeling [523, 524]. With few exceptions [525], all these approaches are deterministic, and consequently, they use exponential retention-time distributions.

These findings can be summarized in the following procedure to obtain equations for a semi-Markov stochastic model:

1. Represent the underlying mechanistic model with the desired physiological structure through a set of phenomenological compartments with their interconnections.
2. Obtain the equivalent semi-Markov representation by specifying the transition probabilities ω_{ij} and the single-passage retention-time distributions $f_i(a)$ for each compartment. Obtain the block diagram representation using the transition probabilities as gain factors and the Laplace transforms of single-passage density functions as transfer functions.
3. Solve the system of algebraic equations in the frequency domain to obtain the transfer function between the input and sampling sites. The Laplace transform of the probability that "a molecule survives in the sampling site after time t" is this transfer function where substitution of the multiplying $\tilde{f}_i(s)$ factor in the sampling site with the corresponding $\tilde{\mathcal{S}}_i(s)$ was made.
4. Evaluate in the time domain the time–amount course by applying traditional inverse Laplace transforms [526] or numerical inversion techniques [520, 527].

Conceiving models based on block diagrams may be quite complex, involving feedback loops and time delays. A paper [528] shows in detail how such a model can be constructed for a pharmacokinetic system. On the other hand, retention-time reversible models can be very powerful and flexible for simulation and data fitting.

Example 6. Simulate a Complex System

The procedure is presented in a complex system involving three compartments. Figure 11.4 illustrates the original model (upper panel) and the semi-Markov model (lower panel), and Figure 11.5 shows the block diagram representation. If we denote by $u(s)$ the input function and by $y_2(s)$ and $y_3(s)$ the output functions at the sampling sites 2 and 3, respectively, we can write

$$y_2(s) = \left[\omega_{12}\tilde{f}_1(s)\,u(s) + \omega_{32}y_3(s)\right]\tilde{f}_2(s)$$

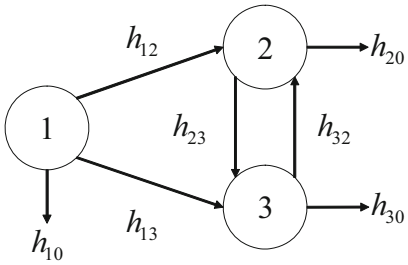


Fig. 11.4 Complex 3-compartment configuration. The administration site is in compartment 1 and the sampling sites are compartments 2 and 3

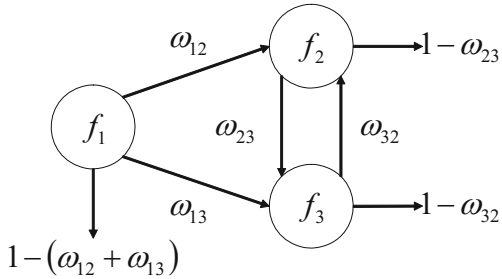
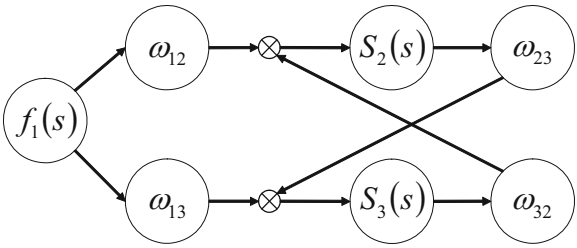


Fig. 11.5 Block diagram representation of the complex system shown in Figure 11.4



and

$$y_3(s) = \left[\omega_{13}\tilde{f}_1(s)\,u(s) + \omega_{23}y_2(s)\right]\tilde{f}_3(s).$$

Solving with respect to $y_2(s)$ and $y_3(s)$, we obtain the transfer functions between the administration and sampling sites:

$$\frac{y_2(s)}{u(s)} = \frac{\tilde{f}_1(s)\left[\omega_{12} + \omega_{13}\omega_{32}\tilde{f}_3(s)\right]\tilde{f}_2(s)}{1 - \omega_{23}\omega_{32}\tilde{f}_2(s)\tilde{f}_3(s)} \tag{11.13}$$

and

$$\frac{y_3(s)}{u(s)} = \frac{\tilde{f}_1(s)\left[\omega_{13} + \omega_{12}\omega_{23}\tilde{f}_2(s)\right]\tilde{f}_3(s)}{1 - \omega_{23}\omega_{32}\tilde{f}_2(s)\tilde{f}_3(s)}. \tag{11.14}$$

The Laplace transform of the probabilities $p_2(t)$ and $p_3(t)$ that "a molecule survives in 2 and 3, respectively, after time t" are obtained by substituting in (11.13), $y_2(s)/u(s)$ and $\tilde{f}_2(s)$ with $\tilde{p}_2(s)$ and $\tilde{\mathcal{S}}_2(s)$, respectively,

$$\tilde{p}_2(s) = \frac{\tilde{f}_1(s)\left[\omega_{12} + \omega_{13}\omega_{32}\tilde{f}_3(s)\right]\tilde{\mathcal{S}}_2(s)}{1 - \omega_{23}\omega_{32}\tilde{f}_2(s)\tilde{f}_3(s)},$$

and in (11.14), $y_3(s)/u(s)$ and $\tilde{f}_3(s)$ by $\tilde{p}_3(s)$ and $\tilde{\mathcal{S}}_3(s)$, respectively,

$$\tilde{p}_3(s) = \frac{\tilde{f}_1(s)\left[\omega_{13} + \omega_{12}\omega_{23}\tilde{f}_2(s)\right]\tilde{\mathcal{S}}_3(s)}{1 - \omega_{23}\omega_{32}\tilde{f}_2(s)\tilde{f}_3(s)}.$$

■

First, Purdue [529] reviewed the use of the semi-Markov theory, from which in principle, the requisite $p_{ij}(t)$ regression function may be determined for arbitrary (nonexponential) retention-time distributions. Although the semi-Markov formulation is elegant, the mechanism determining the sequential location of the particles in the compartmental structure is highly complex, and it may be difficult to write down explicit expressions when one is dealing with a general multicompartment system. The solutions are given in general terms involving an infinite sum of convolutions, and the complexity generally rules out an analytical solution for the $p_{ij}(t)$ function.

11.2.5 Time-Varying Hazard Rates

The initial idea is to use the differential equations of a probabilistic transfer model with hazard rates varying with the age of the molecules, i.e., to enlarge the limiting hypothesis (11.2). The objective is to find nonexponential families of survival distributions that are mathematically tractable and yet sufficiently flexible to fit the observed data. In the simplest case, the differential equation (11.7) links hazard rates and survival distributions. Nevertheless, this relation was at the origin of an erroneous use of the hazard function. In fact, substituting in this relation the age a with the exogenous time t, we obtain

$$\dot{\mathcal{S}}(t) = -h(t)\,\mathcal{S}(t),$$

which looks like the deterministic one-compartment model (8.4) with time-varying fractional flow rate $k(t)$, where the amount of the substance $q(t)$ and $k(t)$ are associated with the $\mathcal{S}(t)$ and $h(t)$, respectively. This correspondence is valid only in exceptional cases, and particularly for multicompartment configurations, the use of a hazard function $h(a)$ as a time-varying fractional flow rate $k(t)$ must be handled with extreme care.

11.2.5.1 One-Compartment Model

Since the exterior time t and the age of the molecules a are the same for the one-compartment model, we can use the previous equation to write

$$\dot{p}(t) = -h(t)\,p(t)\,,\ \ p(0) = 1. \tag{11.15}$$

The solution is given by

$$p(t) = \exp\left[-\int_0^t h(a)\,\mathrm{d}a\right].$$

The closed-form solutions are more difficult to obtain than those previously obtained by means of the survival functions. Numerical integration or quadrature can be used to solve the differential equation or the integral. For instance:

- If the hazard rate is Weibull:

  $$\dot{p}(t) = -\mu\lambda^{\mu}t^{\mu-1}p(t)\,,\ \ p(0) = 1.$$

  This form is very similar to the model often used when the molecules move across fractal media, e.g., the dissolution rate using a time-dependent coefficient given by (5.12) to describe phenomena that take place under dimensional constraints or understirred conditions [17]. The previous differential equation has the solution given by (11.9).
- If the hazard rate is $h(t) = \left(\frac{\alpha}{t} + \beta\right)$:

  $$p(t) = \gamma t^{-\alpha}\exp(-\beta t)\,,$$

  where γ is a normalizing constant. This model involving terms like $t^{-\alpha}$ or $t^{-\alpha}\exp(-\beta t)$ contains only two parameters and seems to be applicable to fit some of the data much better for many drugs [269].

In a pioneer work, Marcus established the link between some usual time-varying forms of $h(t)$ and $f(a)$ in a single compartment [325]. For instance in $h(t) = \left(\frac{\alpha}{t} + \beta\right)$, $\alpha = 1$ leads to $A \sim \mathrm{Gam}(\lambda, \beta)$ and $1 < \alpha < 2$ defines the standard extreme stable-law density with exponent α. In the case of $a = 1.5$, the obtained distribution is known as the retention-time distribution of a Wiener process with drift.

The use of age-dependent hazard rates provides a large increase in modeling flexibility, and such models are currently investigated with increasing interest for the following two reasons, among others:

- Many processes have rates that are inherently age-dependent, e.g., various digestion and enzyme-kinetic processes.

- Some complex standard models with many compartments may be simplified by using approximate age-dependent models with fewer parameters, and thus often with superior subsequent statistical analysis. One such application is the description of mixing in passage models.

From a practical point of view, starting from observed data, we are looking for the retention-time distribution $f(a)$ of molecule ages. Using the data and (11.15), recursive techniques may be applied to reveal an approximative time profile of $h(t)$ (cf. Section 7.7). On a second level, this profile can be identified using retention-time distributions from Table 11.1.

11.2.5.2 Multicompartment Models

This formulation is the time-varying alternative to the probabilistic transfer models assuming constant hazard rates as defined by (11.1), and it can be accommodated by generalizing the Markov processes. These models with age-varying hazard rates are expressed by a set of linear differential equations with time-varying coefficients. One may call them *generalized compartmental models* since they satisfy the equations of a deterministic model with k_{ij} being a function of age a. Nevertheless, reference must always be made to the stochastic origin of these equations and confusion avoided between the exogenous time t and ages a of the molecules in the compartments.

Let us examine now the conditions for which a probabilistic transfer model is equivalent to a retention-time model, both using the same hazard functions. More precisely, for the irreversible multicompartment structures, the study can be reduced to the analysis of an irreversible two-compartment model, where compartment no.1 embodies all compartments before compartment no.2. We have to compare two situations:

- The probabilistic transfer model whose differential form is

$$\dot{p}_2(t) = h_1(t)\, p_1(t) - h_2(t)\, p_2(t) = f_1(t) - h_2(t)\, p_2(t),$$

  where no distinction is made between the time t and ages a.
- The retention-time model expressed by convolution (11.10). The derivation of this convolution product leads to (cf. Appendix E)

$$\dot{p}_2(t) = f_1(t) + f_1 * \dot{\mathcal{S}}_2(t),$$

  where by definition $\dot{\mathcal{S}}_2(t) = -h_2(t)\, \mathcal{S}_2(t)$.

  By merging the last three equations, one has

$$f_1 * [h_2(t)\, \mathcal{S}_2(t)] = h_2(t) f_1 * \mathcal{S}_2(t),$$

which is the condition of equivalence, i.e., $h_2(t)$ must commute to ensure the equivalence between the probabilistic transfer and the retention-time models, both using the same hazard functions. Among the usual hazard functions the exponential distribution has this property. For the reversible multicompartment structures, such a condition further reduces the set of possible distribution functions. For exponential distributions alone we may have such equivalence, but the model degenerates into a Markovian one.

11.2.6 Pseudocompartment Techniques

This section proposes the use of a semi-Markov model with Erlang- and phase-type retention-time distributions as a generic model for the kinetics of systems with inhomogeneous, poorly stirred compartments. These distributions are justified heuristically on the basis of their shape characteristics. The overall objective is to find nonexponential retention-time distributions that adequately describe the flow within a compartment (or pool). These distributions are then combined into a more mechanistic (or physiologically based) model that describes the pattern of drug distribution between compartments. The new semi-Markov model provides a generalized compartmental analysis that can be applied to compartments that are not well stirred.

11.2.6.1 The Erlang-Type Retention-Time Distributions

The Erlang distributions used as retention-time distributions $f_i(a)$ have interesting mathematical properties considerably simplifying the modeling. For the Erlang distribution, it is well known that if ν independent random variables Z_i are distributed according to the exponential distribution

$$Z_i \sim \mathrm{Exp}(\lambda), \qquad i = 1 : \nu,$$

then their summation follows an Erlang distribution:

$$Z = \sum_{i=1}^{\nu} Z_i \sim \mathrm{Erl}(\lambda, \nu).$$

The application of this statement in the present context enables one to represent the process responsible for the retention of molecules by a chain of ν catenary compartments, each of them associated with an exponential retention-time distribution with parameter λ. This compartment chain is well known as the *pseudocompartment* chain, but no physical or mechanistic meaning may be associated with this chain. It simply represents a formal way to take into account the Erlang retention-time distribution.

11.2.6.2 The Phase-Type Retention-Time Distributions

A more general yet tractable approach to semi-Markov models is the phase-type distribution developed by Neuts [530], who showed that any nondegenerate distribution $f(a)$ of a retention time A with nonnegative support can be approximated, arbitrarily closely, by a distribution of phase type. Consequently, all semi-Markov models in the recent literature are special phase-type distribution models. However, the phase-type representation is not unique, and in any case it will be convenient to consider some restricted class of phase-type distributions.

The phase-type distribution has an interpretation in terms of the compartmental model. Indeed, if the phenomenological compartment in the model, which is associated with a nonexponential retention-time distribution, is considered as consisting of a number of pseudocompartments (phases) with movement of particles between these pseudocompartments or out of them, then the retention time of a particle within the entire phenomenological compartment will have a phase-type distribution. The pseudocompartments do not have a *mechanistic* interpretation but rather are a mathematical artifice to generate the desired retention-time distribution.

Using the Markovian formulation, the expanded set of pseudocompartments leads to the solution

$$\mathbf{P}^*(t) = \exp\left(\mathbf{H}^* t\right),$$

where $\mathbf{H}^*$ and $\mathbf{P}^*(t)$ are the transfer-intensity and the state probability matrices, respectively, both associated with the pseudocompartment structure. It has been shown that the solutions for the state probability $p(t)$ of the phenomenological compartment with the assumed retention-time distribution may be obtained by finding appropriate linear combinations of the $p_{ij}^*(t)$. Mathematically, one has

$$p(t) = \underline{b}_1^T \mathbf{P}^*(t)\, \underline{b}_2,$$

where $\underline{b}_1^T = \left[1\ 0\ \ldots\ 0\right]$ and $\underline{b}_2^T = \left[1\ 1\ \ldots\ 1\right]$ are m-dimensional vectors of indicator variables, i.e., 0's or 1's. The elements of $\underline{b}_1$ indicate the origin of particles in the pseudocompartment structure and $\underline{b}_2$ indicates that all the pseudocompartments contribute to build the phenomenological compartment.

Perhaps the most commonly used example of a phase-type distribution is the $\mathrm{Erl}(\lambda, \nu)$ distribution, defined by the catenary system consisting of ν pseudocompartments. According to the phase-type concept for generating distributions, one can find phase-type distributions that exhibit rich kinetic behaviors using concatenation of Erlang distributions associated with several λ's: this case is reported as the *generalized Erlang distribution*. Further kinetic flexibility can be achieved by using feedback pathways and partition of hazard rates in the pseudocompartment structures.

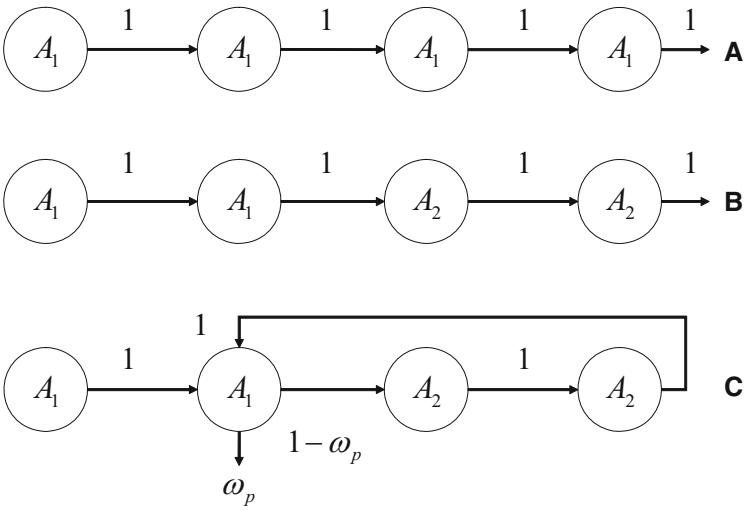


Fig. 11.6 Pseudocompartment configurations generating Erlang (**A**), generalized Erlang (**B**), and phase-type (**C**) distributions for retention times in phenomenological compartments. Retention times are distributed according to $A_1 \sim \text{Exp}(\lambda_1)$ and $A_2 \sim \text{Exp}(\lambda_2)$

To describe heterogeneity within a compartment, Figure 11.6 illustrates three pseudocompartment configurations, each of them involving four pseudocompartments in all.

- The first one (A) is a catenary system with pseudocompartments associated with a λ_1 hazard rate. The transfer-intensity matrix $\mathbf{H}^*$ is

$$\mathbf{H}^* = \begin{bmatrix} -\lambda_1 & \lambda_1 & 0 & 0 \\ 0 & -\lambda_1 & \lambda_1 & 0 \\ 0 & 0 & -\lambda_1 & \lambda_1 \\ 0 & 0 & 0 & -\lambda_1 \end{bmatrix},$$

  and the phase-type distribution generated by this structure is $\text{Erl}(\lambda_1, 4)$.
- Like the previous system, the second (B) is also a catenary one, but two pseudocompartments are associated with the λ_1 hazard rate, and two others with the λ_2 hazard rate. The transfer-intensity matrix is

$$\mathbf{H}^* = \begin{bmatrix} -\lambda_1 & \lambda_1 & 0 & 0 \\ 0 & -\lambda_1 & \lambda_1 & 0 \\ 0 & 0 & -\lambda_2 & \lambda_2 \\ 0 & 0 & 0 & -\lambda_2 \end{bmatrix},$$

  and the generalized Erlang density function is more dispersed with the actual parameter values than the previous one.

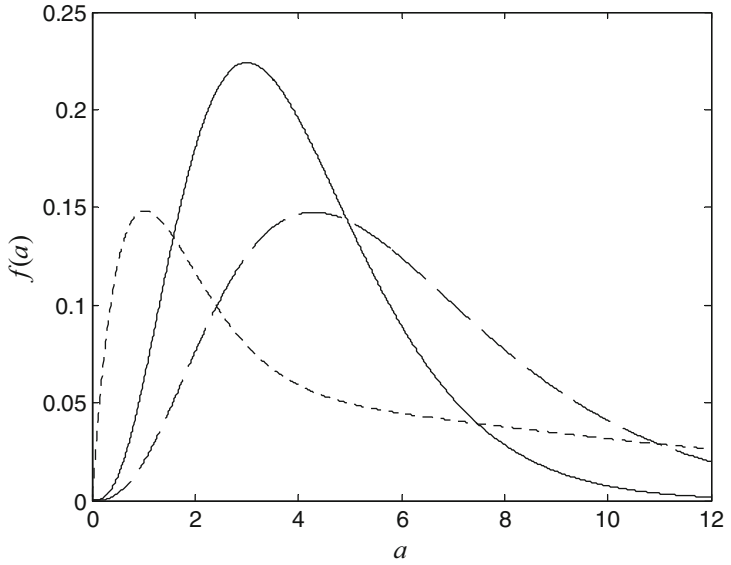


Fig. 11.7 Retention-time densities generated by pseudocompartment configurations: Erlang (*solid line*), generalized Erlang (*dashed line*), and phase-type densities (*dotted line*)

- The third configuration (C) is unusual because the phenomenological compartment output takes place from the second pseudocompartment and the output of the last pseudocompartment is fed back to the second pseudocompartment. The transfer-intensity matrix is

$$\mathbf{H}^* = \begin{bmatrix} -\lambda_1 & \lambda_1 & 0 & 0 \\ 0 & -\lambda_1 & (1-\omega_p)\lambda_1 & 0 \\ 0 & 0 & -\lambda_2 & \lambda_2 \\ 0 & \lambda_2 & 0 & -\lambda_2 \end{bmatrix},$$

  and the resulting density is "long-tailed."

For the three pseudocompartment configurations presented above, Figure 11.7 depicts the obtained density functions with parameters set to $\lambda_1 = 1, \lambda_2 = 0.25\,\mathrm{h}^{-1}$, and $\omega_p = 0.3$.

The phase-type distributions are designed to serve as retention-time distributions in semi-Markov models. To obtain the equations of the model for a phenomenological compartmental configuration, one has to follow the following procedure:

1. Represent the underlying mechanistic model with the desired physiological structure through a set of phenomenological compartments with their interconnections.

2. Express the retention-time distribution for each phenomenological compartment by using phase-type distributions. However, the phase-type distributions for these sites are determined empirically. There is no assurance of finding the "best" phase-type distribution. This step leads to the expanded model involving pseudocompartments generating the desired phase-type distribution.
3. For the resulting model with phase-type distributions, find the expanded transfer-intensity and the state probability matrices of the equivalent Markov model $\mathbf{H}^*$ and $\mathbf{P}^*(t)$, respectively.
4. Simulate the kinetic behavior by combining the $\mathbf{P}^*(t)$ probability functions for the pseudocompartments to obtain the state probabilities $\mathbf{P}(t)$ of a particle belonging to the phenomenological compartments at time t. That is expressed by means of appropriate matrices $\mathbf{B}_1$ and $\mathbf{B}_2$ with indicator variables, i.e., 0's or 1's:

$$\mathbf{P}(t) = \mathbf{B}_1 \mathbf{P}^*(t) \mathbf{B}_2.$$

The elements of $\mathbf{B}_1$ indicate the origin of particles in the pseudocompartment structure and establish the correspondence between the numbering of the original compartments and the sequence of the pseudocompartments. The elements of $\mathbf{B}_2$ indicate the summing of pseudocompartments to yield the phenomenological compartment.

11.2.6.3 Structured Models

Although the structured models are at the origin of the pseudocompartment concept, these models are less well known [531, 532]. The structured models are compartmental systems, but with a structure that describes the dynamics in a physically reasonable way. Imposing some structure on the compartmental model is certainly a way of dealing with possible ill-conditioning of more general models. The resulting model has only a few parameters, and is capable of fitting well some observed data. The proposed structure is more holistic, in the sense that the compartments themselves may not have an obvious physical interpretation, but the system as a whole does. Although the number of compartments is increasing, the number of estimated parameters does not because the model is structured, unlike traditional compartmental modeling. Models of this type include some well-known systems [509], and they have been used as examples in other work [342]. This structured compartmental model has some similarities to the dispersion model [293], but it does have certain advantages.

Faddy [531] consider the compartmental configuration, shown schematically in Figure 11.8, where compartments are numbered 1, 2, ..., m are pseudocompartments, with the starting compartment numbered i_s ($2 \leq i_s \leq m$). The material is transferred between compartments over time according to a Markov process, where the positive parameters h_+, h_-, and h_0 are the hazard rates. Thus a molecule administered to the system would be able to clear the system only via the series of compartments $i_s + 1 : m$, corresponding to $\text{Erl}(h_0, m - i_s + 1)$ distributed retention

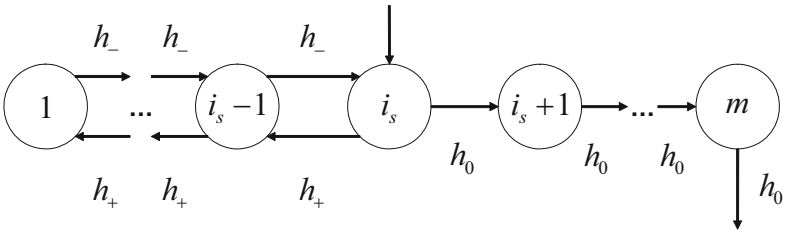


Fig. 11.8 Structured Markovian model. Diffusion is expressed by means of h_+, h_-, and compartments 1 to i_s. Erlang-type elimination is represented by means of h_0 and compartments i_s to m. The drug is given in compartment i_s and cleared from compartment m

times. As previously noted, a large value of m would be exemplified by a "hump" in any observed retention data, corresponding to a delay in clearance of the drug. Pseudocompartments $1, 2, \ldots, i_s$ correspond to the states of a random walk with reflecting barrier at 1, which describes the retention of the drug by movement of elements between nearest-neighbor sites within a heterogeneous peripheral medium. For large i_s, this random walk can be thought of as approximating a diffusion. Such a model thus describes drug kinetics in terms of two components:

- diffusion within the heterogeneous peripheral medium and
- Erlang distributed retention times describing the elimination from the system.

Retention data that after a possible delay in concentration show a sharp decline followed by a long tail would be modeled by $i_s \gg 2$ and $h_0 \gg h_- > h_+$. The condition $h_- > h_+$ ensures that the drift of the random walk (or diffusion) is away from the reflecting barrier. Figure 11.9 illustrates the probability profiles in the distribution and elimination compartments when $m = 20$, $i_s = 15$, $h_+ = 0.1$, $h_- = 0.2$, and $h_0 = 1$.

In summary, any stochastic semi-Markov model may be represented as an expanded Markov model. This simply involves subdividing each compartment into a number of pseudocompartments, leading to a matrix that essentially defines a new expanded compartmental system, but with many more compartments [531]. After passing through a sequence of pseudocompartments, a particle would transfer according to the ω_{ij} transition probabilities. Thus, multicompartment modeling may be done using the definitions and the methodology developed for the probabilistic transfer models. Therefore, the formulation of the probabilistic transfer model is immediate and hence the questions associated with the nature of the eigenvalues and the complexity of the analytical solutions may be attempted using suitable numerical procedures and computer software. The assumption of the Erlang retention-time distributions has several consequences:

- for irreversible models, the eigenvalues of the **H** matrix may be multiple real, leading to a "time-power" solution,
- for reversible models, the eigenvalues may be real or complex multiple values, with negative parts leading to damped oscillations.

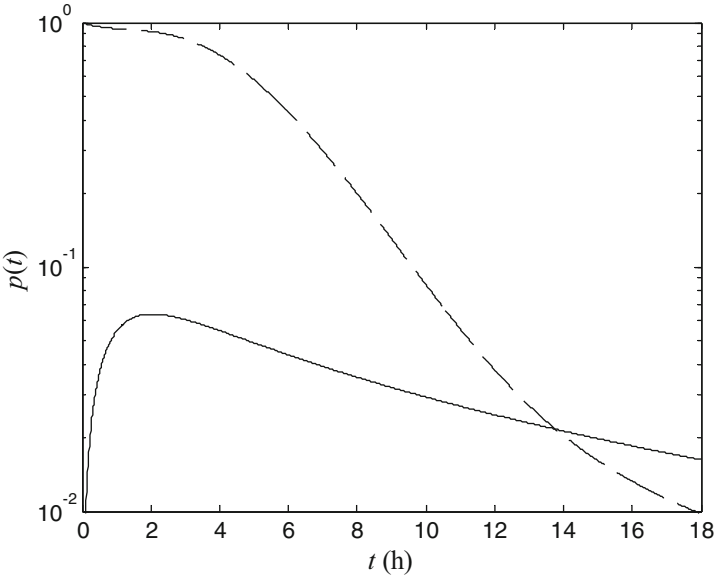


Fig. 11.9 The total probabilities in the distribution (*solid line*) and elimination (*dashed line*) compartments of the Faddy structured model

Erlang- and phase-type distributions provide a versatile class of distributions, and are shown to fit naturally into a Markovian compartmental system, where particles move between a series of compartments, so that phase-type compartmental retention-time distributions can be incorporated simply by increasing the size of the system. This class of distributions is sufficiently rich to allow for a wide range of behaviors, and at the same time offers computational convenience for data analysis. Such distributions have been used extensively in theoretical studies (e.g., [533]), because of their range of behavior, as well as in experimental work (e.g., [534]). Especially for compartmental models, the phase-type distributions were used by Faddy [531] and Matis [326, 331] as examples of "long-tailed" distributions with high coefficients of variation.

11.2.7 A Typical Two-Compartment Model

The mechanistic model is the traditional reversible two-compartment model. For this model, Karalis et al. [535] hypothesized a well-stirred compartment, the central compartment, and a heterogeneous, peripheral compartment. In general, one would assume that the sampling site is a well-stirred medium ensuring sampling feasibility technology where the particles mix quickly and homogeneously with blood, e.g., the central compartment. But such an assumption is not valid for

the peripheral compartment that represents soft tissues, muscles, or bone or other organs, Figure 11.2A. We assume that all molecules are present in compartment 1 at time 0. In the following, we express the heterogeneity in the peripheral compartment in several manners.

11.2.7.1 Semi-Markov Formulation

We propose to use as single-passage retention-time distributions the $A_1 \sim \text{Exp}(\kappa)$ for the central compartment and the $A_2 \sim \text{Gam}(\lambda, \mu)$ distribution for the peripheral compartment and we assume that all molecules are present in compartment 1 at initial time. According to (11.12),

$$\tilde{p}_1(s) = \frac{1}{s + \kappa - \omega_e \kappa \left(\frac{\lambda}{s+\lambda}\right)^{\mu}}. \tag{11.16}$$

Using the numerical inverse Laplace transform and $\kappa = 2\,\text{h}^{-1}$, $\lambda = 1\,\text{h}^{-1}$, $\omega_e = 0.8$, and $\mu = 1, 4, 6$, Figure 11.10 illustrates the $p_1(t)$ time profiles.

Instead of the gamma single-passage distribution for the peripheral compartment, Wise [323] proposed the mixed random walk in series distribution,

$$f(t) \propto t^{-w} \exp\left[-\phi\left(\frac{t}{\mu} + \frac{\mu}{t}\right)\right],$$

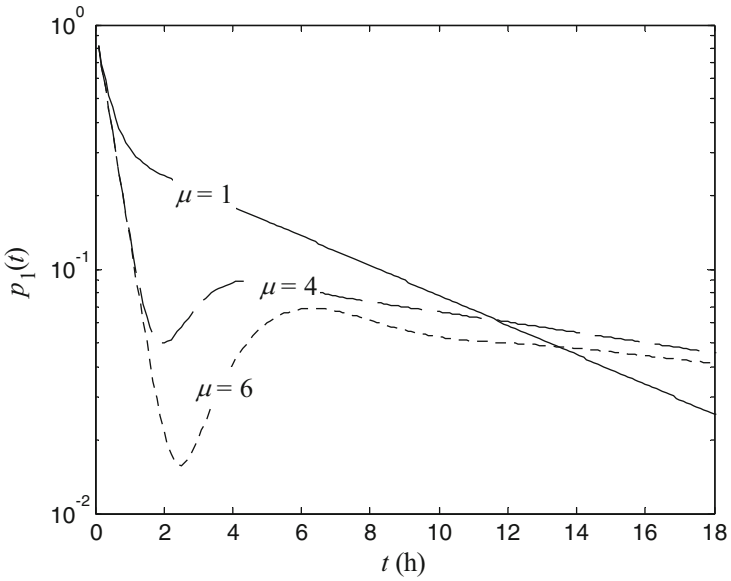


Fig. 11.10 Simulation of time–$p_1(t)$ profiles for $\mu = 1, 4, 6$

to justify the gamma-type function $\gamma t^{-\alpha} \exp(-\beta t)$ often used as an empirical model to fit several series of data. These gamma profiles can also be interpreted in terms of a recirculation process, where the single-passage retention time is the generalized inverse Gaussian distribution [271].

11.2.7.2 Erlang-Type Distribution

We propose the retention-time distributions $A_1 \sim \text{Exp}(\kappa)$ and $A_2 \sim \text{Erl}(\lambda, \nu)$ for the first and second compartments, respectively. The peripheral compartment 2 is then constituted by the ν pseudocompartments that are required to express $\text{Erl}(\lambda, \nu)$. It follows that

$$\kappa = h_{10} + h_{12} \qquad \text{and} \qquad \lambda = h_{21} \qquad \text{and} \qquad \omega_e = \frac{h_{12}}{h_{10} + h_{12}}.$$

The system now becomes an $m = \nu + 1$ compartment model and the probabilistic transfer differential equations are

$$\dot{p}_1 = -\kappa p_1 + \lambda p_m, \qquad \dot{p}_2 = \omega_e \kappa p_1 - \lambda p_2, \qquad \dot{p}_j = \lambda \left(p_{j-1} - p_j\right), \qquad j = 3 : m.$$

In the above equations, p_i represents the probability that a molecule starting in compartment 1 is in compartment i at time t. By using $\tau = \lambda t$ and $\mu = \kappa/\lambda$, one obtains the dimensionless system of differential equations

$$\dot{p}_1 = -\mu p_1 + p_m, \qquad \dot{p}_2 = \omega_e \mu p_1 - p_2, \qquad \dot{p}_j = p_{j-1} - p_j, \qquad j = 3 : m.$$

This model is a special case of the model studied by Matis and Wehrly [536] in which $A_1 \sim \text{Erl}(\lambda_1, \nu_1)$ and $A_2 \sim \text{Erl}(\lambda_2, \nu_2)$ retention-time distributions are associated with the first and second compartments, respectively. The analysis of the characteristic polynomial of this model implies that there are at least two complex eigenvalues, except for the case $\nu = 2$ with parameters satisfying the condition

$$\omega_e < \frac{4(\mu - 1)^3}{27\mu}.$$

The practical significance is that for the above two-compartment models with large ν or large ω_e, the p_i do not have the simple commonly used sum of exponential forms but damped oscillatory ones. According to the μ and ω_e values, one obtains a broad spectrum of models able to fit unusual data profiles. Figure 11.11 illustrates the p_1 profiles for $\nu = 2, 4, 6$ associated with $\kappa = 2\,\text{h}^{-1}$, $\lambda = 1\,\text{h}^{-1}$, and $\omega_e = 0.8$. These simulations are identical to those obtained with the transfer functions in Figure 11.10. Therefore, Erlang distributions are useful for a class of problems in which there is initial dampening of the conditional transfer probability due to such phenomena as noninstant mixing.

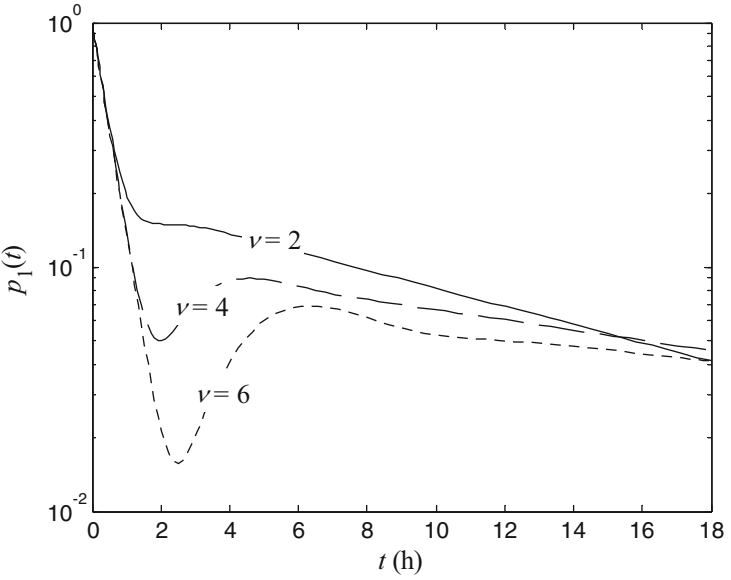


Fig. 11.11 Simulation of time–$p_1\,(t)$ profiles for $\nu = 2, 4, 6$

11.2.7.3 Phase-Type Distribution

We propose the use of the phase-type distributions previously developed as retention-time distributions associated with the peripheral compartment. The numeric values of the parameters are $\kappa = 2\,\mathrm{h}^{-1}$, $\lambda_1 = 1\,\mathrm{h}^{-1}$, $\omega_e = 0.8$, $\lambda_2 = 0.25\,\mathrm{h}^{-1}$, and $\omega_p = 0.3$. For the three cases, the transfer-intensity matrices $\mathbf{H}^*$ of the equivalent Markov model are

$$\mathbf{H}^* = \begin{bmatrix} -\kappa & \omega_e\kappa & 0 & 0 & 0 \\ 0 & -\lambda_1 & \lambda_1 & 0 & 0 \\ 0 & 0 & -\lambda_1 & \lambda_1 & 0 \\ 0 & 0 & 0 & -\lambda_1 & \lambda_1 \\ \lambda_1 & 0 & 0 & 0 & -\lambda_1 \end{bmatrix},$$

$$\mathbf{H}^* = \begin{bmatrix} -\kappa & \omega_e\kappa & 0 & 0 & 0 \\ 0 & -\lambda_1 & \lambda_1 & 0 & 0 \\ 0 & 0 & -\lambda_1 & \lambda_1 & 0 \\ 0 & 0 & 0 & -\lambda_2 & \lambda_2 \\ \lambda_2 & 0 & 0 & 0 & -\lambda_2 \end{bmatrix},$$

$$\mathbf{H}^* = \begin{bmatrix} -\kappa & \omega_e\kappa & 0 & 0 & 0 \\ 0 & -\lambda_1 & \lambda_1 & 0 & 0 \\ \omega_p\lambda_1 & 0 & -\lambda_1 & (1-\omega_p)\,\lambda_1 & 0 \\ 0 & 0 & 0 & -\lambda_2 & \lambda_2 \\ 0 & 0 & \lambda_2 & 0 & -\lambda_2 \end{bmatrix},$$

respectively. Also, the $\mathbf{B}_1$ and $\mathbf{B}_2$ matrices with the indicator variables are

$$\mathbf{B}_1 = \begin{bmatrix} 1\,0\,0\,0\,0 \\ 0\,1\,0\,0\,0 \end{bmatrix} \quad \text{and} \quad \mathbf{B}_2^T = \begin{bmatrix} 1\,0\,0\,0\,0 \\ 0\,1\,1\,1\,1 \end{bmatrix}.$$

Figure 11.12 illustrates the $p_1\,(t)$ state probability of having a particle in the sampled compartment at time t. These $p_1\,(t)$ use the retention-time distributions presented in Figure 11.7. The profile of $p_1\,(t)$ that corresponds to $\mathrm{Erl}(\lambda_1, 4)$ is the same as that drawn in Figure 11.11.

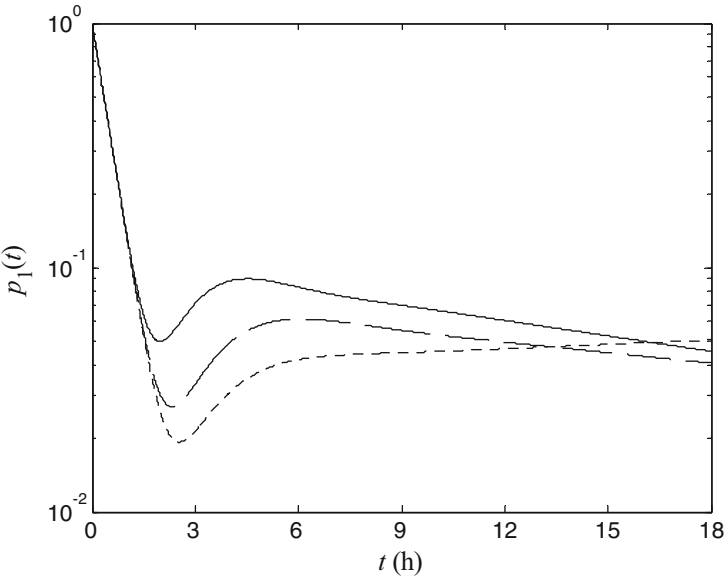


Fig. 11.12 Simulation of time–$p_1\,(t)$ profiles using pseudocompartments to generate Erlang (*solid line*), generalized Erlang (*dashed line*), and phase-type densities (*dotted line*); cf. Figure 11.7

11.3 Time–Concentration Profiles

The probabilistic transfer and retention-time models are models evaluating the transition or retention probabilities that are associated with a single particle. This is why these models are called the *particle models*. In order to account for all the particles in the process and administered amounts, one needs to make further statistical and practical considerations.

11.3.1 Routes of Administration

Let us consider some drug administration practicalities. Up to now, the administered amounts were considered as initial units introduced simultaneously into several compartments at the beginning of the experiment. These amounts were considered as initial conditions to the differential equations describing the studied processes. Nevertheless, this concept seems to have limited applications in pharmacokinetics. In this section, we develop the probabilistic transfer and retention-time models associated with an extravascular or intravascular route of administration.

In both cases, an extra compartment is introduced: the absorption or the infusion balloon compartment for the extravascular and intravascular route, respectively. To model these disposition processes, we again apply probabilistic analysis for these compartments looking for the probability $p\,(t+\Delta t)$ that a particle is present at time $(t+\Delta t)$ in that compartment. Clearly, the necessary events are "that the particle is present at time t," associated with the state probability $p\,(t)$ AND "that it remains in the compartment during the interval from t to $(t+\Delta t)$," associated with the conditional probability $[1-h\Delta t]$. Therefore, the probability of the desired joint event may be written as

$$p\,(t+\Delta t) = p\,(t)\,[1-h\Delta t]\,. \tag{11.17}$$

For the extravascular and intravascular routes, $p\,(t)$ will be referred to as $p_{ev}\,(t)$ and $p_{iv}\,(t)$, and h will be referred to as h_{ev} and h_{iv}, respectively.

The two routes of administration can be formulated as follows:

- In the extravascular case, the compartment is the absorption compartment and the hazard rate h_{ev} represents the absorption rate constant. If we assume that h_{ev} is not dependent on time, rearranging (11.17), taking the limit $\Delta t \to 0$, and solving the so obtained differential equation with initial condition $p_{ev}\,(0) = 1$, we obtain

$$p_{ev}\,(t) = \exp\,(-h_{ev}t)\,.$$

  The retention-time distribution follows:

$$f_{ev}\,(t) = h_{ev}p_{ev}\,(t) = h_{ev}\exp\,(-h_{ev}t)\,.$$

The Laplace transform of the extravascular retention-time distribution is

$$\tilde{f}_{ev}(s) = \frac{h_{ev}}{s + h_{ev}}.$$

- In the intravascular case with a constant rate infusion between the starting time T_S and the ending time T_E, the state probability $p_{iv}(t)$ is given by

$$p_{iv}(t) = \frac{T_E - t}{T_E - T_S}\left[H(t - T_S) - H(t - T_E)\right].$$

Solving (11.17) for h_{iv}, we obtain a time-varying hazard rate h_{iv},

$$h_{iv}(t) = \frac{1}{T_E - t}\left[H(t - T_S) - H(t - T_E)\right],$$

and the retention-time distribution

$$f_{iv}(t) = h_{iv}(t)\, p_{iv}(t) = \frac{1}{T_E - T_S}\left[H(t - T_S) - H(t - T_E)\right]. \tag{11.18}$$

In the above relationships, $H(t)$ is the step Heaviside function. The Laplace transform of the intravascular retention-time distribution is

$$\tilde{f}_{iv}(s) = \frac{1}{s(T_E - T_S)}\left[\exp(-T_S s) - \exp(-T_E s)\right]. \tag{11.19}$$

For both cases, the retention-time distribution functions $f_{ev}(t)$ and $f_{iv}(t)$ are similar to the input functions $v_{ev}(t)$ and $v_{iv}(t)$, respectively, defined for the deterministic models. The only difference is that in the stochastic consideration, the drug amounts are not included is these input functions.

In conclusion, in order to account for the usual routes of administration, one has to expand the system by artifactual compartments and associated retention-time distributions corresponding to the extravascular or intravascular routes. So, the expanded system may now be considered without environmental links and the administration protocol is simply expressed by the initial conditions in the input compartments. In this case, at least one of the m compartments must be considered as the input compartment.

11.3.2 Some Typical Drug Administration Schemes

In the following, we present how to apply the above relationships for the compartmental model shown in Figure 11.13.

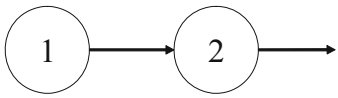


Fig. 11.13
Two-compartment irreversible system

11.3.2.1 Extravascular Case

The most frequent situation is the heterogeneous absorption materialized by retention-time distributions $A_1 \sim \text{Erl}(\lambda, \nu)$ and $A_2 \sim \text{Exp}(\kappa)$ for compartments 1 and 2, respectively. In this configuration, compartment 1 represents the heterogeneous absorption compartment and compartment 2 represents the distribution compartment that is the sampled compartment. The state probability $p(t)$ that "a molecule initially introduced into compartment 1 is in compartment 2 at time t" is evaluated using (11.10) with

$$f_1(a) = \frac{\lambda^\nu a^{\nu-1}}{(\nu-1)!} \exp(-\lambda a) \qquad \text{and} \qquad \mathcal{S}_2(a) = \exp(-\kappa a).$$

The convolution integral can be evaluated using the Laplace transform. In fact,

$$L\{f_1(a)\} = \frac{\lambda^\nu}{(s+\lambda)^\nu} \qquad \text{and} \qquad L\{\mathcal{S}_2(a)\} = \frac{1}{s+\kappa}.$$

In these expressions, L and s denote the Laplace operator and Laplace variable, respectively. The solution is given by

$$p(t) = L^{-1}\left\{\frac{\lambda^\nu}{(s+\kappa)(s+\lambda)^\nu}\right\}. \tag{11.20}$$

The inverse Laplace calculus of (11.20) leads to

$$p(t) = \gamma^\nu \exp(-\kappa t) - \exp(-\lambda t) \sum_{i=1}^{\nu} \gamma^i \frac{(\lambda t)^{\nu-i}}{(\nu-i)!}$$

with $\gamma = \lambda/(\lambda - \kappa)$.

11.3.2.2 Intravascular Case

Here the drug is administered by a constant rate infusion over T hours. This model may be conceived in two different ways:

- Probabilistic transfer model. The model is a special case of the two-compartment model presented in Figure 11.1, where compartment 1 is associated with the infusion balloon and compartment 2 is associated with the central compartment. The links between compartments are specified as $h_{12} = h_{iv}(t)$, $h_{21} = 0$, $h_{10} = 0$, and $h_{20} = h$. The state probabilities associated with compartment 1

are $p_{11}(t) = p_{iv}(t)$ and $p_{21}(t) = 0$. The probabilistic transfer equation for the central compartment 2 is obtained directly from (11.4):

$$\dot{p}_{12}(t) = h_{iv}(t)\, p_{iv}(t) - h p_{12}(t).$$

Given (11.18), the solution of the differential equation is

$$p_{12}(t) = \frac{1}{Th}\left\{\exp\left[-h(t-T)\,H(t-T)\right] - \exp(-ht)\right\}. \tag{11.21}$$

This equation gives the probability that "a molecule set in the infusion balloon 1 at time 0 is present in the central compartment 2 at time t."

- Retention-time model. The model is an irreversible two-compartment model whose solution is given by (11.11):

$$\tilde{p}_2(s) = \tilde{f}_1(s)\,\tilde{\mathcal{S}}_2(s),$$

where $\tilde{f}_1(s) \equiv \tilde{f}_{iv}(s)$ and $\tilde{\mathcal{S}}_2(s) = (s+h)^{-1}$. The solution is given by

$$p_2(t) = \frac{1}{Th}\left\{\exp\left[-h(t-T)\,H(t-T)\right] - \exp(-ht)\right\}.$$

Example 7. Infusion for the Typical Two-Compartment Model

This example concerns the typical two-compartment model previously presented under the semi-Markov formulation (cf. Section 11.2.7). By assuming that molecules are initially present in the central compartment, (11.16) is the Laplace transform of the survival function in that compartment. If now the drug molecules are administered by a constant rate infusion between T_S and T_E, the Laplace transform of the survival function in the central compartment becomes

$$\tilde{p}_1^*(s) = \tilde{f}_{iv}(s)\,\tilde{p}_1(s) = \frac{\exp(-T_S s) - \exp(-T_E s)}{s(T_E - T_S)} \frac{1}{s + \kappa - \omega_e \kappa \left(\frac{\lambda}{s+\lambda}\right)^{\mu}}.$$

This expression is obtained by reporting (11.16) and (11.19) into (11.10). Using the numerical inverse Laplace transform and $\kappa = 2\,\mathrm{h}^{-1}$, $\lambda = 1\,\mathrm{h}^{-1}$, $\omega_e = 0.8$, and $\mu = 1, 4, 6$, Figure 11.14 illustrates the $p_1^*(t)$ time profiles for a 6-h constant-rate infusion. This figure takes into account the infusion duration, whereas Figure 11.10 considers that all molecules are in compartment 1 at initial times. ■

11.3.3 Time–Amount Functions

After specifying the route of drug administration, we now turn to some statistical considerations in order to express the behavior of all particles administered into the system.

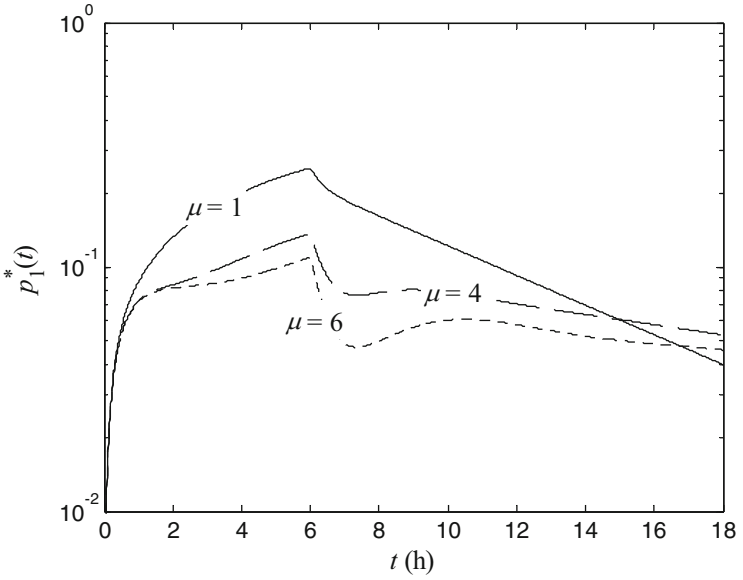


Fig. 11.14 Simulation of time–$p_1^*(t)$ profiles for $\mu = 1, 4, 6$ obtained with a 6-h infusion

11.3.3.1 Number of Particles

Let $\underline{n}_0$ be an m-dimensional deterministic vector representing the number of particles contained in the drug amount $\underline{q}_0$ initially given in each compartment. Also, let $\underline{N}_i(t)$ be an m-dimensional random vector that takes on zero and positive integer values. $\underline{N}_i(t)$ represents, at time t, the random distribution among the m compartments of the number of molecules starting in i. Since all of the molecules are independent by assumption, $\underline{N}_i(t)$ follows a multinomial distribution:

$$\underline{N}_i(t) \sim \text{multinomial}\left[n_{0i}, \underline{p}_i(t)\right], \tag{11.22}$$

where $\underline{p}_i(t)$ is the vector of state probabilities for the molecules starting in i. The expectation vector, variances, and covariances of $\underline{N}_i(t)$ have simple well-known forms:

$$\begin{aligned} E\left[\underline{N}_i(t)\right] &= n_{0i}\underline{p}_i(t), \\ Var\left[N_{ij}(t)\right] &= n_{0i}p_{ij}(t)\left[1 - p_{ij}(t)\right], \\ Cov\left[N_{ij}(t)\, N_{ik}(t)\right] &= -n_{0i}p_{ij}(t)\, p_{ik}(t). \end{aligned}$$

Particles starting in each compartment i contribute to obtaining the number of particles in each compartment:

$$\underline{N}(t) = \sum_{i=1}^{m} \underline{N}_i(t),$$

where $\underline{N}(t)$ is a random vector having expectation, variance, and covariance

$$E[\underline{N}(t)] = \sum_{i=1}^{m} n_{0i} \underline{p}_{\cdot i}(t), \tag{11.23}$$

$$Var\left[N_j(t)\right] = \sum_{i=1}^{m} n_{0i} p_{ij}(t)\left[1 - p_{ij}(t)\right],$$

$$Cov\left[N_j(t) N_k(t)\right] = -\sum_{i=1}^{m} n_{0i} p_{ij}(t) p_{ik}(t),$$

respectively.

11.3.3.2 Repeated Dosage

When drugs are given in repeated dosage, we have to compile the repeated schemes. We assume linearity in mixing multinomial distributions, i.e., if

$$\underline{N}_{ik}(t) \sim \text{multinomial}\left[n_{0ik}, \underline{p}_{\cdot i}(t)\right]$$

for $k = 1 : m_u$, then

$$\underline{N}_i(t) = \sum_{k=1}^{m_u} \underline{N}_{ik}(t) \sim \text{multinomial}\left[\sum_{k=1}^{m_u} n_{0ik}, \underline{p}_{\cdot i}(t)\right]$$

with expectation vector, variances, and covariances

$$E\left[\underline{N}_i(t)\right] = \underline{p}_{\cdot i}(t) \sum_{k=1}^{m_u} n_{0ik},$$

$$Var\left[N_{ij}(t)\right] = p_{ij}(t)\left[1 - p_{ij}(t)\right] \sum_{k=1}^{m_u} n_{0ik},$$

$$Cov\left[N_{ij}(t) N_{ik}(t)\right] = -p_{ij}(t) p_{ik}(t) \sum_{k=1}^{m_u} n_{0ik},$$

respectively. Moreover, if the m_u administrations are delayed by t_k°, one has to substitute in the previous expressions $\underline{p}_{\cdot i}(t)$ with $\underline{p}_{\cdot i}\left(t-t_k^\circ\right) H\left(t-t_k^\circ\right)$.

Example 8. Repeated Infusions for the One-Compartment Model

For the one-compartment model of (11.21), assume that n_0 particles of drug were initially in compartment 2 and then two constant-rate infusions delayed by t° were given in compartment 1. Let n_1 and n_2 be the infused amounts and T_1 and T_2 the infusion times. According to the previous relations, the expectation of the time–amount curve will be

$$E\left[N_2(t)\right] = n_0 \exp(-ht) + \frac{n_1}{T_1 h}\left\{\exp\left[-h(t-T_1) H(t-T_1)\right] - \exp(-ht)\right\}$$

$$+\frac{n_2}{T_2 h}\left\{\exp\left[-h\left(t'-T_2\right) H\left(t'-T_2\right)\right] - \exp\left[-ht' H\left(t'\right)\right]\right\}$$

with $t' = t - t^\circ$.■

11.3.3.3 Drug Amounts

Given the gram-molecular weight of the drug and using Avogadro's number, one converts the number of particles n_{0i} and $\underline{N}(t)$ to the equivalent amounts q_{0i} and $\underline{Q}(t)$, respectively. Thus, the expectation vector, variances, and covariances of the drug amount $\underline{Q}(t)$ in the compartments at time t are

$$E\left[\underline{Q}(t)\right] = \sum_{i=1}^{m} q_{0i}\underline{p}_{\cdot i}(t),$$

$$Var\left[Q_j(t)\right] = \sum_{i=1}^{m} q_{0i} p_{ij}(t)\left[1-p_{ij}(t)\right],$$

$$Cov\left[Q_j(t)\, Q_k(t)\right] = -\sum_{i=1}^{m} q_{0i} p_{ij}(t)\, p_{ik}(t),$$

respectively. In matrix notation, $E\left[\underline{Q}^T(t)\right]$ may also be written as $\underline{q}_0^T \mathbf{P}(t)$. Taking into account (11.3), the expectation of the drug amount becomes

$$E\left[\underline{Q}^T(t)\right] = \underline{q}_0^T \exp(\mathbf{H}t),$$

a similar form to that of deterministic models (8.5).

In conclusion, the solutions $E\left[\underline{Q}^T(t)\right]$ for the expected values for such stochastic models are the same as the solutions $\underline{q}^T(t)$ for the corresponding deterministic

models, and the *transfer-intensity* matrix **H** is analogous to the *fractional flow rates* matrix **K** of the deterministic model. If the hazard rates are constant in time, we have the stochastic analogues of linear deterministic systems with constant coefficients. If the hazard rates depend on time, we have the stochastic analogues of linear deterministic systems with time-dependent coefficients.

So, it is possible to associate some probabilistic interpretations in the deterministic model. From the probabilistic viewpoint $k_{ij}\Delta t$ is the conditional probability that a molecule will be transferred from i to j in the interval t to $t + \Delta t$. Thus $k_{ii}\Delta t$ is the conditional probability that a molecule leaves i in that interval.

If the hazard rate of any single particle out of a compartment depends on the state of the system, the equations of the probabilistic transfer model are still linear, but we have nonlinear rate laws for the transfer processes involved and such systems are the stochastic analogues of nonlinear compartmental systems. For such systems, the solutions for the deterministic model are not the same as the solutions for the mean values of the stochastic model.

Example 9. Two-Compartment Reversible Model

For the model presented in Section 11.2.4 and in the presence of q_{01} and q_{02} amounts of molecules at the starting time in compartments 1 and 2, respectively, the expectation of the time–amount curve in the two compartments will be the inverse Laplace transform of

$$E\left[\tilde{Q}_1(s)\right] = \frac{\left[q_{01} + q_{02}\tilde{f}_2(s)\right]\tilde{\mathcal{S}}_1(s)}{1 - \omega\tilde{f}_1(s)\tilde{f}_2(s)}$$

and

$$E\left[\tilde{Q}_2(s)\right] = \frac{\left[q_{01}\omega\tilde{f}_1(s) + q_{02}\right]\tilde{\mathcal{S}}_2(s)}{1 - \omega\tilde{f}_1(s)\tilde{f}_2(s)}.$$

■

11.3.4 Process Uncertainty or Stochastic Error

So, we find that the mean behavior of the stochastic model is described by the deterministic model we have already developed. The fundamental difference between the stochastic and the deterministic model arises from the chance mechanism in the stochastic model that generates the so-called *process uncertainty*, or *stochastic error*.

11.3.4.1 Spatial Error

The stochastic error is expressed in (11.23) by the variance $Var\left[N_j(t)\right]$ and covariance $Cov\left[N_j(t)\,N_k(t)\right]$ that did not exist in the deterministic model. This error could also be named *spatial stochastic error*, since it describes the process uncertainty among compartments for the same t and it depends on the number of drug particles initially administered into the system. For the sake of simplicity, assume $n_{0i} = n_0$ for each compartment i. From the previous relations, the coefficient of variation $CV_j(t)$ associated with a time curve $N_j(t)$ in compartment j is

$$CV_j(t) = \frac{\sqrt{Var\left[N_j(t)\right]}}{E\left[N_j(t)\right]} = \sqrt{\frac{1}{n_0}\frac{\sum_{i=1}^{m}\left[1-p_{ij}(t)\right]}{\sum_{i=1}^{m}p_{ij}(t)}}.$$

CV varies as $1/\sqrt{n_0}$ and it is not a small number for dosages involving few particles or drugs administered at very low doses; otherwise, $CV \ll 1$, as is typical in pharmacokinetics [537, 538]. From a mechanistic point of view, if the number of molecules present is not large, the concentration as a function of time will show the random fluctuations we expect from chance occurrences. However, if the number is very large, these fluctuations will be negligible, and for purposes of estimation, the stochastic error may be omitted in comparison with the measurement error.

11.3.4.2 Serial Error

An important generalization concerns the multinomial distribution of observations at different times. To deal with this, we analyze in the Markovian context the prediction of the statistical behavior of particles at time $t+t^\circ$ based on the observations at t, i.e., the state about the conditional random variable $\left[\underline{N}_i\left(t+t^\circ\right) \mid n_i(t)\right]$. As previously, in common use is the multinomial distribution

$$\left[\underline{N}_i\left(t+t^\circ\right) \mid n_i(t)\right] \sim \text{multinomial}\left[n_i(t)\,,\underline{p}_i\left(t,t+t^\circ\right)\right]$$

using the transfer probability $\underline{p}_i\left(t,t+t^\circ\right)$ with elements $p_{ij}\left(t,t+t^\circ\right)$. For the standard Markov process, the above expression is reduced to

$$\left[\underline{N}_i\left(t+t^\circ\right) \mid n_i(t)\right] \sim \text{multinomial}\left[n_i(t)\,,\underline{p}_i\left(t^\circ\right)\right], \tag{11.24}$$

where $\underline{p}_i\left(t^\circ\right)$ is the state probability with elements $p_{ij}\left(t^\circ\right)$. The conditional expectation of $E\left[\underline{N}_i\left(t+t^\circ\right) \mid n_i(t)\right]$ is $n_i(t)\,\underline{p}_i\left(t^\circ\right)$, and whatever the particles' origin,

$$E\left[\underline{N}\left(t+t^{\circ}\right) \mid \underline{n}\left(t\right)\right] = \sum_{i=1}^{m} n_i\left(t\right) \underline{p}_{i}\left(t^{\circ}\right), \tag{11.25}$$

$$Var\left[N_j\left(t+t^{\circ}\right) \mid \underline{n}\left(t\right)\right] = \sum_{i=1}^{m} n_i\left(t\right) p_{ij}\left(t^{\circ}\right)\left[1-p_{ij}\left(t^{\circ}\right)\right],$$

$$Cov\left[N_j\left(t+t^{\circ}\right) N_k\left(t+t^{\circ}\right) \mid \underline{n}\left(t\right)\right] = -\sum_{i=1}^{m} n_i\left(t\right) p_{ij}\left(t^{\circ}\right) p_{ik}\left(t^{\circ}\right).$$

The expressions (11.24) and (11.25) correspond to (11.22) and (11.23), respectively. The latter expressions can be obtained from the former ones by substituting t with 0 and t° with t. Since $\underline{N}\left(t+t^{\circ}\right)$ is conditioned to the random $\underline{n}\left(t\right)$, the total expectation theorem leads unconditionally to (cf. Appendix D)

$$E\left[\underline{N}\left(t+t^{\circ}\right)\right] = \sum_{i=1}^{m} E\left[N_i\left(t\right)\right] \underline{p}_{i}\left(t^{\circ}\right),$$

and the total variance theorem leads to

$$\begin{aligned} Var\left[N_j\left(t+t^{\circ}\right)\right] &= \sum_{i=1}^{m} E\left[N_i\left(t\right)\right] p_{ij}\left(t^{\circ}\right)\left[1-p_{ij}\left(t^{\circ}\right)\right] \\ &\quad + \sum_{i=1}^{m} Var\left[N_i\left(t\right)\right] p_{ij}\left(t^{\circ}\right), \\ Cov\left[N_j\left(t+t^{\circ}\right) N_k\left(t+t^{\circ}\right)\right] &= -\sum_{i=1}^{m} E\left[N_i\left(t\right)\right] p_{ij}\left(t^{\circ}\right) p_{ik}\left(t^{\circ}\right) \\ &\quad - \sum_{i=1}^{m} Var\left[N_i\left(t\right)\right] p_{ij}\left(t^{\circ}\right) p_{ik}\left(t^{\circ}\right). \end{aligned}$$

The covariance structure following the chain binomial distribution [330, 514] introduces a serial covariance process error [539]. It is expressed by

$$E\left[N_j\left(t\right) N_k\left(t+t^{\circ}\right) \mid \underline{n}\left(t\right)\right] = \sum_{i=1}^{m} N_j\left(t\right) N_i\left(t\right) p_{ik}\left(t^{\circ}\right),$$

and using the same unconditional approach,

$$E\left[N_j\left(t\right) N_k\left(t+t^{\circ}\right)\right] = E\left[N_j^2\left(t\right)\right] p_{jk}\left(t^{\circ}\right) + \sum_{\substack{s=1 \\ s\neq j}}^{m} E\left[N_j\left(t\right) N_i\left(t\right)\right] p_{ik}\left(t^{\circ}\right).$$

Hence for all $j = 1 : m$ and $k = 1 : m$,

$$Cov\left[N_j(t), N_k\left(t + t^\circ\right)\right] = p_{jk}\left(t^\circ\right) Var\left[N_j(t)\right] + \sum_{\substack{s=1 \\ s \neq j}}^{m} Cov\left[N_j(t), N_k(t)\right] p_{ik}\left(t^\circ\right),$$

which can be expressed in terms of the n_{0i} and $p_{ij}(t)$ $(i, j = 1 : m)$ using (11.23). This error could be named *temporal stochastic error*, since it describes the error correlation between two time instants for a couple of compartments. These results agree with the equations of Kodell and Matis [540] in the two-compartment case that they discussed. The above derivations apply equally to the time-dependent Markov process if we replace $p_{ij}(t^\circ)$ by $p_{ij}(t, t + t^\circ)$. The additional difficulties in the time-dependent case come in the computation of $p_{ij}(t_\circ, t)$. It is important to note that in the general case, the stochastic errors have slight serial correlation and hence are not independent. In the pharmacokinetic context where the number of molecules is large, the serial error may be neglected in comparison with the measurement error.

In principle, the general objective is to solve for the distribution of the random vector $\underline{N}(t)$, which might then be compared with the deterministic solution. However, the first and the second moments are sufficient for many applications using least squares procedures, since the mean value function gives the regression model and the second moments provide information useful in weighting the data and in identifying the model. Hence, one focus only on these moments, and, for simplicity, one considers only the expectations and variances. The covariance structure, where the $\underline{N}(t)$ are interrelated both temporally and serially, must be used together with the measurement error.

Finally, note also that we do not use the count of particles that have gone to the environment. This can be recovered from the original counts and the counts in the other compartments. Use of that count would introduce an exact linear dependence in the data.

11.3.5 Distribution of Particles and Process Uncertainty

To illustrate the process uncertainty, we present the case of the two-compartment model, Figure 11.1. Equation (11.5) associated with the transfer-intensity matrix **H** was used to simulate the random distribution of particles, which expresses the process uncertainty.

11.3.5.1 The Time Profile of the Distribution of Particles

After obtaining the state probabilities and setting the distributional assumptions, it is interesting to simulate the probabilistic behavior of the system, i.e., evaluate

$\Pr\left[N_j(t) = n\right], n = 0, \ldots, \infty$ and $j = 1, 2$. For a given n, $\Pr\left[N_j(t) = n\right]$ is the joint probability of the $n+1$ possible mutually exclusive events that "i particles originated in compartment 1" AND "$n - i$ particles originated in compartment 2 are present in compartment j at t" with $i = 0 : n$. Because the particles behave independently

$$\Pr\left[N_j(t) = n\right] = \sum_{i=0}^{n} \Pr\left[i; n_{01}, p_{1j}(t)\right] \Pr\left[n - i; n_{02}, p_{2j}(t)\right],$$

where $\Pr[i; n, p]$ is the binomial distribution giving the probability of obtaining i tiles among n with prior probability p.

Using hazard rates $h_{10} = 0.5$, $h_{20} = 0.1$, $h_{12} = 1$, and $h_{21} = 0.1\,\mathrm{h}^{-1}$, and initial conditions $\underline{n}_0^T = [100\ 50]$, Figures 11.15 and 11.16 show the time profile of $\Pr\left[N_j(t) = n\right]$ (the $n = 0$ levels were not shown). In these figures for a given time t and a fixed level n, the disk area is proportional to the associated probability $\Pr\left[N_j(t) = n\right]$. Thus for each t, the sum of areas is equal to 1. It is noted that a fixed n has chances to occur at several t, and for a fixed t, the probability is widespread over a range of n values. This phenomenon is the process uncertainty or stochastic error.

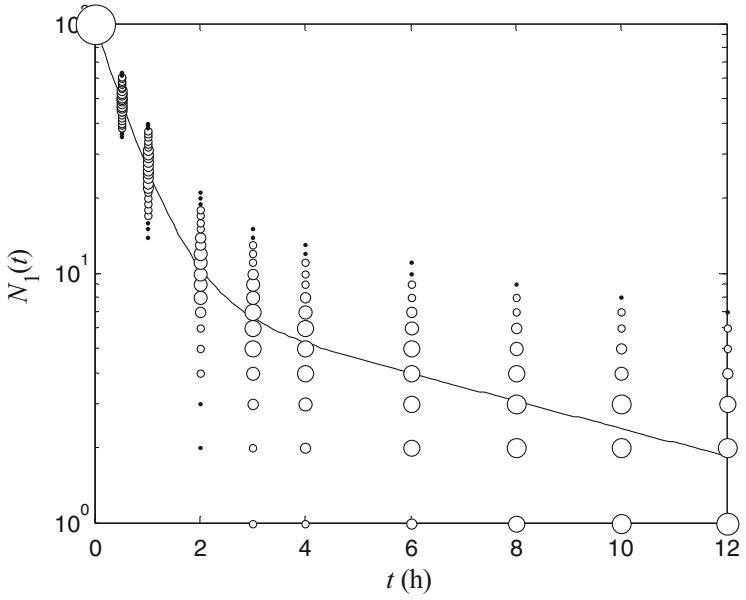


Fig. 11.15 Probabilistic behavior of the particles observed in compartment 1. The *solid line* is the solution of the deterministic model. The area of a disk located at coordinates (t, n) is proportional to $\Pr[N_1(t) = n]$

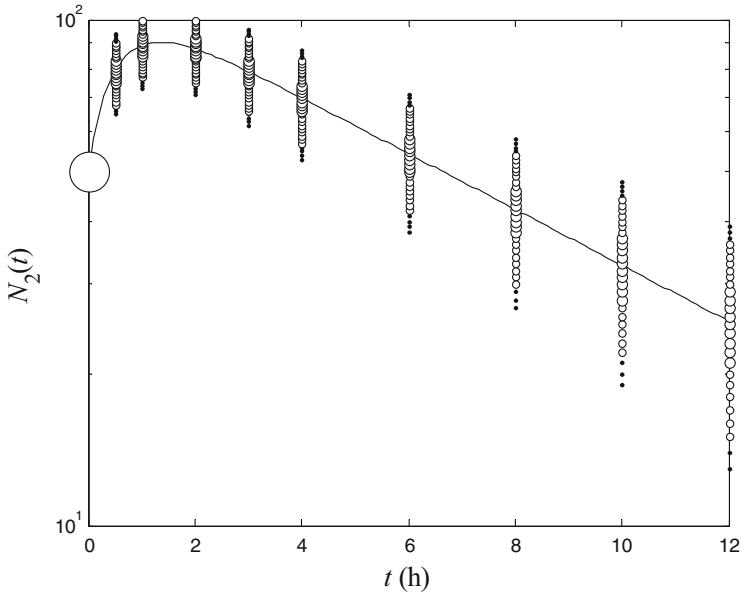


Fig. 11.16 Probabilistic behavior of the particles observed in compartment 2. The *solid line* is the solution of the deterministic model. The area of a disk located at coordinates (t, n) is proportional to $\Pr[N_2(t) = n]$

11.3.5.2 The Process Uncertainty and the Serial Correlation

Assuming as initial conditions first $\underline{n}_0^T = [10\ 5]$ and then $10\underline{n}_0$, Figures 11.17 and 11.18 illustrate:

- the time–particle-count profiles for the two compartments $E\left[\underline{N}_j(t)\right], j = 1, 2,$
- the confidence intervals computed as $E\left[\underline{N}_j(t)\right] \pm \sqrt{Var\left[N_j(t)\right]}, j = 1, 2,$ and
- random data generated from the binomial distribution, $\text{Bin}\left[n_{0i}, p_{ij}(t)\right]$, with prior probabilities computed from (11.5).

These profiles were normalized with respect to the initial condition in each compartment. The wider confidence intervals correspond to the initial conditions $\underline{n}_0$, and the narrower confidence intervals to $10\underline{n}_0$. Even without measurement error, fluctuations in the predicted amounts expressing the process uncertainty were observed: the lower the number of molecules initially present in the compartments, the higher the observed fluctuations.

For $t^\circ = 1, 2, 4$, Figure 11.19 illustrates the correlation coefficients

$$Cor\left[N_j(t), N_k\left(t + t^\circ\right)\right]$$

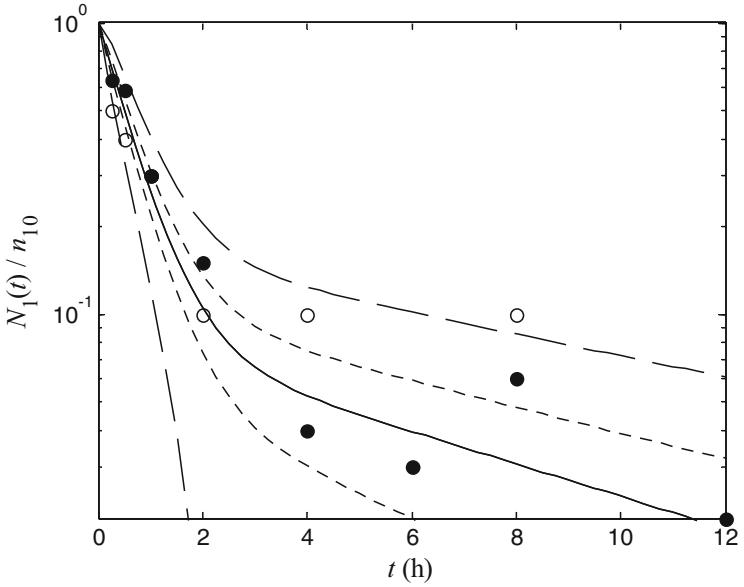


Fig. 11.17 Normalized particle-count profiles in compartment 1. *Dashed line* and *open circles* for low initial conditions, and *dotted line* and *full circles* for high initial conditions

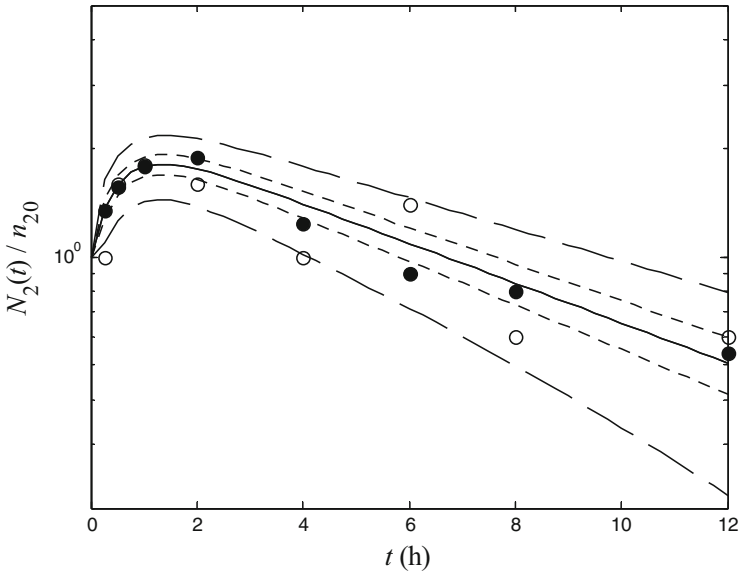


Fig. 11.18 Normalized particle-count profiles in compartment 2. Symbols as in Figure 11.17

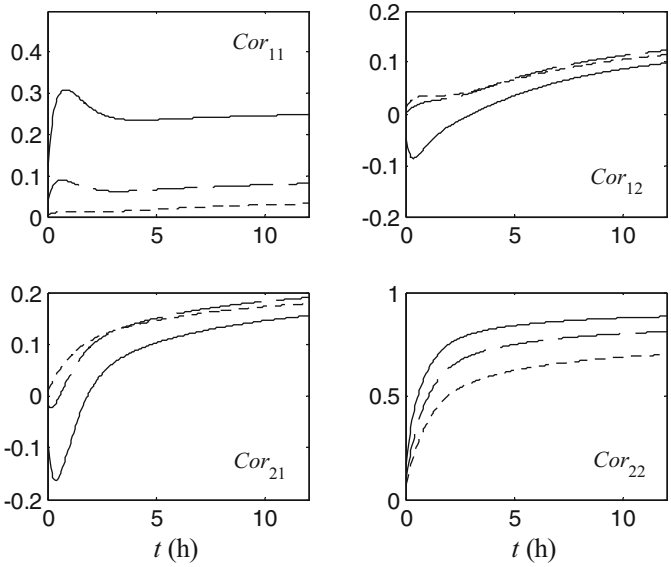


Fig. 11.19 Autocorrelations and cross-compartment serial correlations with increased values of delay $t^{\circ} = 1, 2, 4$ (*solid*, *dashed*, and *dotted lines*, respectively)

computed from covariances

$$Cov\left[N_j(t), N_k\left(t+t^{\circ}\right)\right]$$

between the same and different compartments. The autocorrelations

$$Cor\left[N_1(t), N_1(t+t^{\circ})\right] \text{ and } Cor\left[N_2(t), N_2(t+t^{\circ})\right]$$

vanish with increasing t° and they are always positive. $Cor\left[N_2(t), N_2(t+t^{\circ})\right]$ reaches high levels because particles stay longer in compartment 2 when trapped by the slow hazard rate h_{21}. The cross-correlations $Cor\left[N_j(t), N_k(t+t^{\circ})\right], j \neq k$, are low in absolute value. Sensitivity analysis reveals that the inter-compartment hazard rates h_{12} and h_{21} highly influence autocorrelations, while cross-correlations are more influenced by h_{10} and h_{20}, the elimination rates of particles to the environment.

11.3.6 Time Profiles of the Model

According to definitions (8.3) and (11.6), the relationship between clearance, volume of distribution, and hazard rate is again recalled:

$$CL(t) = V(t)\, h(t).$$

This relationship is now considered as time-dependent because of $h(t)$, the age-dependent hazard rate in the retention-time models, or because of $V(t)$, the time-varying volume of distribution. For all the above models, the time–concentration curve $E[C(t)]$ in each observed compartment is obtained by dividing $E[Q(t)]$ by $V(t)$. For the simplest one-compartment model, two different interpretations may arise:

- The volume of distribution is assumed constant. In this case,

$$E[C(t)] = \frac{E[Q(t)]}{V} = \frac{q_0 \mathcal{S}(t)}{V}$$

  and $E[C(t)]$ is directly proportional to the survival function $\mathcal{S}(t)$. Also, the clearance $CL(t)$ becomes an age-dependent parameter proportional to the hazard rate $h(t)$.
- The clearance is assumed constant. In this case,

$$E[C(t)] = \frac{E[Q(t)]}{V(t)} = q_0 \mathcal{S}(t) \frac{h(t)}{CL} = \frac{q_0}{CL} f(t)$$

  and $E[C(t)]$ is directly proportional to the density function $f(t)$. Also, the volume $V(t)$ becomes an age-dependent parameter inversely proportional to the hazard rate $h(t)$.

In other words, the expectation of the amount behaves always as the survival function $\mathcal{S}(t)$ but the expectation of the concentration behaves either as the density function $f(t)$ if CL is assumed constant, or as $\mathcal{S}(t)$ if V is assumed constant. Consequently, given a set of observed data, we may have indication that the process has a constant V if the best fitting is obtained by using the survival function. Conversely, if the best fitting is obtained by using the density function, the process is rather driven by a constant CL. Figure 11.20 simulates one-compartment retention-time models with initial conditions and compares the time–concentration curves obtained under the hypothesis of a constant V or a constant CL. It is noticeable that:

- the exponential retention-time distribution did not discriminate between the two hypotheses, and
- for the other distributions, the constant CL hypothesis yields a maximum in the time–concentration curve.

After bolus administration and keeping the CL constant, Weiss [270] obtained the simple time–concentration profile

$$c(t) = \gamma t^{-(1-\mu)} \exp(-\lambda t)$$

by assuming the $\text{Gam}(\lambda, \mu)$ retention-time distribution for the particle ages in the single compartment. Under the same conditions, Piotrovskii [541] assumed the $\text{Wei}(\lambda, \mu)$ retention-time distribution, but both models constrained the shape

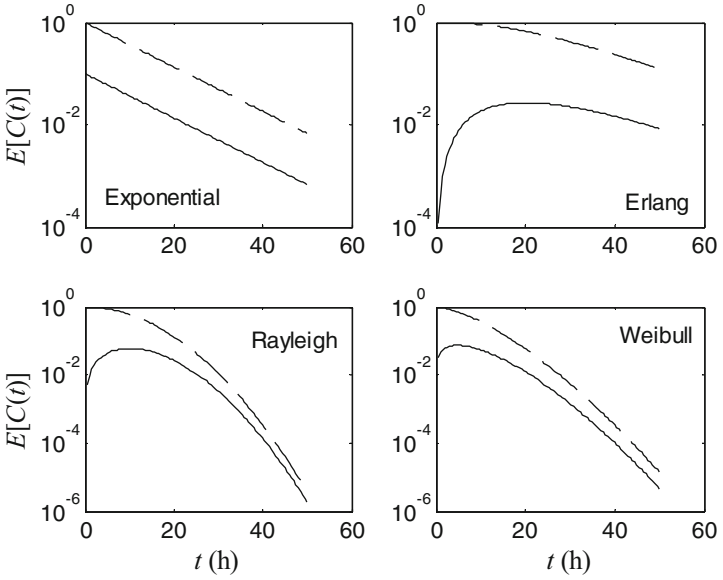


Fig. 11.20 Time–concentration curves for the hypotheses of a constant V (*dashed line*) and a constant CL (*solid line*)

parameter ($0 < \mu < 1$) in order to ensure monotonically decreasing kinetic profiles. Nevertheless, there is no indication of biological or numerical nature excluding cases with $\mu > 1$ that lead to profiles similar to those shown in Figure 11.20 and, in several cases [323], experimental data lead to negative powers of time less than -1 that contrast with the positivity of μ. Therefore, some assumptions become questionable, e.g., the simple compartmental structure, or the time-constancy of CL, or the choice for the retention-time distribution.

11.4 Random Hazard-Rate Models

In the models of the previous section, the stochastic nature of the system was due to the random movement of homogeneous individual particles. They are probabilistic transfer models or retention-time models expressing that the molecules are retained, or trapped by cells or otherwise fixed components of the process. In this way, these models express the structural heterogeneity that may originate from the time courses of particles through media that are inhomogeneous or from the retention of particles by organs in the body that are characterized by heterogeneous or fractal structures, e.g., the liver and lung.

Let us now consider a class of models that introduce particle heterogeneity through random rate coefficients. In this conceptualization, the particles are assumed

different due to variability in such characteristics as age, size, molecular conformation, or chemical composition. The hazard rates h are now considered to be random variables that vary, influenced by extraneous sources of fluctuation as though stochastic processes were added on to the hazard rates. This approach corresponds to a physiologically realistic mechanism by which the hazard rates fluctuate in an apparently random manner because of influences from other parts of the real system affecting them but that are not included in the model. The random variable h is associated with a specific probability density function $f(h)$.

Hazard rates are heterogeneous particle models expressing a functional heterogeneity. As such they contrast with probabilistic transfer and retention-time models, which assume homogeneous particles and express a structural heterogeneity. As pointed out in Chapter 7, these heterogeneities may be described by simple empirical models with time-varying parameters. Using stochastic modeling, these heterogeneities may also be expressed in a different manner. In fact, the combination of the resulting stochasticities will provide a rich collection of models. Matis and Wehrly [329] call P1 stochasticity the variability induced by structural heterogeneity, and P2 stochasticity the variability induced by functional heterogeneity.

We now consider models that combine the sources of stochastic variability identified previously [542]. The experimental context reproducing the randomness of h can be conceived as follows:

- Assume that m_0 independent *units* were introduced initially into the system with a transfer mechanism whose hazard rate h applies to *all units* in the experiment. The random movement of individual *units* in the heterogeneous process will result in a state probability $p(t,h)$ depending on the specific h of all units in that experiment. Using the binomial distribution, the *conditional* expectation and variance are

$$E[N(t) \mid h] = m_0 p(t,h),$$
$$Var[N(t) \mid h] = m_0 p(t,h)[1 - p(t,h)].$$

- Let also n_0 be replicates of the above experiment where the hazard rate varies from experiment to experiment with probability density function $f(h)$. From the previous relations, the *unconditional* expectation and variance are (cf. Appendix D)

$$E[N(t)] = E_h E[N(t) \mid h] = m_0 p(t),$$
$$Var[N(t)] = E_h Var[N(t) \mid h] + Var_h E[N(t) \mid h]$$
$$= m_0 p_S(t) + m_0^2 p_F(t),$$

with

$$p(t) = \int_h p(t,h) f(h)\, \mathrm{d}h, \tag{11.26}$$

$$p_S(t) = \int_h p(t,h)\,[1 - p(t,h)] f(h)\, \mathrm{d}h = p(t) - \int_h p^2(t,h) f(h)\, \mathrm{d}h,$$

$$p_F(t) = \int_h [p(t,h) - p(t)]^2 f(h)\, \mathrm{d}h = \int_h p^2(t,h) f(h)\, \mathrm{d}h - [p(t)]^2 .$$

The variance expression is composed of two terms: $m_0 p_S(t)$ generalizes the variance of a standard binomial distribution and is attributable to the stochastic transfer mechanism (structural heterogeneity) and $m_0^2 p_F(t)$ reflects the random nature of h (functional heterogeneity).

The random hazard rate model is easily obtained from the above by considering a *single unit*, $m_0 = 1$, and n_0 particles initially administered into the system. The first two moments are obtained by summing n_0 independent and identically distributed experiments:

$$E[N(t)] = n_0 p(t), \tag{11.27}$$

$$Var[N(t)] = n_0 [p_S(t) + p_F(t)] = n_0 p(t)\,[1 - p(t)] .$$

These relations are analogous to (11.23); the only difference is that in (11.27), $p(t)$ mixes the conditional $p(t,h)$ with the distribution $f(h)$.

11.4.1 Probabilistic Models with Random Hazard Rates

The solution of the probabilistic transfer equations leads to the exponential model (11.3). The presence of negative exponentials in the model may simplify somewhat the choice of distribution associated with the random hazard rate. In fact, the elements $p(t,h)$ of the state probability matrix $\exp(\mathbf{H}t)$ in (11.3) are exponentials, and integrating (11.27) over the random variable h, we obtain

$$\int_h \exp(-ht) f(h)\, \mathrm{d}h = \mathcal{M}(-t),$$

where $\mathcal{M}(-t)$ is the *moment generating function* of h. Hence, the parameters of the assumed "mixing" distribution $f(h)$ for the population of heterogeneous particles may be estimated directly by fitting $\mathcal{M}(-t)$ to data.

For the one-compartment model with n_0 initial conditions, the distribution of the random hazard rate h can be simply mixed with the state probability $p(t,h) = \exp(-ht)$, and relations (11.27) become

Table 11.2 Density and moment generating functions.

	$f(h)$	$\mathcal{M}(-t)$
Exp(κ)	$\kappa \exp(-\kappa h)$	$(1+t/\kappa)^{-1}$
Erl(λ, ν)	$\frac{\lambda^{\nu} h^{\nu-1}}{(\nu-1)!} \exp(-\lambda h)$	$(1+t/\lambda)^{-\nu}$
Gam(λ, μ)	$\lambda (\lambda h)^{\mu-1} \exp(-\lambda h) / \Gamma(\mu)$	$(1+t/\lambda)^{-\mu}$
Rec(α, β)	$1/\beta, 0 \leq (h-\alpha)/\beta \leq 1$	$\exp(-\alpha t)[1-\exp(-\beta t)]/(\beta t)$

$$E[N(t)] = n_0 \mathcal{M}(-t),$$
$$Var[N(t)] = n_0 \left[\mathcal{M}(-t) - \mathcal{M}^2(-t)\right].$$

By using the moment generating functions of Table 11.2, one directly obtains the following cases (the variance expressions were omitted for simplicity):

- Discrete distribution:

$$E[N(t)] = n_0 \sum_{i=1}^{m} p_i \exp(-\kappa_i t)$$

  associated with the distribution function $\Pr[h = \kappa_i] = p_i, i = 1 : m$. In principle, one could fit this model with multiple rates to data and estimate the κ_i parameters. However, in practice the estimation can be difficult even for $m = 3$, and becomes particularly hazardous for any real application with $m > 3$ [542].
- Rectangular distribution:

$$E[N(t)] = n_0 \exp(-\alpha t) \frac{1-\exp(-\beta t)}{\beta t}.$$

  This model is an analogue to the previous model with multiple rates in that the m specified fractions are replaced by a continuous rectangular rate distribution [542, 543].
- Gamma distribution:

$$E[N(t)] = n_0 (1+t/\lambda)^{-\mu}. \tag{11.28}$$

  This is the most widely applied distribution for h. When the shape parameter is an integer, one obtains the Erlang distribution. Hence, the one-compartment stochastic model leads to power-law profiles involving λ and μ parameters.

In the following, we show how to apply probabilistic transfer models with random hazard rates associated with the administration and elimination processes in a single-compartment configuration.

11.4.1.1 Hazard Rate for the Absorption Process

We report the one-compartment probabilistic transfer model receiving the drug particles by an absorption process. In this model, the elimination rate h was fixed and the absorption constant h_{ev} was random. For the stochastic context, the difference $h_{ev} - h = w$ is assumed to follow the gamma distribution, i.e., $W \sim \text{Gam}(\lambda, \mu)$ with density $f(w; \lambda, \mu)$ and $E[W] = \mu/\lambda$.

The state probability for a particle with given h_{ev} to be in the central compartment at time t is

$$\begin{aligned} p(t, w) &= \frac{h_{ev}}{h_{ev} - h} \left[\exp(-ht) - \exp(-h_{ev}t)\right] \\ &= \frac{h + w}{w} \exp(-ht) \left[1 - \exp(-wt)\right]. \end{aligned} \tag{11.29}$$

Irrespective of the individual h_{ev}, the state probability is the mixture

$$p(t) = \int_w p(t, w) f(w; \lambda, \mu) \, \mathrm{d}w.$$

But for the gamma density the following hold:

$$\begin{aligned} x f(x; \lambda, \mu) &= \frac{\mu}{\lambda} f(x; \lambda, \mu + 1), \\ \frac{1}{x} f(x; \lambda, \mu) &= \frac{\lambda}{\mu - 1} f(x; \lambda, \mu - 1), \end{aligned} \tag{11.30}$$

allowing us to compute

$$\begin{aligned} &\int_h p(t, h) f(h) \, \mathrm{d}h \\ &= \exp(-ht) \left\{ 1 - \mathcal{M}(-t; \lambda, \mu) + \frac{h\lambda}{\mu - 1} \left[1 - \mathcal{M}(-t; \lambda, \mu - 1)\right] \right\} \\ &\int_h p^2(t, h) f(h) \, \mathrm{d}h \\ &= \exp(-2ht) \left\{ 1 - 2\mathcal{M}(-t; \lambda, \mu) + \mathcal{M}(-2t; \lambda, \mu) \right. \\ &\quad + \frac{2h\lambda}{(\mu - 1)} \left[1 - 2\mathcal{M}(-t; \lambda, \mu - 1) + \mathcal{M}(-2t; \lambda, \mu - 1)\right] \\ &\quad \left. + \frac{h^2\lambda^2}{(\mu - 1)(\mu - 2)} \left[1 - 2\mathcal{M}(-t; \lambda, \mu - 2) + \mathcal{M}(-2t; \lambda, \mu - 2)\right] \right\} \end{aligned}$$

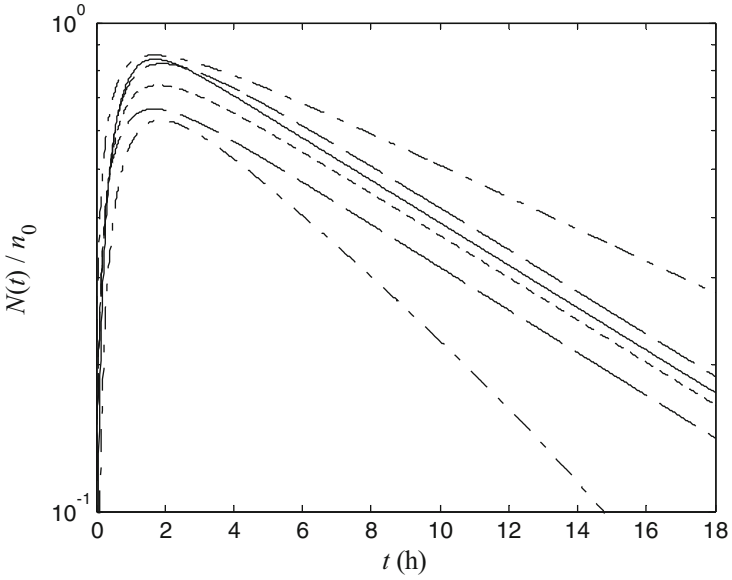


Fig. 11.21 For the absorption model, expected profile (*solid line*), confidence corridors (*mixed* and *dashed lines* for functional and structural heterogeneity, respectively), and profile with the mean coefficient value (*dotted line*)

with $\mu > 2$ and $\mathcal{M}(-t;\lambda,\mu)$ given in Table 11.2. From (11.26) and (11.27), one obtains the expected profile $E[N(t)] = n_0 p(t)$ and the standard deviations $\sqrt{n_0 p_S(t)}$ and $\sqrt{n_0 p_F(t)}$ associated with the structural and functional heterogeneity, respectively. For $n_0 = 10$, $h = 0.1\,\mathrm{h}^{-1}$, $\lambda = 1.5\,\mathrm{h}^{-1}$, and $\mu = 2.5$, Figure 11.21 shows the expected profile, the confidence corridors computed from the previous standard deviations, and the profile $n_0 p(t, \mu/\lambda)$ obtained from (11.29) using the expected value of the random variable W. All these profiles were normalized with respect to the initial condition n_0. We note the larger variability associated with the functional heterogeneity compared to that associated with the structural one, and the difference between the expected profile of the model with a random rate coefficient and the profile of the model with a fixed coefficient evaluated at the mean rate.

11.4.1.2 Hazard Rate for the Elimination Process

We present the one-compartment case in which the drug amount n_0 is given over a period T by a constant-rate infusion. Assuming a random hazard rate h over the molecules, the state probability that "a molecule associated with a hazard rate h is in the compartment at time t" is

$$p(t,h) = \frac{1}{Th} \begin{cases} 1 - \exp(-ht), & t \le T, \\ \exp[-h(t-T)] - \exp(-ht), & T < t. \end{cases} \tag{11.31}$$

Let h be gamma distributed, i.e., $H \sim \text{Gam}(\lambda, \mu)$. By using properties (11.30) with $\mu > 1$ for the gamma distribution, the resulting probability that "a molecule regardless of its hazard rate is in the compartment at time t" will be

$$p(t) = \frac{\lambda}{T(\mu-1)} \begin{cases} 1 - (1 + t/\lambda)^{-(\mu-1)}, & t \le T, \\ [1 + (t-T)/\lambda]^{-(\mu-1)} - (1 + t/\lambda)^{-(\mu-1)}, & T < t. \end{cases}$$

From (11.26) and (11.27), we obtain the expected profile and standard deviation. For $n_0 = 10$, $\lambda = 1.5\,\text{h}^{-1}$, and $\mu = 2.5$, Figure 11.22 shows the expected profile, the confidence corridors computed from the standard deviation, and the profile $n_0 p(t, \mu/\lambda)$ obtained from (11.31) using the expected value of the random variable H. All these profiles were normalized with respect to the initial condition n_0. As for the absorption process, we note the difference between the expected profile of the model with a random rate coefficient and the profile of the model with a fixed coefficient evaluated at the mean rate.

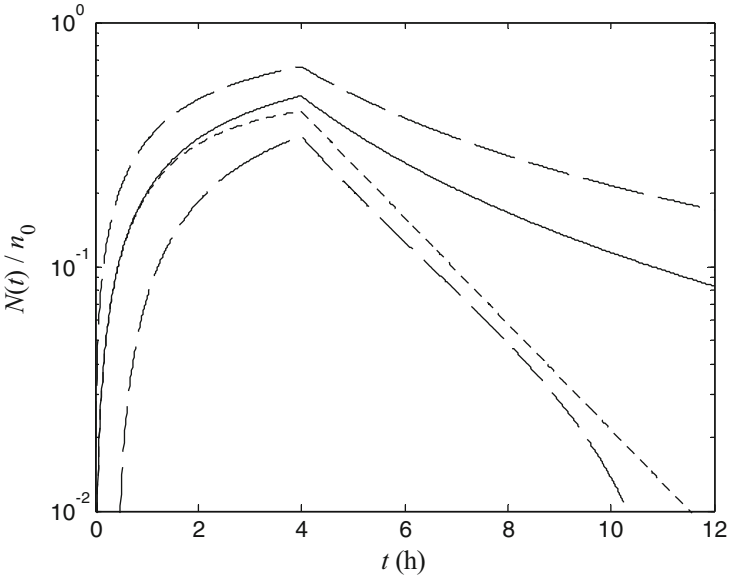


Fig. 11.22 For the elimination model, expected profile (*solid line*), confidence corridors (*dashed lines*), and profile with the mean coefficient value (*dotted line*)

For $\mu = 1$, the gamma distribution is reduced to an exponential one, and following the same procedure,

$$p(t) = \frac{\lambda}{T} \begin{cases} \ln(1 + t/\lambda), & t \leq T, \\ \ln(1 + t/\lambda) - \ln[1 + (t - T)/\lambda], & T < t. \end{cases}$$

In this case, following long-term infusion, no asymptotic behavior can be reached as t goes to infinity, i.e., no steady state exists.

Note that when the drug is given by a short infusion, i.e., $T \to 0$, the above expressions for $p(t)$ are reduced to (11.28).

Each molecule has its own hazard rate, and if we assume a constant volume of distribution V, each molecule will have its own clearance defined as $CL \triangleq Vh$. Then CL becomes a random variable, and there follows the distribution of h with expectation $E[CL] = VE[h] = V\mu/\lambda$. Regardless of the molecule's clearance, the systemic clearance may be obtained on the basis of the expected profile $E[N(t)]$ using either the plateau evaluation during a long-term infusion or the total area under the curve. Both evaluations give $CL = V(\mu - 1)/\lambda$. Note that for $\mu = 1$, the systemic clearance cannot be defined albeit individual molecular clearances exist. The discrepancy between $E[CL]$ and CL is due to the randomness of the model parameter h.

The discrepancy mentioned above in the parameter space is at the origin of the often reported discrepancy in the output space. When a rate coefficient is a random variable, the *expected amount* of a model with a *random rate* coefficient will always exceed the *amount* of a model with a fixed coefficient evaluated at the *mean rate*. It is a widespread conjecture in modeling that for systems with linear kinetics, the deterministic solution is identical to the mean value from any stochastic formulation. This conjecture, however, clearly does not hold when the rate coefficient is a random variable. In fact, the function $\exp(-kt)$ is convex, and using *Jensen's inequality* [544] we can prove that for any t,

$$E[\exp(-kt)] \geq \exp(-E[k]\, t),$$

which permits us to conclude that the kinetic profile of a homogeneous substance is always faster than that of a heterogeneous compound for which the mean rate is the same as the rate of the homogeneous one. Therefore, the mean of the stochastic model exceeds the deterministic model evaluated at the mean rates, $E[N(t)] > n(t, E[CL])$, and this is why $CL < E[CL]$.

For models using the pseudocompartment techniques to express the retention-time distribution, the same procedure as for the probabilistic transfer models can be applied to incorporate the randomness of the distribution parameter. Also, for simple situations, several assumed probability density functions of h that are rich in form yet parsimonious in parameters have been suggested by Matis [329, 543, 545]. Although these models are lengthy, they have few parameters and may be fitted to data using standard nonlinear least squares computer programs. Clearly, these models represent the union of many mechanisms that have been observed

in experimental studies to be of interest in retention-time modeling. These models have considerable appeal analytically because the parameters are identifiable, the regression functions are not necessarily monotonic, and most of the previous models are special cases of this mixture model.

11.4.2 Retention-Time Models with Random Hazard Rates

Like the previous ones, these models are two-level models. Now, the retention-time model substitutes the probabilistic transfer model in the first level, and in the second level, parameters of this model are assumed to be random and they are associated with a given distribution. Consider, for instance, the one-compartment model with Erlang retention times where the parameter λ is a random variable expressing the heterogeneity of the molecules. Nevertheless, even for the simplest one-compartment case, the model may reach extreme complexity. In these cases, analytical solutions do not exist and numerical procedures have to be used to evaluate the state probability profiles.

This approach is presented for the two-compartment model of Section 11.2.7. At the second level in (11.16), we assume that λ is a gamma-distributed random variable, $\Lambda \sim \text{Gam}(\lambda_2, \mu_2)$. The Laplace transform of the state probability is

$$p_1(t) = \frac{\lambda_2}{\Gamma(\mu_2)} \int_\lambda (\lambda_2 \lambda)^{\mu_2 - 1} \exp(-\lambda_2 \lambda) L^{-1} \left\{ \frac{1}{s + \kappa - \omega_e \kappa \left(\frac{\lambda}{s+\lambda}\right)^\mu} \right\} \mathrm{d}\lambda.$$

This expression was computed by the numerical inverse Laplace transform embedded in the numerical quadrature. As previously, we used $\mu = 1, 4, 6$, $E[\Lambda] = \mu_2/\lambda_2 = 1$, and $\lambda_2 = 4\,\mathrm{h}^{-1}$. Figure 11.23 illustrates the influence of the μ parameter on the shape of the state probability profile: the larger μ, the most pronounced the rebound form of the profile. For comparison, cf. Figure 11.10, which was obtained with fixed $\lambda = 1\,\mathrm{h}^{-1}$.

Actually, the inverse problem should be solved, i.e., given the data $n(t)$ containing errors, obtain a plausible candidate $f(h)$ associated with a known function $p(t, h)$. This function, termed *kernel*, is assumed to be a retention-time distribution other than an exponential one; otherwise, the problem has a tractable solution by means of the moment generating functions as presented earlier. This part aims to supply some indications on how to select the density of h. For a given probability density function $f(h)$, one has to mix the kernel with $f(h)$:

$$n(t) = n_0 \int_h p(t, h) f(h)\, \mathrm{d}h. \tag{11.32}$$

Equation (11.32) is a linear Fredholm integral equation of the first kind. It is also known as an *unfolding* or *deconvolution* equation. One can preanalyze the data and try to solve this first-kind integral equation. Besides the complexity of this equation, there is a paucity of numerical methods for determining the unknown function $f(h)$

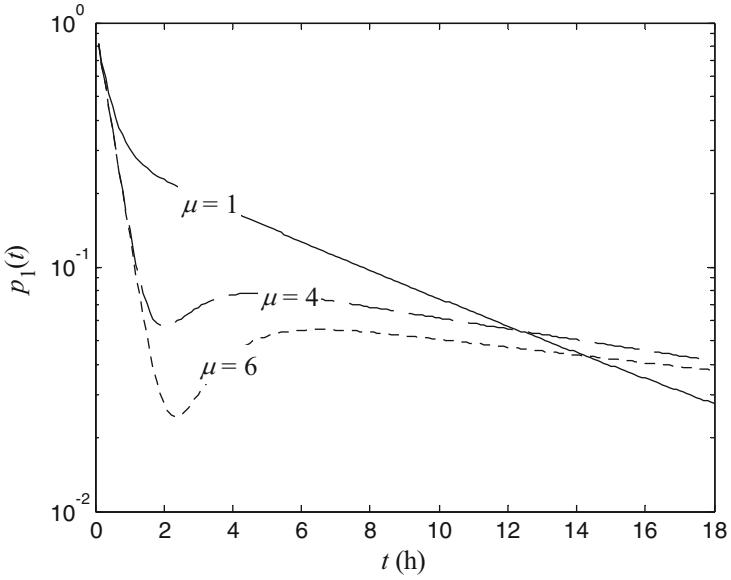


Fig. 11.23 Simulation of time–$p_1(t)$ profiles by assuming $\lambda \sim \text{Gam}(\lambda_2, \mu_2)$ and for $\mu = 1, 4, 6$

[212, 546] with special emphasis on methods based on the principle of maximum entropy [211, 547]. The so-obtained density function may be approximated by several models, gamma, Weibull, Erlang, etc., or by phase-type distributions.

Although attractive at first, there are some problems associated with the random rate models. Here, we really have a two-level stochastic model in that the parameters of the basic model contain a stochastic process. So, the first important problem is how to partition the contribution of the basic probabilistic or retention-time model, and the contribution of the random rate distribution model. This cannot be decided on the basis of empirical time–concentration data alone. The case is that of fitting a sum of exponentials model to the time–concentration data and then to assume that the number of compartments in the system is at least as large as the number of exponential terms required to achieve an acceptable fitting. This practice is inappropriate and may be very misleading when a random rate coefficient is present. Indeed, for a biphasic distribution of time–concentration data, the biexponential model is used with the common hypothesis that the underlying mechanism is a deterministic two-compartment model. But it is apparent that a one-compartment model with h having two possible outcomes has a biexponential function for its mean value. It follows therefore that one cannot imply that a multiexponential fitting of the observed mean value is sufficient evidence of a multicompartment system [542]. A second problem is related to the choice of stochastic processes to be added to a transfer coefficient. Since no transfer coefficient may ever be negative, distributions such as the normal are excluded, but log-normal, gamma, or Weibull would be acceptable.

11.5 The Kolmogorov or Master Equations

Given a compartmental structure, the probabilistic transfer, the retention-time, as well as the random hazard rate models were first conceived to express the probability of a particle transferring between compartments. In a next step and for the multinomial distribution, the model was extended to all particles administered into the system and expectations, variances, and covariances were obtained. In the last step, $\Pr\left[N_j(t) = n\right]$, $n = 0, \ldots, \infty$ and $j = 1 : m$, was obtained, i.e., the probability of having a given number of particles n in a given compartment j at time t. In a reverse way and for a given compartmental structure, one could model $\Pr\left[N_j(t) = n\right]$ and subsequently obtain the statistical distribution of particle transfer among compartments. The probabilistic transfer formulation is rather focused on compartmental modeling and describes mainly diffusion processes. Additionally, probabilistic transfer models can also be proposed for processes involving chemical, metabolic, and enzymatic reactions as well as for release, transport, and absorption phenomena. In this section, enlarged modeling concepts will be used to take into account all these processes without exclusively referring to the special case of compartmental modeling.

In a general context, suppose a given volume V contains a spatially homogeneous mixture of N_i particles from m different populations of initial size n_{0i} $(i = 1 : m)$. Suppose further that these m populations can interact through $m_\circ$ specified reaction or diffusion channels R_l $(l = 1 : m_\circ)$. These processes are assumed to be characterized by the *probability of an elementary event per unit time* that depends only on the physical properties of the diffusing or reacting particles and on the real system environment such as temperature, pressure, etc. Then, we may assert the existence of $m_\circ$ constants h_l that are the hazard rates as they were defined in (11.1) for the probabilistic transfer models. Now they are reformulated and the elementary transfer event is designated by a single index l, instead the double index ij denoting the start and end compartments.

The h_l are used to express the conditional probability of "changes in the population sizes for the R_l reaction from t to $t + \Delta t$ given the system in $\underline{n}(t)$ at t." These probabilities are described by means of the *intensity functions* $I_{\varphi_{l,1},\ldots,\varphi_{l,m}}(N_1, \ldots, N_m)$, whereby

$$\begin{aligned} &\Pr\left[N_1 \text{ changes by } \varphi_{l,1}, \ldots, N_m \text{ changes by } \varphi_{l,m}\right] \\ &= I_{\varphi_{l,1},\ldots,\varphi_{l,m}}(N_1, \ldots, N_m)\,\Delta t + o(\Delta t) \end{aligned}$$

with $\varphi_{l,i}$ denoting the changes in population i by the R_l reaction. Analysis of the mechanistic behavior of a population of reacting particles inside V leads to the intensity functions [548] of the form

$$I_{\varphi_{l,1},\ldots,\varphi_{l,m}}(N_1, \ldots, N_m) = h_l N_1^{\psi_{l,1}} \ldots N_m^{\psi_{l,m}} \tag{11.33}$$

involving the hazard rates h_l and the number of particles $\psi_{l,i}$ from the population i implied in the R_l reaction. The intensity functions are ψ_l-order elementary processes in the previous equation, with

$$\psi_l = \sum_{i=1}^{m} \psi_{l,i}.$$

The definition of the h_l hazard rates and the model of (11.33) are the only required hypotheses to formulate the stochastic movement or reaction of particles in a spatial homogeneous mixture of m-particle populations interacting through $m_\circ$ reactions.

To calculate the stochastic time evolution of the system, the key element is the *grand probability function*

$$p_{n_1,\ldots,n_m}(t) = \Pr\left[N_1(t) = n_1, \ldots, N_m(t) = n_m\right],$$

i.e., the joint probability that "there will be in the system n_1 particles of the 1-st population, ..., and n_m particles of the m-th population at time t." The abundance of particles at t can be viewed as a random vector $\underline{N}(t) = [N_1(t), \ldots, N_m(t)]^T$ and the objective is to solve for $p_{\underline{n}}(t)$, for any $t > 0$. One standard approach for solving for the grand probability function is to use equations known as the *Kolmogorov differential equations* or known also as the *master equation* in chemical engineering. This equation may be obtained by using the addition and multiplication laws of probability theory to write $p_{\underline{n}}(t + \Delta t)$ as the sum of the probabilities of the $1 + m_\circ$ different ways in which the system can arrive at the state $\underline{n}(t)$ at time $t + \Delta t$:

$$p_{\underline{n}}(t + \Delta t) = p_{\underline{n}}(t)\left[1 - \Delta t \sum_{l=1}^{m_\circ} a_l\right] + \Delta t \sum_{l=1}^{m_\circ} b_l. \tag{11.34}$$

Here we have defined the quantities a_l and b_l by

$$a_l \triangleq I_{\varphi_{l,1},\ldots,\varphi_{l,m}}(N_1, \ldots, N_m),$$
$$b_l \triangleq p_{n_1-\varphi_{l,1},\ldots,n_m-\varphi_{l,m}}(t)\, I_{\varphi_{l,1},\ldots,\varphi_{l,m}}\left(n_1 - \varphi_{l,1}, \ldots, n_m - \varphi_{l,m}\right).$$

Thus:

- The quantity $a_l \Delta t$ is the probability that "an R_l reaction occurs in Δt, given the system in $\underline{n}(t)$" and the first term in (11.34) is the probability that "the system will be in the state $\underline{n}(t)$ at time t, and then remains in that state in $(t, t + \Delta t)$."
- The quantity $b_l \Delta t$ gives the probability that "the system has one R_l reaction *removed* from the state $\underline{n}(t)$ at time t, and then undergoes an R_l reaction in $(t, t + \Delta t)$." Thus, b_l will be the product of $p_{\underline{n}}(t)$ evaluated at the appropriate once-removed state at t, TIMES the l-th intensity function evaluated in that once-removed state.

Subtracting $p_{\underline{n}}(t)$ in (11.34), dividing by Δt, and taking the limit as $\Delta t \to 0$, one has

$$\dot{p}_{\underline{n}}(t) = -p_{\underline{n}}(t) \sum_{l=1}^{m_\circ} a_l + \sum_{l=1}^{m_\circ} b_l \tag{11.35}$$

for $n_i > 0$ and the appropriate boundary conditions for each $n_i = 0$. The initial conditions are

$$p_{\underline{n}}(0) = \prod_{i=1}^{m} \delta(n_i - n_{0i}),$$

where $\delta(n - n_0)$ is the *Dirac delta function*. These equations yield the desired probability distribution for $\underline{N}(t)$. This is an infinite system of linear differential equations in the state probabilities expressed by the Kolmogorov equations. Although the system is infinite, the probabilities associated with states much larger than $\sum_i n_{0i}$ become minute.

An important property of the stochastic version of compartmental models with linear rate laws is that the mean of the stochastic version follows the same time course as the solution of the corresponding deterministic model. That is not true for stochastic models with nonlinear rate laws, e.g., when the probability of transfer of a particle depends on the state of the system. However, under fairly general conditions the mean of the stochastic version approaches the solution of the deterministic model as the number of particles increases. It is important to emphasize for the nonlinear case that whereas the deterministic formulation leads to a *finite* set of *nonlinear* differential equations, the master equation generates an *infinite* set of *linear* differential equations although the rate laws are nonlinear.

Besides the hypothesis of spatially homogeneous processes in this stochastic formulation, the particle model introduces a structural heterogeneity in the media through the scarcity of particles when their number is low. In fact, the number of differential equations in the stochastic formulation for state probability keeps track of all of the particles in the system, and therefore it accounts for particle scarcity. The presence of several differential equations in the stochastic formulation is at the origin of the uncertainty, or stochastic error, in the process. The deterministic version of the model is unable to deal with stochastic error, but as stated in Section 11.3.4, that is reduced to zero when the number of particles is very large. Only in this last case can the set of Kolmogorov differential equations be adequately approximated by the deterministic formulation, involving a set of differential equations of fixed size for the states of the process.

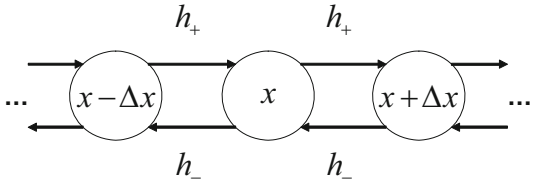


Fig. 11.24 Two-way catenary compartment model

11.5.1 Master Equation and Diffusion

As an example application, we will develop the master equation for a fragment of a two-way catenary compartment model around three compartments spaced by Δz, as illustrated in Figure 11.24. By assuming only one particle in movement, the master equation gives

$$p_{\ldots 010\ldots}(t+\Delta t) = p_{\ldots 010\ldots}(t)\left[1-\Delta t\left(h_- + h_+\right)\right] + p_{\ldots 100\ldots}(t)\, h_+ \Delta t + p_{\ldots 001\ldots}(t)\, h_- \Delta t.$$

The subscripts $\ldots 010\ldots$, $\ldots 100\ldots$, and $\ldots 001\ldots$ indicate that the particle is located in the z, $z-\Delta z$, or $z+\Delta z$ compartment, respectively. For writing convenience, we denote these probabilities by $p(z,t)$, $p(z-\Delta z,t)$, and $p(z+\Delta z,t)$. Assuming equal probabilities that the particle jumps to the nearest site to its left or right, i.e., $h_-\Delta t = h_+\Delta t = 0.5$, the previous equation becomes

$$p(z,t+\Delta t) - p(z,t) = \frac{1}{2}\left[p(z-\Delta z,t) - 2p(z,t) + p(z+\Delta z,t)\right].$$

Similarly to (2.7), we define

$$\mathcal{D} \triangleq \frac{1}{2}\frac{(\Delta z)^2}{\Delta t}.$$

Dividing these last two equations term by term, we obtain

$$\frac{p(z,t+\Delta t) - p(z,t)}{\Delta t} = \mathcal{D}\frac{p(z-\Delta z,t) - 2p(z,t) + p(z+\Delta z,t)}{(\Delta z)^2}.$$

Taking the limits $\Delta t \to 0$ and $\Delta z \to 0$ and keeping $(\Delta z)^2/\Delta t = 2\mathcal{D}$ constant in the limiting case, the previous equation gives the diffusion equation (2.18) in one dimension.

Therefore, the solution of the master equation can be thought of as a Markovian random walk in the space of reacting or diffusing species. It measures the probability $p(z,t)$ of finding the walker in a particular position z at any given time t. Furthermore, by taking into account the number of particles in the compartments,

probabilities can be converted to concentrations to obtain the second Fick's law, (2.16). If we consider an asymmetric walk where $h_- \neq h_+$, we obtain the diffusion equation with the drift velocity of the walker [549]. Moreover, if the transfer probabilities h_- and h_+ depend on the number of walkers present at a given time, the master equation corresponds to a nonlinear situation leading to anomalous diffusion, as presented in Section 2.2 for fractals and disordered media.

A good review of the master equation approach to chemical kinetics has been given by McQuarrie [550]. Jacquez [366] presents the master equation for the general m-compartment, the catenary, and the mammillary models. That author further develops the equation for the one- and two-compartment models to obtain the expectation and variance of the number of particles in the model. Many others consider the m-compartment case [510, 513, 551], and Matis [552] gives a complete methodological rule to solve the Kolmogorov equations.

In any particular case, the master equation is fairly easy to write; however, solving it is quite another matter. The number of problems for which the master equation can be solved analytically is even less than the number of problems for which the deterministic corresponding equations can be solved analytically. In addition, unlike the reaction equations (linear, nonlinear, etc.), the master equation does not readily lend itself to numerical solution, owing to the number and nature of its independent variables. In fact, the master equation is a generic form that when expanded, leads to the set of Kolmogorov differential equations whose number is equal to the product of population size for all the reactants. In short, although the master equation is both exact and elegant, it is usually not very useful for making practical numerical calculations.

We can, however, analyze these problems within the framework of the stochastic formulation by looking for an exact solution, or by using the probability generating functions, or the stochastic simulation algorithm.

11.5.2 Exact Solution in Matrix Form

For simple cases with populations of small sizes, one can express the Kolmogorov equations in a matrix form. The elements of the grand probability function $p_{n_1,\ldots,n_m}(t)$ can be considered in a vector form:

$$\begin{aligned}\underline{p}(t)^T = [&p_{0,\ldots,0}(t), \quad p_{1,\ldots,0}(t), \quad \ldots, \quad p_{n_{01},\ldots,0}(t), \quad \ldots,\\ &\ldots, \qquad \ldots, \qquad \ldots, \qquad \ldots, \qquad \ldots,\\ &p_{0,\ldots,n_{0m}}(t), p_{1,\ldots,n_{0m}}(t), \ldots, p_{n_{01},\ldots,n_{0m}}(t), \ldots].\end{aligned}$$

And the Kolmogorov equations may be written as

$$\dot{\underline{p}}(t) = \underline{p}(t)\,\mathbf{R},$$

where **R** is a constant-coefficient matrix. For models with finite **R**, e.g., when only initial conditions are present in an open system, one can proceed to find numerical solutions for the probability distributions by the direct solution of the above differential equation. In the general context, the dimension of **R** is infinite and the previous equation rules out a direct exact solution. One option for such problems is to truncate the set of differential equations for some large upper bound of population size, and then proceed to find directly a close approximate solution for the size distributions. Therefore, one can truncate the distribution of $\underline{N}(t)$ at some large upper value and solve the resulting finite system. This useful option will be illustrated in Section 11.5.5.1.

11.5.3 Cumulant Generating Functions

A very useful tool for finding analytically the distribution of $\underline{N}(t)$ is to obtain and solve partial differential equations for the associated cumulant generating functions. The *moment generating function*, denoted by $\mathcal{M}(\underline{\theta}, t)$, is defined for a multivariate integer-valued variable $\underline{N}(t)$ as

$$\mathcal{M}(\underline{\theta}, t) \triangleq \sum_{(n_1,\ldots,n_m \geq 0)} p_{n_1,\ldots,n_m}(t) \prod_{i=1}^{m} \exp(\theta_i n_i), \tag{11.36}$$

where $\underline{\theta}$ is a dummy variable. The *cumulant generating function*, denoted by $\mathcal{K}(\underline{\theta}, t)$, is defined as

$$\mathcal{K}(\underline{\theta}, t) \triangleq \log \mathcal{M}(\underline{\theta}, t) \tag{11.37}$$

with power series expansion

$$\mathcal{K}(\underline{\theta}, t) = \sum_{(s_1,\ldots,s_m \geq 0)} \kappa_{s_1,\ldots,s_m}(t) \prod_{i=1}^{m} \frac{\theta_i^{s_i}}{s_i!}. \tag{11.38}$$

This equation formally defines the joint $(s_1, \ldots, s_m)$-th cumulants, $\kappa_{s_1,\ldots,s_m}(t)$ as the coefficients in the series expansion of $\mathcal{K}(\underline{\theta}, t)$. The multiple summations in (11.36) and (11.38) on n_i and s_i, respectively, require that at least one n_i or s_i be different from 0. Hence, the approach to deriving differential equations for cumulants is simple in practice.

Statistical characteristics of the random vector $\underline{N}(t)$ can be directly obtained from cumulants $\kappa_{s_1,\ldots,s_m}(t)$ with all $s_i = 0$ except:

- $s_j = 1$ to calculate expectation $E\left[N_j(t)\right]$, or
- $s_j = 2$ to calculate variance $Var\left[N_j(t)\right]$, or
- $s_j = 1$ and $s_k = 1$ to calculate covariance $Cov\left[N_j(t) N_k(t)\right]$, etc.

There are also a number of advantages to using cumulant generating functions instead of probability or moment generating functions. For instance, in the univariate case:

- The cumulant functions provide a basis for parameter estimation using weighted least squares. The expected value function $\kappa_1(t)$ could serve as the regression function, the variance function $\kappa_2(t)$ supplies the weights, and $\kappa_3(t)$ provides a simple indicator of possible departure from an assumed symmetric distribution.
- The cumulant structure provides a convenient characterization for some common distributions:
  1. for the Poisson distribution, all cumulants are equal, i.e., $\kappa_i = c$ for all i, and
  2. for the Gaussian distribution, all cumulants above order two are zero, i.e., $\kappa_i = 0$ for $i > 2$.
- The low-order cumulants may be utilized to give saddle-point approximations of the underlying distribution [552, 553].

Partial differential equations may be written directly using an infinitesimal generator technique, called the *random-variable technique*, given in Bailey [554]. For intensity functions of the form (11.33), we define the operator notation

$$I_{\varphi_{l,1},\ldots,\varphi_{l,m}}\left(\frac{\partial}{\partial\theta_1},\ldots,\frac{\partial}{\partial\theta_m}\right)\mathcal{M}(\underline{\theta},t) = h_l\frac{\partial^{\psi_l}\mathcal{M}(\underline{\theta},t)}{\partial\theta_i^{\psi_{l,1}}\ldots\partial\theta_i^{\psi_{l,m}}}.$$

Using this notation, the moment generating function is given in [554] (p. 73):

$$\frac{\partial\mathcal{M}(\underline{\theta},t)}{\partial t} = \sum_{l=1}^{m_\circ}\left[\exp\left(\sum_{i=1}^{m}\theta_i\varphi_{l,i}\right)-1\right]I_{\varphi_{l,1},\ldots,\varphi_{l,m}}\left(\frac{\partial}{\partial\theta_1},\ldots,\frac{\partial}{\partial\theta_m}\right)\mathcal{M}(\underline{\theta},t)\,. \tag{11.39}$$

The boundary condition for this partial differential equation is obtained from (11.36). By multiplying both sides of this relationship by $\mathcal{M}^{-1}$ and using the definition of the cumulant generating function, the partial differential equation of the cumulant generating function is derived. The operator equation approach is very useful. Therefore, this approach is easily applied to density-dependent models, for which the intensity functions involve higher powers of $\underline{N}$ leading to nonlinear partial differential equations. The approach also extends to multiple populations. For cases in which analytical solutions are not available, one might solve for the cumulant of desired order instead of using series expansions and after equating coefficients of powers of $\underline{\theta}$.

For linear systems, the differential equation for the j-th cumulant function is linear and it involves terms up to the j-th cumulant. The same procedure will be followed subsequently with other models to obtain analogous differential equations, which will be solved numerically if analytical solutions are not tractable. Historically, numerical methods were used to construct solutions to the master equations, but these solutions have pitfalls that include the need to approximate

higher-order moments as a product of lower moments, and convergence issues [550]. What was needed was a general method that would solve this sort of problem, and that came with the stochastic simulation algorithm.

11.5.4 Stochastic Simulation Algorithm

A computational method was developed by Gillespie in the 1970s [548, 555] from premises that take explicit account of the fact that the time evolution of a spatially homogeneous process is a discrete, stochastic process instead of a continuous, deterministic process. This computational method, which is referred to as the *stochastic simulation algorithm*, offers an alternative to the Kolmogorov differential equations that is free of the difficulties mentioned above. The simulation algorithm is based on the *reaction probability density function* defined below.

Let us now consider how we might go about simulating the stochastic time evolution of a dynamic system. If we are given that the system is in state $\underline{n}(t)$ at time t, then essentially all we need in order to "move the system forward in time" are the answers to two questions: "when will the next random event occur," and "what kind of event will it be?" Because of the randomness of the events, we may expect that these two questions will be answered in only some probabilistic sense.

Prompted by these considerations, Gillespie [555] introduced the reaction probability density function $p(\varkappa, l)$, which is a joint probability distribution on the space of the continuous variable $\varkappa$ $(0 \leq \varkappa < \infty)$ and the discrete variable l $(l = 1 : m_\circ)$. This function is used as $p(\varkappa, l)\,\Delta\varkappa$ to define the probability that "given the state $\underline{n}(t)$ at time t, the next event will occur in the infinitesimal time interval $(t + \varkappa, t + \varkappa + \Delta\varkappa)$, AND will be an R_l event." Our first step toward finding a legitimate method for assigning numerical values to $\varkappa$ and l is to derive, from the elementary conditional probability $h_l \Delta t$, an analytical expression for $p(\varkappa, l)$. To this end, we now calculate the probability $p(\varkappa, l)\,\Delta\varkappa$ as the product $p_0(\varkappa)$, the probability at time t that "no event will occur in the time interval $(t, t + \varkappa)$" TIMES $a_l \Delta\varkappa$, the subsequent probability that "an R_l event will occur in the next differential time interval $(t + \varkappa, t + \varkappa + \Delta\varkappa)$":

$$p(\varkappa, l)\,\Delta\varkappa = p_0(\varkappa)\, a_l \Delta\varkappa. \tag{11.40}$$

The probability of more than one reaction occurring in $(t + \varkappa, t + \varkappa + \Delta\varkappa)$ is $o(\Delta\varkappa)$.

In order to appreciate $p_0(\varkappa)$, the probability that "no event occurs in $(t, t + \varkappa)$," imagine the interval $(t, t + \varkappa)$ to be divided into L subintervals of equal length $\varepsilon = \varkappa / L$. The probability that none of the events $R_1, \ldots, R_{m_\circ}$ occurs in the first ε subinterval $(t, t + \varepsilon)$ is

$$\prod_{\nu=1}^{m_\circ} [1 - a_\nu \varepsilon + o(\varepsilon)] = 1 - \sum_{\nu=1}^{m_\circ} a_\nu \varepsilon + o(\varepsilon) = 1 - a_0 \varepsilon + o(\varepsilon)$$

if we put $a_0 \triangleq \sum_{\nu=1}^{m_\circ} a_\nu$. This is also the subsequent probability that no event occurs in $(t + \varepsilon, t + 2\varepsilon)$, and then in $(t + 2\varepsilon, t + 3\varepsilon)$, and so on. Since there are L such ε subintervals between t and $t + \varkappa$, then $p_0(\varkappa)$ can be written as

$$p_0(\varkappa) = [1 - a_0 \varepsilon + o(\varepsilon)]^L = \left[1 - \frac{a_0 \varkappa}{L} + o\left(L^{-1}\right)\right]^L .$$

This is true for any $L > 1$, and in particular, for infinitely large L. By using the limit formula for the exponential function,

$$\lim_{\nu \to \infty} \left(1 - \frac{x}{\nu}\right)^\nu = \exp(-x),$$

the probability $p_0(\varkappa)$ becomes

$$\lim_{L \to \infty} p_0(\varkappa) = \exp(-a_0 \varkappa).$$

Inserting the previous expression in (11.40), we arrive at the following exact expression for the reaction probability density function:

$$p(\varkappa, l) = a_l \exp(-a_0 \varkappa). \tag{11.41}$$

Thus, we observe that $p(\varkappa, l)$ depends, through the quantity in the exponential, on the parameters for all events (not just R_l) and on the current sizes of populations for all particles (not just the R_l reactants).

Even though it may be impossible to solve a complicated dynamic system exactly, Gillespie's method can be used to numerically simulate the time evolution of the system [548]. In this method, implied events are thought of as occurring with certain probabilities, and the events that occur change the probabilities of subsequent events. This stochastic simulation algorithm has been shown to be physically and mathematically well grounded from a kinetic point of view, and rigorously equivalent to the spatial homogeneous master equation, yet surprisingly simple and straightforward to implement on a computer [548, 550, 555]. In the limit of large numbers of reactant molecules, the supplied results are entirely equivalent to the solution of the traditional kinetic differential equations derived from the mass balance law [548]. As presented here, the stochastic simulation algorithm is applicable only to spatially homogeneous systems. Work toward extending the algorithm to accommodate particle diffusion in spatially heterogeneous systems is currently in progress.

For most macroscopic dynamic systems, the neglect of correlations and fluctuations is a legitimate approximation [550]. For these cases the deterministic and

stochastic approaches are essentially equivalent, and one is free to use whichever approach turns out to be more convenient or efficient. If an analytical solution is required, then the deterministic approach will always be much easier than the stochastic approach. For systems that are driven to conditions of instability, correlations and fluctuations will give rise to transitions between nonequilibrium steady states and the usual deterministic approach is incapable of accurately describing the time behavior. On the other hand, the stochastic simulation algorithm is directly applicable to these studies.

11.5.4.1 Implementation

This algorithm can easily be implemented in an efficient modularized form to accommodate quite large reaction sets of considerable complexity [555]. For an easy implementation, the joint distribution can be broken into two disjoint probabilities using Bayes's rule $p(\varkappa, l) = p(\varkappa)p(l \mid \varkappa)$. But note that $p(\varkappa)$ may be considered as the marginal probability of $p(\varkappa, l)$, i.e.,

$$p(\varkappa) = \sum_{l=1}^{m_\circ} p(\varkappa, l),$$

and substituting this into (11.41) leads to values for its component parts:

$$p(\varkappa) = a_0 \exp(-a_0 \varkappa) \tag{11.42}$$

and

$$p(l \mid \varkappa) = \frac{a_l}{a_0}. \tag{11.43}$$

Given these fundamental probability density functions, the following algorithm can be used to carry out the reaction set simulation:

- Initialization:
  1. Set values for the h_l.
  2. Set the initial number n_{0i} of the m reactants.
  3. Set $t = 0$, and select a value for t_{sim}, the maximum simulation time.
- Loop:
  1. Compute the intensity functions a_l and $a_0 \triangleq \sum_{\nu=1}^{m_\circ} a_\nu$.
  2. Generate two random numbers r_1 and r_2 from a uniform distribution on $[0, 1]$.
  3. Compute the next time interval $\varkappa = \ln(1/r_1)/a_0$. Draw from the probability density function (11.42).
  4. Select the reaction to be run by computing l such that

$$\sum_{\nu=1}^{l-1} a_\nu < r_2 a_0 \leq \sum_{\nu=1}^{l} a_\nu .$$

Draw from the probability density function (11.43).

5. Adjust $t = t + \varkappa$ and update the n_l values according to the R_l reaction that just occurred.
6. If $t > t_{\text{sim}}$, then terminate. Otherwise, go to 1.

By carrying out the above procedure from time 0 to time t_{sim}, we evidently obtain only one possible realization of the stochastic process. In order to get a statistically complete picture of the temporal evolution of the system, we must actually carry out several independent realizations or "runs." These runs must use the same initial conditions but different starting numbers for the uniform random number generator in order for the algorithm to result in different but statistically equivalent chains. If we make K runs in all, and record the population sizes $n_i(k,t)$ in run k at time t ($i = 1:m$ and $k = 1:K$), then we may assert that the average number of particles at time t is

$$\overline{n}_i(t) \approx \frac{1}{K} \sum_{k=1}^{K} n_i(k,t),$$

and the fluctuations that may reasonably be expected to occur about this average are

$$s_i(t) \approx \left\{ \frac{1}{K} \sum_{k=1}^{K} [n_i(k,t) - \overline{n}_i(t)]^2 \right\}^{1/2}.$$

The approximately equal signs in the previous relations become equality signs in the limit $K \to \infty$. However, the fact that we obviously cannot pass to this limit of infinitely many runs is not a practical source of difficulty. On the one hand, if $s_i(t) \ll \overline{n}_i(t)$, then the results $n_i(k,t)$ will not vary much with k; in that case the estimate of $\overline{n}_i(t)$ would be accurate even for $K = 1$. On the other hand, if $s_i(t) \gtrsim \overline{n}_i(t)$, a highly accurate estimate of $\overline{n}_i(t)$ is not necessary. Of more practical significance and utility in this case would be the approximate range over which the numbers $n_i(k,t)$ are scattered for several runs k. In practice, somewhere between 3 and 10 runs should provide a statistically adequate picture of the state of the system at time t.

The computer storage space required by the simulation algorithm is quite small. This is an important consideration, since charges at most large computer facilities are based not only on how long a job runs but also on how much memory storage is used.

Because the speed of the stochastic simulation algorithm is linear with respect to the number of reactions, adding new reaction channels will not greatly increase the simulation runtime, i.e., doubling either the number of reactions or the number of

reactant species, doubles (approximately) the total algorithm runtime. The algorithm speed depends more on the number of molecules. This is seen by noting that the computation of the next time interval in $\varkappa = \ln(1/r_1)/a_0$ depends on the reciprocal of a_0, a term representing the number of molecules in the simulation. If the reaction set contains at least one second-order reaction, then a_0 will contain at least one multiplication product of two species in the population. In this case the simulation speed will fall off like the reciprocal of the square of the population. Recent improvements to the algorithm are helping to keep the runtime in check [556, 557].

11.5.4.2 Extensions

The simulation algorithm might allow one to deal in an approximate way with spatial heterogeneities. The basic idea is to divide the volume V into a number of subvolumes V_μ ($\mu = 1 : M$) in such a way that spatial homogeneity may be assumed within each subvolume. Each subvolume V_μ would then be characterized by its own (uniformly distributed) particle populations $N_{1\mu}(t), \ldots, N_{m\mu}(t)$ and also a set of hazard rates $h_{l\mu}$ appropriate to the physicochemical characteristics inside V_μ. For instance, in order to apply the simulation algorithm to a collection of cells, the original algorithm must be extended to accommodate the introduction of spatial dependencies of the concentration variables. Introducing the spatial context into the stochastic simulation algorithm using the subvolumes V_μ may be materialized by a rectangular array of square cells with only nearest-neighbor, cell–cell interactions. In this model of interacting cells, it is assumed that each cell is running its own internal program of biochemical reactions.

The fact that simulation of any given reaction generates its own "local" simulation time steps poses the problem of synchronization of the internal simulation times of cells. In the simplest case with no specific interaction affecting the order of the reaction, converting the algorithm from what is essentially a spatial-scanning method to a temporal-scanning method can solve this problem. This is accomplished by first making an initial spatial scan through all of the cells in the array, and inserting the cells into a priority queue that is ordered from shortest to longest local cell time. All succeeding iterations are then based on the temporal order of the cells in the priority queue. In other words, a cell is drawn from the queue, calculations are performed on the reaction set for that cell, and then the cell is placed back on the queue in its new temporally ordered position.

The use of a priority queue to order the cells was a unique innovation, and it solves the synchronizing problem inherent in a multicellular situation. Not only does this allow an easy mechanism for intercellular signaling: this methodology can also readily accommodate local inhomogeneities in molecular populations. Work has been done that extends the stochastic simulation algorithm to reaction-diffusion processes, and the modification to the method is straightforward. Diffusion is considered to be just another possible chemical event with an associated probability

[558]. As with all the other chemical events, diffusion is assumed to be intracellular and the basic idea behind this approach is incorporated into the simulation.

The justification for using the stochastic approach, as opposed to the simpler mathematical deterministic approach, was that the former presumably took account of *fluctuations* and *correlations*, whereas the latter did not. Oppenheim et al. [559] subsequently demonstrated that the stochastic formulation reduces to the deterministic formulation in the thermodynamic limit (wherein the size of particle populations and the containing volume all approach infinity in such a way that the particle concentrations approach finite values). Experience indicates that for most systems, the constituent particle populations need to have sizes only in the hundreds or thousands in order for the deterministic approach to be adequate. Thus, for most systems the differences between the deterministic and stochastic formulations are purely academic, and one is free to use whichever formulation turns out to be more convenient or efficient. However, near state instabilities in certain nonlinear systems, fluctuations, and correlations can produce dramatic effects, even for a huge number of particles [560]. For these systems the stochastic formulation would be the more appropriate choice.

Among the three presented approaches to solve the Kolmogorov or master equations, the partial differential equations for cumulant generating functions are most adequate for the estimation problem. The exact solution using the **R** matrix can never be applied in a real context because of the astronomic requirement of memory for storing and matrix processing operations. The stochastic simulation algorithm is an elegant tool for simulating and analyzing the system, but as a nonparametric approach it is not adequate for the estimation problem.

11.5.5 Simulation of Linear and Nonlinear Models

The two-compartment model and the model of the enzymatic reaction (cf. Sections 11.1.2 and 8.5.1, respectively) will be presented as typical cases for linear and nonlinear models, respectively. For these simulations, the model parameters were set as follows:

- For the compartmental system: $h_{10} = 0.5$, $h_{20} = 0.1$, $h_{12} = 1$, $h_{21} = 0.1\,\mathrm{h}^{-1}$.
- For the enzyme reaction: $k_{+1} = 1$, $k_{-1} = 0.5$, $k_{+2} = 1\,\mathrm{h}^{-1}$.

11.5.5.1 Exact Solution

Initial conditions for the compartmental model and the enzymatic reaction were set to $\underline{n}_0^T = [10\ 5]$, and $s_0 = 10$, $e_0 = 5$, and $c_0 = 0$, respectively. These values are very low regarding the experimental reality, but they were deliberately chosen as such to facilitate the computation of the exact solution.

Two-Compartment Model First, we develop full probabilistic transfer modeling. Consider the number of particles in the first and second compartments being n_1 and n_2, respectively, at time $t + \Delta t$, where Δt is some small time interval. There are a number of mutually exclusive ways in which this event could have come about, starting from time t. Specifically, they are:

- to have size (n_1, n_2) at time t with no change from t to $t + \Delta t$,
- to have size $(n_1 + 1, n_2)$ at time t with only a single irreversible elimination by the h_{10} way in the next interval Δt,
- to have size $(n_1, n_2 + 1)$ at time t with only a single irreversible elimination by the h_{20} way in the next interval Δt,
- to have size $(n_1 + 1, n_2 - 1)$ at time t with only a single reversible particle transfer to compartment 2 by the h_{12} way in Δt,
- to have size $(n_1 - 1, n_2 + 1)$ at time t with only a single reversible particle transfer to compartment 1 by the h_{21} way in Δt, and
- other ways that involve two or more independent changes of unit size in the interval Δt.

Because this set of mutually exclusive "pathways" to the desired event at $t + \Delta t$ is exhaustive, the probability of size n_1, n_2 at $t + \Delta t$ may be written as the sum of the individual probabilities of these pathways. Symbolically, using the assumptions for possible changes, one has for suitably small Δt,

$$\begin{aligned} p_{n_1,n_2}(t+\Delta t) = {} & p_{n_1,n_2}(t)\left[1 - h_{10}n_1\Delta t - h_{20}n_2\Delta t - h_{12}n_1\Delta t - h_{21}n_2\Delta t\right] \\ & + p_{n_1+1,n_2}(t)\left[h_{10}(n_1+1)\Delta t\right] \\ & + p_{n_1,n_2+1}(t)\left[h_{20}(n_2+1)\Delta t\right] \\ & + p_{n_1+1,n_2-1}(t)\left[h_{12}(n_1+1)\Delta t\right] \\ & + p_{n_1-1,n_2+1}(t)\left[h_{21}(n_2+1)\Delta t\right] + o(\Delta t), \end{aligned}$$

where $o(\Delta t)$ denotes terms of higher order than Δt associated with multiple independent changes. Subtracting $p_{n_1,n_2}(t)$, dividing by Δt, and taking the limit as $\Delta t \to 0$, one has

$$\begin{aligned} \dot{p}_{n_1,n_2}(t) = {} & h_{10}\left[(n_1+1)p_{n_1+1,n_2}(t) - n_1 p_{n_1,n_2}(t)\right] \\ & + h_{20}\left[(n_2+1)p_{n_1,n_2+1}(t) - n_2 p_{n_1,n_2}(t)\right] \\ & + h_{12}\left[(n_1+1)p_{n_1+1,n_2-1}(t) - n_1 p_{n_1,n_2}(t)\right] \\ & + h_{21}\left[(n_2+1)p_{n_1-1,n_2+1}(t) - n_2 p_{n_1,n_2}(t)\right] \end{aligned} \tag{11.44}$$

for $n_1, n_2 > 0$, with boundary conditions for either $n_1 = 0$ or $n_2 = 0$. Initial conditions are $p_{n_{01},n_{02}}(0) = 1$ and $p_{n_1,n_2}(0) = 0$ for $n_1 \neq n_{01}$ and $n_2 \neq n_{02}$. The solution of this set of differential equations yields the desired probability distribution for $[N_1(t), N_2(t)]^T$.

Second, considering now the exchange processes between compartment and environment as the set of first-order reactions,

$$\begin{aligned} N_1 &\xrightarrow{h_{10}} \text{environment}, & I_{-1,0} &= h_{10}N_1, \\ N_2 &\xrightarrow{h_{20}} \text{environment}, & I_{0,-1} &= h_{20}N_2, \\ N_1 &\xrightarrow{h_{12}} N_2, & I_{-1,1} &= h_{12}N_1, \\ N_2 &\xrightarrow{h_{21}} N_1, & I_{1,-1} &= h_{21}N_2, \end{aligned} \tag{11.45}$$

we can obtain the master equation (11.44) directly from the (11.35) formulation. For this model, there were two interacting populations ($m = 2$) and four ($m_\circ = 4$ in equation 11.45) intensity functions. Since only one particle from each population was implied in these intensity functions, all $\psi_{l,i}$ exponents were equal to one.

The possible states in each compartment are $n_{01} + n_{02} + 1$. Therefore $\mathbf{R}$ is a 256-dimensional matrix. The initial condition for the master equation is $p_{10,5}(0) = 1$. Figures 11.25 and 11.26 show the associated probabilities for each state as functions of time for the central and peripheral compartments, respectively. In these figures the disk area is proportional to the associated probability, the full markers are the expected values, and the solid lines the solution of the deterministic model. As already mentioned, we note that the expectation of the stochastic model follows the time profile of the deterministic system.

Enzymatic Reaction The usual stochastic approach begins by focusing attention on the probability function $p_{s,e,c}(t)$, which is defined to be the probability of finding

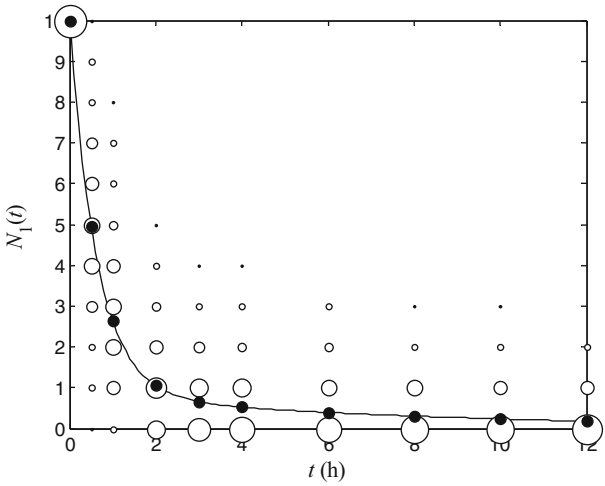


Fig. 11.25 The exact solution of the Kolmogorov equations associating marginal probabilities with the number of particles in compartment 1. The *solid line* is the solution of the deterministic model. The areas of disks located at coordinates (t, n_1) are proportional to $p_{n_1}(t)$

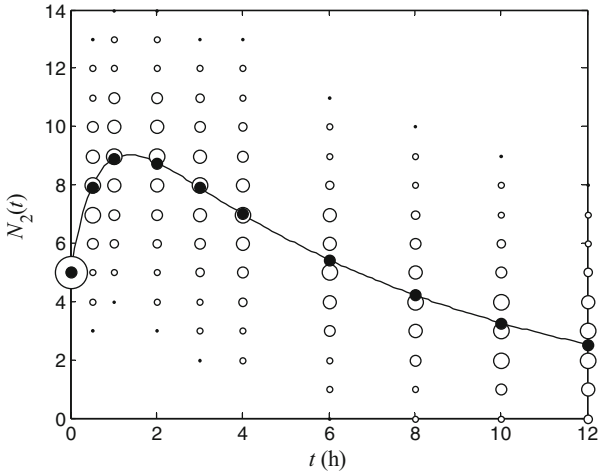


Fig. 11.26 The exact solution of the Kolmogorov equations associating marginal probabilities with the number of particles in compartment 2. The *solid line* is the solution of the deterministic model. The areas of disks located at coordinates (t, n_2) are proportional to $p_{n_2}(t)$

s molecules of substrate S, e molecules of enzyme E, and c molecules of complex C at time t. From (8.7), the intensity functions are

$$\begin{aligned} I_{-1,-1,1} &= k_{+1} se, \\ I_{1,1,-1} &= k_{-1} c, \\ I_{0,1,-1} &= k_{+2} c. \end{aligned}$$

From the conservation law of enzyme sites, $e = e_0 - c$, the enzyme population size can be substituted with the previous relation involving the initial enzyme amount e_0 and the current complex population size c. The intensity functions become

$$\begin{aligned} I_{-1,1} &= k_{+1} s\,(e_0 - c)\,, \\ I_{1,-1} &= k_{-1} c, \\ I_{0,-1} &= k_{+2} c, \end{aligned} \tag{11.46}$$

and applying the standard rules of probability theory and the (11.35) formulation, it is a straightforward matter to deduce the master equation:

$$\begin{aligned} \dot{p}_{s,c}(t) = {} & k_{+1}\left[(s+1)(e_0 - c + 1)\, p_{s+1,c-1}(t) - s\,(e_0 - c)\, p_{s,c}(t)\right] \\ & + k_{-1}\left[(c+1)\, p_{s-1,c+1}(t) - c p_{s,c}(t)\right] \\ & + k_{+2}\left[(c+1)\, p_{s,c+1}(t) - c p_{s,c}(t)\right]. \end{aligned} \tag{11.47}$$

In principle, this time-evolution equation can be solved subject to the given initial condition $p_{s,c}(0) = \delta(s - s_0)\,\delta(c - c_0)$ to obtain $p_{s,c}(t)$ uniquely for all $t > 0$. The number of product molecules can be recovered as $s_0 - (s + c)$. For a given s_0 and e_0, a computer solution is constrained not only by run time but also by the amount of computer memory that would be required just to store the current values of the function $p_{s,c}(t)$ on the two-dimensional integer lattice space of the variables S and C. The master equation may be solved exactly only when s_0 and e_0 are small. For this model, there were two interacting populations ($m = 2$), and three ($m_\circ = 3$, in equation 11.46) intensity functions. Since only one particle from each population was implied in these intensity functions, all $\psi_{l,i}$ exponents were equal to one.

The possible states for substrate are 11 and 6 for the complex. **R** is a 66-dimensional matrix and the initial condition for the master equation is $p_{10,0}(0) = 1$. Figures 11.27 and 11.28 show the associated probabilities for each state as functions of time for the substrate and the complex, respectively. As previously, the full markers are the expected values and the solid lines the solution of the deterministic model. Notably, the expectation of the stochastic model does not follow the time profile of the deterministic system. This is the main characteristic of nonlinear systems.

11.5.5.2 Cumulant Generating Functions

To illustrate how to proceed using the cumulant generating functions, the well-known two-compartment model and the enzymatic reaction will be presented as

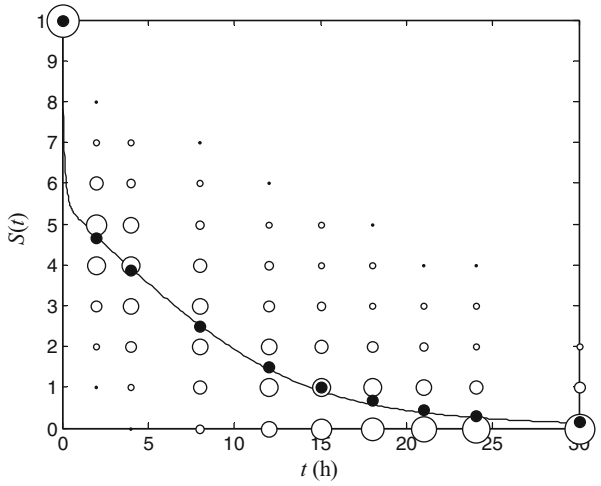


Fig. 11.27 The exact solution for the substrate, $s(t)$, of the Kolmogorov equations associating marginal probabilities with the number of particles. The *solid line* is the solution of the deterministic model. The areas of disks located at coordinates (t, s) are proportional to $p_s(t)$

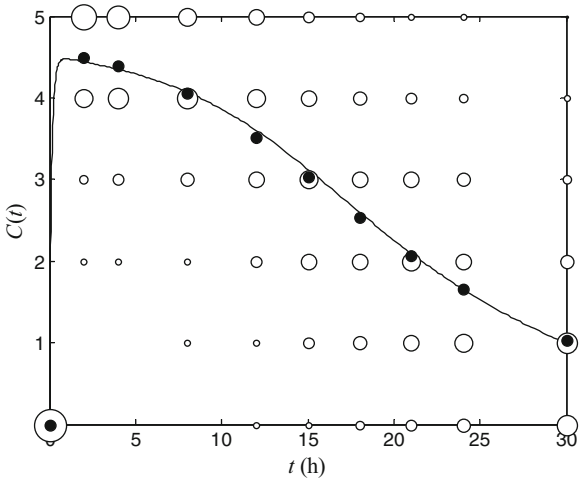


Fig. 11.28 The exact solution for the complex, $c\,(t)$, of the Kolmogorov equations associating marginal probabilities with the number of particles. The *solid line* is the solution of the deterministic model. The areas of disks located at coordinates (t, c) are proportional to $p_c\,(t)$

examples of linear and nonlinear systems, respectively. In these examples, there are two interacting populations ($m = 2$) and the cumulant generating function is

$$\mathcal{K}\,(\theta_1, \theta_2, t) = \sum_{i,j \geq 0} \kappa_{ij}\,(t)\,\frac{\theta_1^i \theta_2^j}{i!j!} \tag{11.48}$$

Initial conditions for the compartmental model and the enzymatic reaction were set to $\underline{n}_0^T = [100\ 50]$, and $s_0 = 100$, $e_0 = 50$, and $c_0 = 0$, respectively. These values are higher than those used previously and they are more likely to resemble experimental reality.

Two-Compartment Model The model assumptions in (11.45) were substituted directly into operator equation (11.39), which was transformed via (11.37) to yield

$$\begin{aligned}\frac{\partial \mathcal{K}}{\partial t} &= \{h_{10}\,[\exp\,(-\theta_1) - 1] + h_{12}\,[\exp\,(-\theta_1 + \theta_2) - 1]\}\,\frac{\partial \mathcal{K}}{\partial \theta_1} \\ &\quad + \{h_{20}\,[\exp\,(-\theta_2) - 1] + h_{21}\,[\exp\,(\theta_1 - \theta_2) - 1]\}\,\frac{\partial \mathcal{K}}{\partial \theta_2}.\end{aligned}$$

Upon substituting the series expansion (11.48) into the previous equation and equating coefficients of θ_1 and θ_2, one has the following differential equations for the cumulant functions:

$$\begin{aligned}\dot{\kappa}_{10} &= -\,(h_{10} + h_{12})\,\kappa_{10} + h_{21}\kappa_{01}, \\ \dot{\kappa}_{01} &= h_{12}\kappa_{10} - (h_{20} + h_{21})\,\kappa_{01},\end{aligned}$$

which are stochastic analogues of the deterministic formulation (8.4) and of the probabilistic transfer model (11.4). The equations for higher-order cumulants were obtained by equating coefficients of second- and third-order terms of $\theta_1^i \theta_2^j$. Especially for the second cumulants, the equations are

$$\begin{aligned}
\dot{\kappa}_{20} &= -2\left(h_{10}+h_{12}\right)\kappa_{20} + 2h_{21}\kappa_{11} + \left(h_{10}+h_{12}\right)\kappa_{10} + h_{21}\kappa_{01}, \\
\dot{\kappa}_{02} &= -2\left(h_{20}+h_{21}\right)\kappa_{02} + 2h_{12}\kappa_{11} + \left(h_{20}+h_{21}\right)\kappa_{01} + h_{12}\kappa_{10}, \\
\dot{\kappa}_{11} &= -\left(h_{10}+h_{12}+h_{20}+h_{21}\right)\kappa_{11} + h_{12}\left(\kappa_{20}-\kappa_{10}\right) + h_{21}\left(\kappa_{02}-\kappa_{01}\right).
\end{aligned}$$

Since $p_{n_{01},n_{02}}(0) = 1$ and $p_{n_1,n_2}(0) = 0$ for $n_1 \neq n_{01}$ and $n_2 \neq n_{02}$, from (11.36) and (11.38) one has $\mathcal{M}(\theta_1,\theta_2,t) = \exp\left(\theta_1 n_{01} + \theta_2 n_{02}\right)$ and $\mathcal{K}(\theta_1,\theta_2,t) = \theta_1 n_{01} + \theta_2 n_{02}$, respectively. Initial conditions for the cumulant differential equations are obtained by equating $\mathcal{K}(\theta_1,\theta_2,t)$ with the terms of the power expansion in (11.48): $\kappa_{ij}(0) = 0$ except for $\kappa_{10}(0) = n_{01}$ and $\kappa_{01}(0) = n_{02}$.

Simulations of these equations confirm the probabilistic behavior and the time profile of the distribution of particles that were already shown in Section 11.3.5 in Figures 11.15 and 11.16. Through the $\kappa_{11}(t)$ profile, this analysis reveals the statistical independence of the population sizes. Since the system has no entry, the two variables are negatively linked with an extreme value about 0.5 h as shown in Figure 11.29. This link is stronger when h_{12} and h_{21} are high compared to h_{10} and h_{20}.

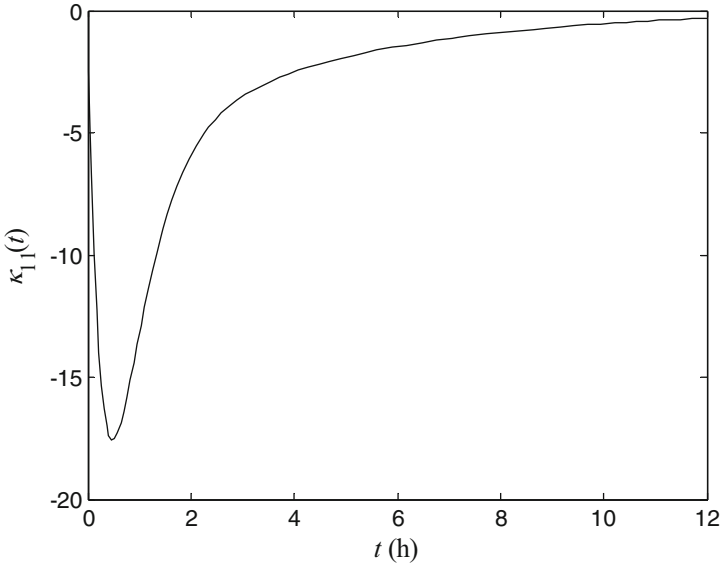


Fig. 11.29 Cumulant $\kappa_{11}(t)$ profile expressing the statistical dependence of the population sizes for the compartmental model

Enzymatic Reaction The intensity functions (11.46) were substituted directly into operator equation (11.39), which was transformed by (11.37) to yield

$$\frac{\partial \mathcal{K}}{\partial t} = k_{+1}\left[\exp\left(-\theta_1+\theta_2\right)-1\right]\left[e_0\frac{\partial \mathcal{K}}{\partial \theta_1}-\left(\frac{\partial^2 \mathcal{K}}{\partial \theta_1 \partial \theta_2}+\frac{\partial \mathcal{K}}{\partial \theta_1}\frac{\partial \mathcal{K}}{\partial \theta_2}\right)\right]$$
$$+\left\{k_{-1}\left[\exp\left(\theta_1-\theta_2\right)-1\right]+k_{+2}\left[\exp\left(-\theta_2\right)-1\right]\right\}\frac{\partial \mathcal{K}}{\partial \theta_2}.$$

Upon substituting the series expansion (11.48) into the previous equation, and equating the coefficients of θ_1 and θ_2, one has the following differential equations for the expected value functions:

$$\dot{\kappa}_{10} = -k_{+1}\left(e_0-\kappa_{01}\right)\kappa_{10}+k_{-1}\kappa_{01}+k_{+1}\kappa_{11},$$
$$\dot{\kappa}_{01} = k_{+1}\left(e_0-\kappa_{01}\right)\kappa_{10}-\left(k_{-1}+k_{+2}\right)\kappa_{01}-k_{+1}\kappa_{11}.$$

These equations are not equivalent to the deterministic formulation given by (8.8). The last term $k_{+1}\kappa_{11}$ involving the stochastic interaction in the previous equations expresses the main difference between deterministic and stochastic solutions for a nonlinear system.

The cumulant differential equations truncated at third order are

$$\dot{\kappa}_{20} = k_{+1}\left(A-2B\right)+k_{-1}\left(\kappa_{01}+2\kappa_{11}\right),$$
$$\dot{\kappa}_{02} = k_{+1}\left(A+2C\right)+k_0\left(\kappa_{01}-2\kappa_{02}\right),$$
$$\dot{\kappa}_{11} = k_{+1}\left(B-A-C\right)-k_{-1}\left(\kappa_{01}-\kappa_{02}\right)-k_0\kappa_{11},$$
$$\dot{\kappa}_{30} = k_{+1}\left(3B-A-3D\right)+k_{-1}\left(\kappa_{01}+3\kappa_{11}+3\kappa_{21}\right),$$
$$\dot{\kappa}_{21} = k_{+1}\left(A+C+D-2B-2F\right)+k_{-1}\left(\kappa_{02}-\kappa_{01}+2\kappa_{12}-2\kappa_{11}\right)-k_0\kappa_{21},$$
$$\dot{\kappa}_{12} = k_{+1}\left(B-A-2C-E+2F\right)+k_{-1}\left(\kappa_{03}+\kappa_{01}-2\kappa_{02}\right)+k_0\left(\kappa_{11}-2\kappa_{12}\right),$$
$$\dot{\kappa}_{03} = k_{+1}\left(A+3C+3E\right)+k_0\left(\kappa_{01}+3\kappa_{02}-3\kappa_{03}\right),$$

with

$$A = \left(e_0-\kappa_{01}\right)\kappa_{10}-\kappa_{11},$$
$$B = \left(e_0-\kappa_{01}\right)\kappa_{20}-\left(\kappa_{21}+\kappa_{10}\kappa_{11}\right),$$
$$C = \left(e_0-\kappa_{01}\right)\kappa_{11}-\left(\kappa_{12}+\kappa_{10}\kappa_{02}\right),$$
$$D = \left(e_0-\kappa_{01}\right)\kappa_{30}-\left(2\kappa_{20}\kappa_{11}+\kappa_{10}\kappa_{21}\right),$$
$$E = \left(e_0-\kappa_{01}\right)\kappa_{12}-\left(2\kappa_{02}\kappa_{11}+\kappa_{10}\kappa_{03}\right),$$
$$F = \left(e_0-\kappa_{01}\right)\kappa_{21}-\left(\kappa_{20}\kappa_{02}+\kappa_{11}^2+\kappa_{10}\kappa_{12}\right),$$
$$k_0 = k_{-1}+k_{+2}.$$

In these equations, the contributions of the fourth- and higher-order cumulants are neglected.

From the above, we remark again as for the first-order cumulant that the differential equations for the second-order cumulants κ_{20}, κ_{02}, and κ_{11} imply the third-order cumulants κ_{12} and κ_{21} and so on. This can be generalized by noting that the differential equation for the j-th cumulant function for a ψ-degree power in the intensity function model involves terms up to the $(j + \psi)$-th cumulant. Obviously, this fact rules out exact solutions, such as those previously found for the linear kinetic model, for the present equations. A standard approach to this problem has been to assume that the population size variable follows a Gaussian distribution, and set to 0 all cumulants of order 3 or higher. One can also intend [552] to find approximating cumulant functions using a "cumulant truncation" procedure. In this approach, one approximates the cumulant functions of any specific order, say j, of a ψ-degree power model by solving a system of up to the first $(j + \psi)$ cumulant functions with all higher-order cumulants set to 0.

Initial conditions for the cumulant differential equations are $\kappa_{ij}(0) = 0$ except for $\kappa_{10}(0) = s_0$. Setting to 0 all cumulants of order 4 or higher, simulations of these equations confirm the expected behaviors and the associated confidence intervals. Through the $\kappa_{11}(t)$ profile shown in Figure 11.30, this analysis reveals the statistical independence of the population sizes. Moreover, $\kappa_{11}(t)$ magnified by k_{+1} evaluates the discrepancy between the deterministic and the stochastic solution: the

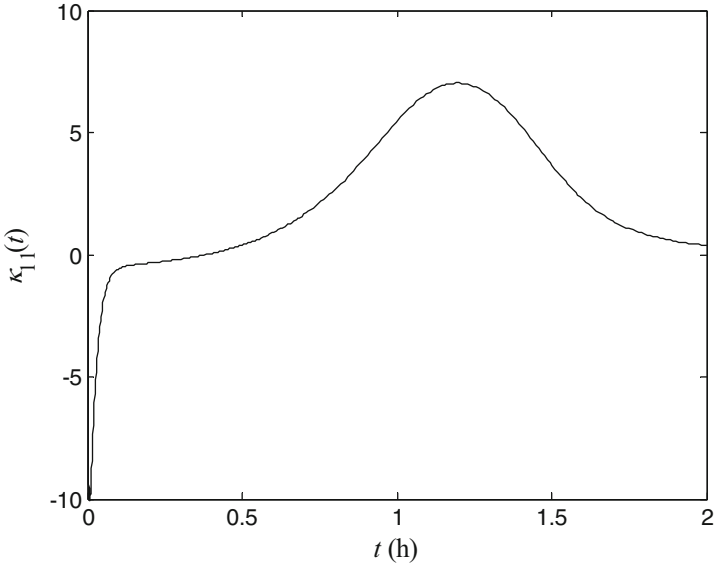


Fig. 11.30 Cumulant $\kappa_{11}(t)$ profile expressing the statistical dependence of the population sizes for the enzymatic model

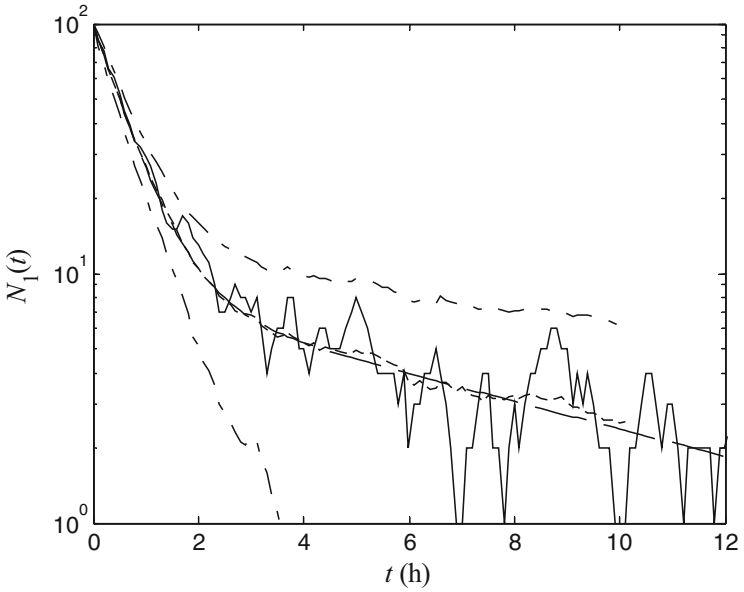


Fig. 11.31 The deterministic profile (*dashed line*), typical run (*solid line*), average (*dotted line*), and confidence corridor (*dashed-dotted line*) in compartment 1

substrate is overestimated at the early time of reaction by the deterministic model and underestimated over 0.5 h with a maximum about 1.2 h.

11.5.5.3 Stochastic Simulation Algorithm

As previously, initial conditions for the compartmental model and the enzymatic reaction were set to $\underline{n}_0^T = [100\ 50]$, and $s_0 = 100$, $e_0 = 50$, and $c_0 = 0$, respectively. Figures 11.31 and 11.32 show the deterministic prediction, a typical run, and the average and confidence corridor for 100 runs from the stochastic simulation algorithm for the compartmental system and the enzyme reaction, respectively. Figures 11.33 and 11.34 show the coefficient of variation for the number of particles in compartment 1 and for the substrate particles, respectively.

"On average," the solutions supplied by the deterministic system and the stochastic method are in close agreement, but the stochastic approach captures the fluctuations in the system. In comparing Figures 11.33 and 11.34, it is clear that when the number of molecules is large, the fluctuations might take the appearance of noise. But when there are small numbers of molecules, the fluctuations may in fact no longer be just noise but a significant part of the signal. Whether these fluctuations make a difference in the basic behavior of the system depends on the characteristics of that particular system. The system may also move between situations in which the fluctuations do and do not matter. However, when it is known that the system

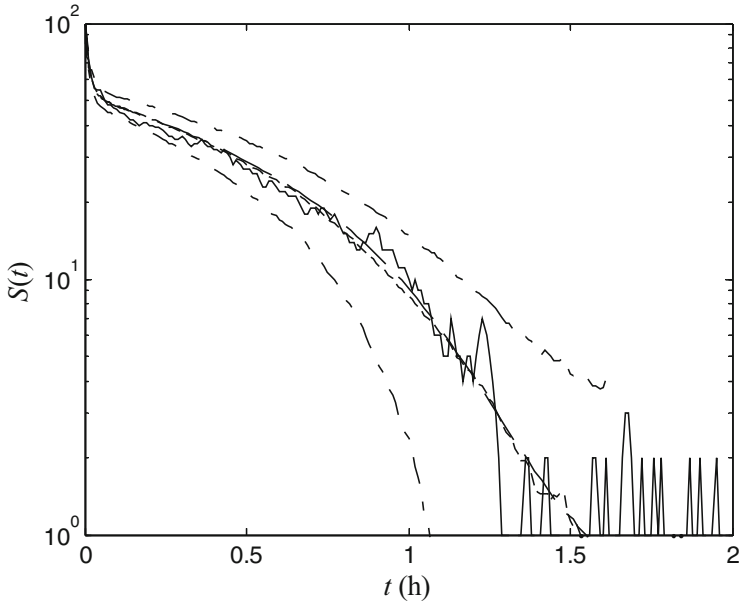


Fig. 11.32 The deterministic profile (*dashed line*), typical run (*solid line*), average (*dotted line*), and confidence corridor (*dashed-dotted line*) for substrate particles

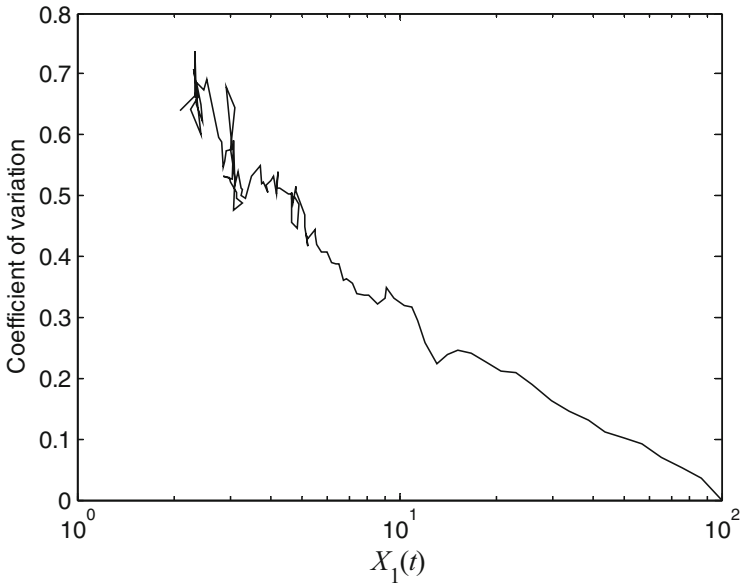


Fig. 11.33 Coefficient of variation for the particles in compartment 1

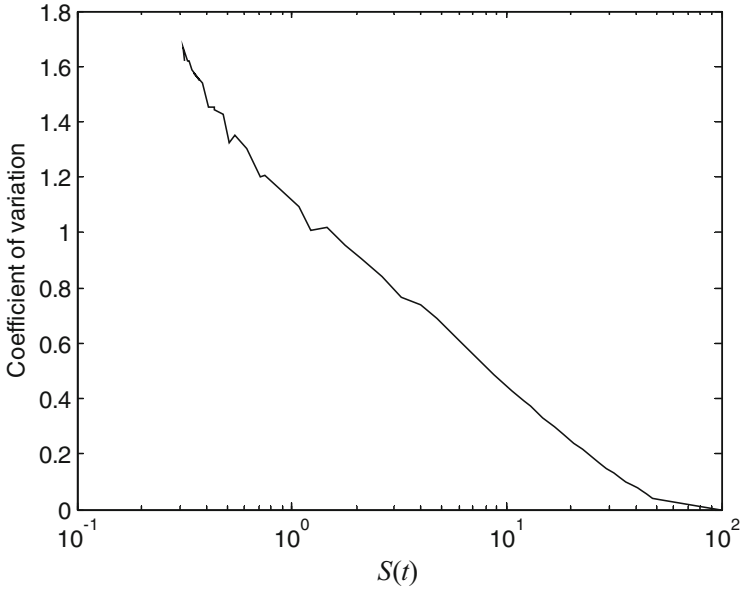


Fig. 11.34 Coefficient of variation for the substrate particles

contains small numbers of molecules and the network is nonlinear, the stochastic approach appears to be a more appropriate method, because both of these situations will magnify any fluctuations that already exist in the system.

11.6 Fractals and Stochastic Modeling

In the classical book [4], the distinct models dealing with ion channel kinetics are extensively discussed. One of the most important results is the connection established between fractal scaling and stochastic modeling. Based on experimental data, Liebovitch et al. [561] assessed the dependence of the effective kinetic constant $k_\circ$ on the sufficient time scale for detection $t_\circ$ by a fractal scaling relationship:

$$\log k_\circ (t_\circ) = \log \alpha + \left(1 - d_f\right) \log t_\circ, \tag{11.49}$$

where α is a constant and d_f is the fractal dimension. Moreover, the effective kinetic constant $k_\circ (t_\circ)$ can be considered as the conditional probability per unit time that the channel changes state (open vs. closed), i.e., $k_\circ (t_\circ)$ is considered as the hazard function $h(t_\circ)$ defined by (11.6). In that case, the survival function $\mathcal{S}(t_\circ)$ is the cumulative probability $\Pr[T_\circ > t_\circ]$ that the duration of the open (or closed) state $T_\circ$ is greater than $t_\circ$. Solving (11.6) and using the fractal scaling relationship (11.49), we obtain

$$\mathcal{S}(t_\circ) = \exp\left[-\frac{\alpha}{2-d_f}(t_\circ)^{2-d_f}\right],$$

which is the Weibull *survival* function already mentioned in Table 11.1. When the fractal dimension is close to 2, the previous equation takes the form of a power-law of time:

$$\mathcal{S}(t_\circ) = g\left(2-d_f\right)(t_\circ)^{-\alpha},$$

where $g\left(2-d_f\right)$ is a function of the fractal dimension. This form is equivalent to (2.8) in Chapter 2. From this development, we note the correspondence between the time scale sufficient for detection, $t_\circ$, and the age a of particles in a given compartment. This short presentation illustrates how fractality could be incorporated in retention-time distributions.

All the stochastic models presented here may include multiple compartments, age-varying rates, and heterogeneous particles with random rate coefficients, and their mathematical solutions tend to have various forms, e.g., exponential form, power function, damped oscillatory regimens, etc. This formulation concerns systems that are discrete in space, i.e., the particle can be located in one of a number of discrete compartments, and continuous in time; i.e., the particle is located continuously in one compartment until a transition occurs that discretely moves it to another compartment. In fact, stochastic models, and especially semi-Markov models, are tools for analyzing data when the response of interest is the time up to the occurrence of some event. Such events are generically referred to as *failures*, although the event may, for instance, be the ability of a power system to supply energy on demand without local failures or large-scale blackouts, the operating hours of replaced parts in equipment already in field use, industrial product testing, or the change of residence in a demographic study.

Certainly at the beginning, stochastic modeling had applications in the field of reliability, a relatively new field whose conception is primarily due to the complexity, sophistication, and automation inherent in modern technology. The problems of maintenance, repair, and field failures became severe for the military equipment used in World War II. In the late 1940s and early 1950s reliability engineering appeared on the scene [562–564]. For instance, the analysis of a process with operative and failure states can be based on the model presented in Figure 11.2C. Compartments 1 and 2 correspond to states in which the process is operable and failed, respectively. A_1 corresponds to the time before failure, and A_2 to the time needed to repair. Lastly, ω and $1-\omega$ correspond to the probabilities of entering reparable and irreparable failure states, respectively.

Recently, pharmacodynamicists have become interested in stochastic modeling for analyzing failure time data associated with pharmacological treatments [565]. Despite the unquestionable erudition of stochastic modeling, only a few of the stochastic models proposed to account for the observed biological data enjoy widespread use. The main reasons are that parameter estimation of stochastic

processes in biology is a relatively recent enterprise and that a number of models involve the application of fairly advanced statistics that typically lie beyond the scope and knowledge of experimental biologists.

11.7 Stochastic vs. Deterministic Models

In many cases and with an acceptable degree of accuracy, the time evolution of a dynamic system can be treated as a continuous, deterministic process. For deterministic processes the law of mass conservation is well grounded in experiments and also leads to equations that can be readily solved. Besides the great importance of the differential equation approach for either compartmental analysis or analysis of reactions involved in a living system, we should not lose sight of the fact that the physical basis for this method leaves something to be desired. The approach evidently assumes that the time evolution of a real process is both *continuous* and *deterministic*. However, time evolution of such a system is not a continuous process, because particle population sizes can obviously change only in discrete integer amounts. Moreover, time evolution is not a deterministic process. Even if we put aside quantum considerations and regard particle motions as governed by the equations of classical mechanics, it is impossible to predict the exact particle population size at some future time unless we take into account the precise positions and velocities of all the particles in the system.

This criticism has been supported by several recent experimental results that strongly suggest that several processes, like ecological systems, microscopic biological systems, and nonlinear systems driven to conditions of instability, in fact behave stochastically. So it was not until the early 1950s that it became clear that in small systems the law of mass conservation breaks down and that even small fluctuations in the number of molecules may be a significant factor in the behavior of the system [566]. Therefore, the equations obtained by using the law of mass conservation to describe fluctuations in the particle population sizes can be a serious shortcoming. Implicit in using the law of mass conservation are the key assumptions of continuity and determinism that could be warranted when there is a large number of the molecules of interest. These assumptions are reasonable for some systems of reactants, like a flask in the chemistry lab, but they are questionable when it comes to small living systems like cells and neurological synapses.

For instance, it turns out that inside a cell the situation is not continuous and deterministic, and that random fluctuations drive many of the reactions. With regard to the continuity assumption, it is important to note that individual genes are often present only in one or two copies per cell and that the regulatory molecules are typically produced in low quantities [567]. The low number of molecules may compromise the notion of continuity and consequently that of homogeneity. As for determinism, the rates of some of these reactions are so slow that many minutes may pass before, for instance, the start of mRNA transcription after the necessary molecules are present. This may call into question the notion of deterministic change

due to the fluctuations in the timing of cellular events. As a consequence, two regulatory systems having the same initial conditions might ultimately settle into different states, a phenomenon strengthened by the small numbers of molecules involved. This phenomenon is already reported as *sensitivity to initial conditions* (cf. Section 3.4) and it is characteristic of a nonlinear system exhibiting chaotic behavior. Thus, heterogeneity may be at the origin of fluctuations, and fluctuations are the prelude of instability and chaotic behavior.

Consequently, the observed process uncertainty may actually be an important part of the system and the expression of a structural heterogeneity. When the fluctuations in the system are small, it is possible to use the traditional deterministic approach. But when fluctuations are not negligibly small, the obtained differential equations will give results that are at best misleading, and possibly very wrong if the fluctuations can give rise to important effects. With these concerns in mind, it seems only natural to investigate an approach that incorporates small volumes and small numbers of particle populations and that may actually play an important part.

However, research along these lines is relatively scarce. The mathematical biology community continues to produce work that ignores the fact that there is a very different world inside a small biological system where topological heterogeneity prevails over homogeneity. So we turned to methods that are better able to capture the inherent stochastic nature of the system like the previously developed probabilistic transfer model, which expresses a structural heterogeneity and generates the process uncertainty corresponding to the observed fluctuations in the real process.

Aside from the continuity assumption and the discrete reality discussed above, deterministic models have been used to describe only those processes whose operation is fully understood. This implies a perfect understanding of all direct variables in the process and also, since every process is part of a larger universe, a complete comprehension of how all the other variables of the universe interact with the operation of the particular subprocess under study. Even if one were to find a real-world deterministic process, the number of interrelated variables and the number of unknown parameters are likely to be so large that a complete mathematical analysis would probably be so intractable that one might prefer to use a simpler stochastic representation. A small, simple stochastic model can often be substituted for a large, complex deterministic model since the need for the detailed causal mechanism of the latter is supplanted by the probabilistic variation of the former. In other words, one may deliberately introduce simplifications or "errors in the equations" to yield an analytically tractable stochastic model from which valid statistical inferences can be made, in principle, on the operation of the complex deterministic process.

For modeling purposes, the complexity pictured by heterogeneity undoubtedly requires more much knowledge than homogeneous conditions. If homogeneity prevails over heterogeneity, deterministic models may be good candidates to describe the real process. Conversely, the huge amount of knowledge needed to describe heterogeneity could be summarized only by the statistical concepts provided by stochastic modeling approaches.

Stochastic models have much to offer at the present time in strengthening the theoretical foundation and in extending the practical utility of the widespread deterministic models. After all, in a mathematical sense, the deterministic model is a special limiting case of a stochastic model.

The stochastic formulation was proposed to account for the heterogeneity in biological media since it supplies tractable forms to fit the data. These forms involve time-varying parameters in dynamic modeling. But it is unlikely to have parameters depending on time through a single maturation or age dependence. We believe that internal dynamic states of the process are involved in these time-dependencies. Introduction of these states leads to nonlinear dynamic modeling associated with various levels of stability. Naturally occurring, the nonlinear model may exhibit chaotic behavior. Thus, one must frequently expect chaotic-like behavior when the process is heterogeneous. In contrast, it is impossible to expect chaotic properties with homogeneous processes.

Part IV
Modeling in Pharmacodynamics

Pharmacodynamics has been described as "what the drug does to the body," in contrast to pharmacokinetics, which tells "what the body does to the drug." The simplest pharmacodynamic study is a phase I clinical trial, in which the purpose is to relate dose to toxicity. In a more general context, pharmacodynamics may concern efficacy as well as toxicity, or both. Such studies complement the more frequently performed pharmacokinetic studies, and pharmacokinetic-dynamic studies are becoming increasingly popular as a means of encompassing complete organism responses to pharmacological interventions. They can synthesize and integrate knowledge gained from multiple experiments, in vitro and in vivo, and can be used for forecasting or confirming (validation) the results of drug administration or of clinical trials.

In comparison to pharmacokinetic modeling, the task of defining appropriate structural and statistical models to represent and forecast pharmacodynamic outcomes is more complex. This complexity first relies on the difficulty of proper selection of the outcome measures in order to relate concentration to quantifiable effect. However, investigators tend to focus on those efficacy measures that are most reproducible or convenient for quantification, but these may bear little or no relationship to the therapeutically relevant effect. Hence, the importance of the appropriate choice of short-term markers predictive of long-term clinical outcomes has repeatedly been emphasized [568].

Efficacy measures can be categorized as *biomarkers*, *surrogate markers*, or *clinical outcomes*. While the clinical outcome is the ultimate efficacy measure quantifying the direct benefit to a patient, it is often difficult to measure. Instead, clinical outcomes are often predicted from surrogate markers that can readily and sooner be observed and can easily be quantified. The essential feature of a surrogate marker is that it predicts clinical outcome. Before surrogate endpoints can be used to predict clinical outcomes, however, it is crucial to validate them and prove their relevance. Biomarkers are measurable physiological or biochemical parameters that reflect some pharmacodynamic activity of the investigated drug, even if they are not directly related to clinical outcome. They may be useful to get insight into the

overall pharmacodynamic behavior of a compound of interest. Surrogate markers are clinically validated biomarkers.

The complexity in pharmacodynamic studies also arises from the wide variety of pharmacological responses that are often discontinuous. These responses are classified as dichotomous, categorical, time-to-event, or counts:

- When the observed response is *dichotomous*, the probability of response is related to the level of the pharmacodynamic stimulus.
- *Categorical* responses can be regarded as generalized dichotomous responses: several distinct responses are possible, not just one [569].
- *Time-to-event* data record the time elapsing from some natural origin of the time scale to an event [565].
- The *count model* describes data that report the number of events in a volume of space or time.

Despite the problems associated with constructing appropriate efficacy measures, a plethora of classical pharmacodynamic models have been developed. Accordingly, in this last part of the book, we first present in Chapter 12 the classical modeling approaches used to analyze the observed pharmacodynamic effects. In Chapter 13, we present concepts and applications of nonlinear dynamics to pharmacodynamics, placing particular emphasis on chemotherapy and drugs affecting the central nervous system, the cardiovascular system, and the endocrine function.

Chapter 12
Classical Pharmacodynamics

The master of the oracle at Delphi does not say anything and does not conceal anything, only hints.

Heraclitus of Ephesus (544–483 BC)

Receptors are the most important targets for therapeutic drugs [570]. Therefore, it is important to explore the mechanisms of receptor modulation and drug action in intact in vivo systems. Also, the need for a more mechanism-based approach in pharmacokinetic-dynamic modeling has been increasingly recognized [571, 572]. Hill [573] made the first explicit mathematical model of simulated drug action to account for the time courses and concentration–effect curves obtained when nicotine was used to provoke contraction of the frog rectus abdominis muscle.

Simple mathematical calculations by the first pharmacologists in the 1930s indicated that structurally specific drugs exert their action in very small doses and do not act on all molecules of the body but only on certain ones, those that constitute the drug receptors. For example, Clark [574] calculated that ouabain applied to the cells of the heart ventricle, isolated from the toad, would cover only 2.5% of the cellular surface. These observations prompted Clark [574, 575] to apply the mathematical approaches used in enzyme kinetics to the effects of chemicals on tissues, and this formed the basis of the occupancy theory for drug–receptor interaction. Thus, pharmacological receptor models preceded accurate knowledge of receptors by many years.

12.1 Occupancy Theory in Pharmacology

According to the occupancy theory, which has evolved chronologically from the original work of Clark [574, 575], the drug effect is a function of two processes:

- binding of drug to the receptor and drug-induced activation of the receptor, and
- propagation of this initial receptor activation into the observed pharmacological effect, where the intensity of the pharmacological effect is proportional to the number of receptor sites occupied by drug.

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_12

Therefore, the drug–receptor interaction follows the law of mass action and may be represented by the equation

$$\begin{gathered}\gamma\,[\text{drug molecules}]+[\text{receptor}]\underset{k_{-1}}{\overset{k_{+1}}{\rightleftarrows}}[\text{drug–receptor complex}]\\ [\text{drug–receptor complex}]\overset{k_2}{\Longrightarrow}[\text{pharmacological effect}]\,,\end{gathered}\tag{12.1}$$

where γ molecules of drug activate a receptor and give an activated receptor usually called the drug–receptor complex. Although γ is defined as the number of molecules interacting with one receptor, it is in practice merely used to provide better data fits. Rate constants k_{+1}, k_{-1} characterize the association and dissociation of the complex, respectively. The ratio k_{-1}/k_{+1} is defined in pharmacology as the dissociation constant k_D of the complex. The proportionality constant k_2 relates the drug–receptor complex concentration $\upsilon\,(t)$ with the pharmacological effect $E\,(t)$, through the equation

$$E\,(t)=k_2\upsilon\,(t)\,.\tag{12.2}$$

When the total number of receptors r_0 is occupied, the effect will be maximal:

$$E_{\max}=k_2 r_0.\tag{12.3}$$

For drug concentration $c\,(t)$ and a total receptor concentration r_0 we thus have

$$\dot{\upsilon}\,(t)=k_{+1}c^{\gamma}\,(t)\,[r_0-\upsilon\,(t)]-k_{-1}\upsilon\,(t)\,,\;\;\upsilon\,(0)=0.\tag{12.4}$$

In the equilibrium state ($\dot{\upsilon}\,(t)=0$ assumption **H1**) we have

$$\upsilon^*=\frac{r_0 c^{*\gamma}}{k_D+c^{*\gamma}},\tag{12.5}$$

where c^*, υ^* are the drug and drug–receptor complex concentrations in the equilibrium, respectively. By combining the last equation with (12.2) and (12.3), we obtain the working equation for the so-called *sigmoid* $E_{\max}$ *model*:

$$E^*=\frac{E_{\max}c^{*\gamma}}{k_D+c^{*\gamma}},\tag{12.6}$$

where E^* is the pharmacological effect at equilibrium. From the last equation, it can be seen that the dissociation constant k_D expresses also the γ-power of drug concentration needed to induce half maximal effect ($E_{\max}/2$). When γ is set to 1, the model is called the *basic* $E_{\max}$ *model*, but this model offers less flexibility in the shape of the function compared to the sigmoid $E_{\max}$ model.

Assuming relatively rapid drug–receptor equilibrium with respect to $c\,(t)$ variations, then $c^*\approx c\,(t)$ (assumption **H2**), so the previous equation becomes

$$E^*(t) = \frac{E_{\max} c^{\gamma}(t)}{k_D + c^{\gamma}(t)}, \tag{12.7}$$

where $E^*(t)$ indicates that the effect is driven by the pharmacokinetic time.

With $\gamma = 1$, (12.6) has been used extensively in pharmacology to describe the effect of chemicals on tissues in the modified form:

$$E^* = \frac{\varepsilon r_0 c^*}{k_D + c^*},$$

where ε is the intrinsic efficacy (inherent ability of the chemical to induce a physiological response). In other words, ε is the proportionality constant k_2 relating the receptor density r_0 with the maximal effect $E_{\max}$ (12.3). In order to avoid the use of the efficacy term (due to its ad hoc nature), Black and Leff [576] introduced in 1983 the operational model of drug action

$$E^* = \frac{\rho E_{\max} c^*}{k_D + (\rho + 1) c^*},$$

where ρ is equal to the ratio of the receptor density over the concentration of the complex that produces 50% of the maximal tissue response. In reality, this constant ratio characterizes the propensity of a given chemical–tissue system to yield a response.

Since the development of the occupancy theory, the mathematical models used to explain the action of ligands at receptors have been subject to continuous development prompted by new experimental observations. Currently, pharmacological studies deal with drug–receptor or drug–tissue interactions to get estimates for receptor (tissue) affinity and capacity. Thus, the operational model enjoys widespread application in the field of functional receptor pharmacology [577]. Although this model is routinely applied to in vitro studies, the estimates for receptor affinity and capacity can be used for prediction of the effect in vivo. In principle, k_D should be of the same order as the unbound Ec_{50}^{γ}, where Ec_{50} is the concentration at half maximal effect in vivo. In this context, Visser et al. [578] correlated the in vitro measurements with in vivo observations in rats when studying the effect of γ-aminobutyric acid receptor modulators on the electroencephalogram.

12.2 Empirical Pharmacodynamic Models

Combined pharmacokinetic-dynamic studies seek to characterize the time course of drug effects through the application of mathematical modeling to dose–effect–time data. This definition places particular emphasis on the time course of drug action. Pharmacodynamics is intrinsically related to pharmacokinetics, which encompasses the study of movement of drugs into, through, and out of the body. The term

pharmacodynamic models exclusively refers to those models that relate drug concentration with the pharmacological effect.

The most common function used to relate drug concentration c with effect is the $E_{\max}$ model:

$$E = \frac{E_{\max} c^{\gamma}}{Ec_{50}^{\gamma} + c^{\gamma}}, \tag{12.8}$$

where $E_{\max}$ is the maximum effect and Ec_{50} is the concentration at half the maximal observable in vivo effect. Equation (12.8) corresponds to (12.6) with Ec_{50}^{γ} substituting k_D. It is also clear that (12.8) is a static nonlinear model in which c corresponds to the equilibrium point c^*. If we consider c as a time course $c\,(t)$, we must implicitly assume that equilibrium is achieved rapidly throughout $c\,(t)$, so $c^* \equiv c\,(t)$ (assumption **H2**).

If a baseline E_0 is introduced to the previous equation,

$$E = E_0 \pm \frac{E_{\max} c^{\gamma}}{Ec_{50}^{\gamma} + c^{\gamma}},$$

we obtain the $E_{\max}$ model describing either stimulation or inhibition of the effect by the concentration of the drug. Parameters $E_{\max}$, Ec_{50}, and γ are assumed constant and independent of the drug dose as well as the drug and receptor concentrations.

Other simpler empirical models have also been used since the early days of pharmacodynamics [579, 580] to describe the drug concentration–effect relationship. The linear model relies on a linear relationship between E and c:

$$E = \alpha c + \beta, \tag{12.9}$$

where α is the slope indicating the sensitivity of the effect to concentration changes. The intercept β can be viewed as the baseline effect. Equation (12.9) reveals that the linearity between c and E is unlimited, and this feature is undoubtedly a drawback of the model. Besides, a log-linear model between E and c can also be considered:

$$E = \alpha \log (c) + \beta. \tag{12.10}$$

Due to the logarithmic expression of concentration in this model a larger concentration range is related "linearly" with the effect. As a rule of thumb, 20 to 80% of the concentration range of the $E_{\max}$ model can be approximately described with (12.10).

Although these empirical approaches may quantify and fit the data well, they do not offer a physical interpretation of the results.

12.3 Pharmacokinetic-Dynamic Modeling

In the mid-1960s, G. Levy [579, 580] was the first to relate the pharmacokinetic characteristics with the in vivo pharmacological response of drug using the above-mentioned linear models. In fact, as the pharmacological responses $E(t)$ and the drug concentration $c(t)$ can be observed simultaneously and repeatedly as a function of time, a combined pharmacokinetic-dynamic model is needed to describe these time courses. From the simple models, the discipline of pharmacokinetic-dynamic modeling emerged gradually, and in actuality even complex physiological processes controlling drug response can be modeled. The key mechanisms intrinsic to pharmacokinetic-dynamic models are the following:

- the processes may take place under either equilibrium or nonequilibrium conditions for the pharmacodynamic part,
- the binding of drug with the receptor may either be reversible or irreversible, and
- the bound drug may induce its effect directly or indirectly.

A general scheme for the basic components of pharmacokinetic-dynamic models is depicted in Figure 12.1. According to this scheme, the drug at the prereceptor phase is considered to distribute to an effect compartment; then it reacts with the receptors under equilibrium (direct link, assumption **H3**) or nonequilibrium (indirect-link) conditions, and finally, at the postreceptor phase, the activated receptors can either produce the response directly (direct response, assumption **H4**) through the transducer function $\mathcal{T}$ (which is usually a proportionality constant like k_2 in equation 12.1) or they can interfere with an endogenous or already existing process that produces the final response (indirect response). In fact, all the processes

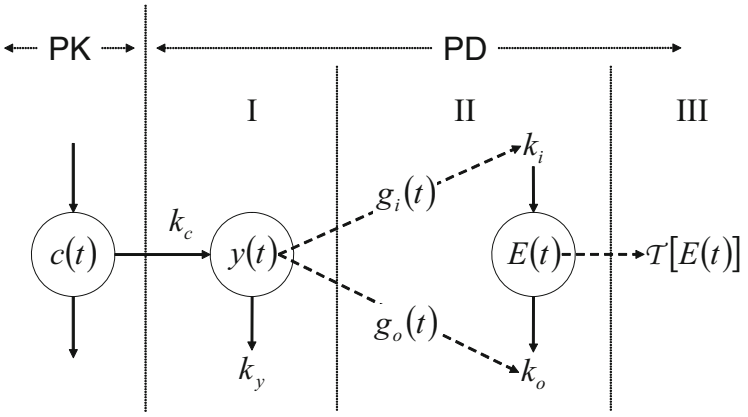


Fig. 12.1 Schematic of the basic processes involved in pharmacokinetic (**PK**)-dynamic (**PD**) models. The phases I, II, and III refer to processes that take place in the prereceptor, receptor, and postreceptor proximity, respectively. The symbols are defined in the text

of the general model depicted in Figure 12.1 are not necessarily incorporated in the final model used in practice. Almost always, one of these steps is considered to be the limiting one, and the model reduces to one of the basic models described below.

12.3.1 Link Models

During the first decades of the development of pharmacokinetic science, a lag time in pharmacological response after intravenous administration was often treated by applying a compartmental approach. If the plasma concentration declined in a biexponential manner, the observed pharmacodynamic effect was fitted to plasma or "tissue" compartment concentrations. Due to the lag time of effects, a successful fit was sometimes obtained between effect and tissue drug level [581]. However, there is no a priori reason to assume that the time course of a drug concentration at the effect site must be related to kinetics in tissues that mainly cause the multiexponential behavior of the plasma time–concentration course. A lag time between drug levels and dynamic effects can also occur for drugs described by a one-compartment model.

Segre [582] was the first author to consider the possibility that the time course of pharmacological effect could itself be used to describe the transfer rate of a drug to the biophase. Thus, the lag time of the effect was modeled by including two hypothetical tissue compartments between the plasma compartment and the pharmacodynamic response compartment.

The idea of Segre was further developed, in an elegant way, by Sheiner and associates [583, 584] by linking the effect compartment to a kinetic model. This approach has since been called the *link model*. The time course of the drug in the effect site is determined by the rates of transfer of material into and from the effect compartment; the lag time of the effect-site concentration is controlled by the elimination rate constant of the effect compartment. The beauty of this approach is that instead of relating the pharmacodynamic response to drug concentrations in some more or less well-defined tissue, it is related to the plasma drug level, which in clinical practice is of great importance.

12.3.1.1 Direct Link

Strictly speaking, pharmacodynamic models are employed to relate the receptor site drug concentration to pharmacological response at any given time using data mainly from in vivo experiments. However, the receptor site drug concentration normally cannot be measured directly. Thus, the simplest pharmacokinetic-dynamic mechanistic model arises from assuming that the drug concentration in the blood, $c\,(t)$ (far left compartment of Figure 12.1), is the same at the receptor site, $y\,(t)$. Strictly speaking, this assumption expresses a prereceptor equilibrium (**H3**) and the resulting model does not utilize concentrations at the effect site.

Further, under equilibrium conditions **H1**, we can use (12.6) to relate the pharmacological effect E^* with the drug concentrations c^*, or in addition, use (12.7) to relate the time courses $E^*(t)$ and $c^*(t)$ under the supplementary assumption **H2**. Thus, the simplest mechanistic models are once again the basic and the sigmoid $E_{\max}$ models, but now they have a specific physical interpretation in terms of drug–receptor reaction kinetics.

As is implicit from all the above, the measured concentration in plasma is directly linked to the observed effect for these simple mechanistic, pharmacokinetic-dynamic models. Accordingly, these models are called *direct-link models* since the concentrations in plasma can be used directly in (12.6) and (12.7) for the description of the observed effects. Under the assumptions of the direct-link model, plasma concentration and effect maxima will occur at the same time, that is, no temporal dissociation between the time courses of concentration and effect is observed. An example of this can be seen in the direct-link sigmoid $E_{\max}$ model of Racine-Poon et al. [585], which relates the serum concentration of the anti-immunglobulin E antibody CGP 51901, used in patients for the treatment of seasonal allergic rhinitis, with the reduction of free anti-immunglobulin E.

Under the assumptions of the direct-link model, neither a counterclockwise (Figure 12.2) nor a clockwise hysteresis loop (Figure 12.4) will be recorded in an effect vs. concentration plot. In principle, the shape of the effect vs. concentration plot for an ideal direct-link model will be a curve identical to the specific

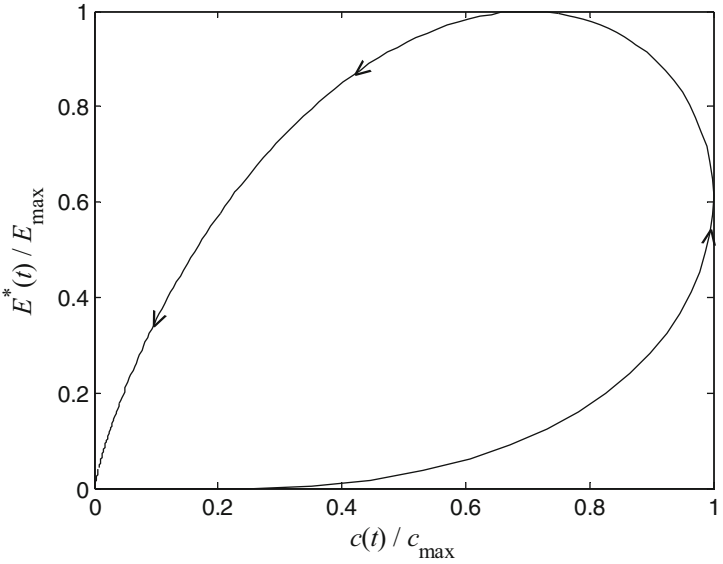


Fig. 12.2 Normalized effect–plasma drug concentration state space for the indirect-link model. As time flows (indicated by arrows) a counterclockwise hysteresis loop is formed. The rate constant for drug removal from the effect compartment k_y characterizes the temporal delay, that is, the degree of hysteresis

pharmacodynamic model, relating effect with concentration, e.g., linear for a linear pharmacodynamic model, sigmoid for the sigmoid $E_{\max}$ model etc.

12.3.1.2 Indirect Link: The Effect-Compartment Model

In the direct-link model, concentration–effect relationships are established without accounting for intrinsic pharmacodynamic temporal behavior, and relationships are valid only under the assumption of effect site, prereceptor equilibrium **H3**. In contrast, indirect-link models are required if there is a temporal dissociation between the time courses of concentration and effect, and the observed delay in the concentration–effect relationship is most likely caused by a functional delay between the concentrations in the plasma and at the effect site.

When a lag time of $E(t)$ is observed with respect to the $c(t)$ time course, the use of a combined pharmacokinetic-dynamic model, the indirect-link model, is needed to relate the drug concentration $c(t)$ to the receptor site drug concentration $y(t)$ (which cannot be measured directly) and the $y(t)$ to the pharmacological response $E(t)$.[1]

The effect-compartment model relaxes the assumption **H3** and it stems from the assumption of prereceptor nonequilibrium between drug concentration in the blood or plasma $c(t)$ and the receptor site $y(t)$. According to this model, an additional compartment is considered, the effect (or biophase) compartment, and concentration $y(t)$ in that compartment reacts with the receptors, Figure 12.1.

Notation:

- V_c and V_y denote the apparent volumes of distribution of the plasma and effect compartments, respectively.
- k_c and k_y denote the first-order rate constants for the drug transfer from plasma to effect site and for drug elimination from the effect site, respectively.

Then assuming that the mass-flux equality holds for the effect compartment, i.e., $V_c k_c = V_y k_y$, the drug concentration $y(t)$ in the effect compartment can be described by the linear differential equation

$$\dot{y}(t) = k_y [c(t) - y(t)], \quad y(0) = 0. \tag{12.11}$$

This equation can be solved by applying the Laplace transformation and convolution principles (cf. Appendix E):

$$y(t) = k_y \tilde{y}(t), \tag{12.12}$$

[1] In the classical pharmacokinetic-pharmacodynamic literature, the effect-site concentration and the effect-site elimination rate constant are denoted by c_E and k_{E0}, respectively. Here, the symbols $y(t)$ and k_y are used instead.

where $\tilde{y}(t)$ is defined as the *apparent effect site* drug concentration and it is given by

$$\tilde{y}(t) = \int_0^t c(t') \exp\left[-k_y (t - t')\right] \mathrm{d}t'.$$

The time symbols t', t denote the temporal dissociation between the time courses of concentration and effect, respectively. For various types of drug administration, the function $c(t)$ is known and therefore analytic solutions for $\tilde{y}(t)$ have been obtained using the integral defined above. Substituting (12.12) into (12.7), we obtain the fundamental equation for the $E_{\max}$ indirect-link model:

$$E^*(t) = \frac{E_{\max}^{\gamma} \tilde{y}^{\gamma}(t)}{\tilde{y}_{50}^{\gamma} + \tilde{y}^{\gamma}(t)}, \tag{12.13}$$

where $\tilde{y}_{50}$ is the apparent effect site drug concentration producing 50% of the maximum effect.

In this model, the rate constant k_y was originally considered to reflect a distributional delay of drug from plasma to the effect compartment. However, it can also be regarded as a constant producing the delay in effects in relation to plasma, irrespective of whether this is caused by distributional factors, receptor events, production of a mediator of any kind, etc.

The basic feature of the indirect-link model is the counterclockwise hysteresis loop that is obtained from plotting the observed values of the effect vs. the observed plasma drug concentration values, Figure 12.2. In other words, the effect is delayed compared to the plasma drug concentration and this is reflected in the effect–concentration state space.

Numerous applications of pharmacokinetic-dynamic models incorporating a biophase (or effect) compartment for a variety of drugs that belong to miscellaneous pharmacological classes, e.g., anesthetic agents [586], opioid analgesics [587–589], barbiturates [590, 591], benzodiazepines [592], antiarrhythmics [593], have been published. The reader can refer to a handbook [594] or recent reviews [572] for a complete list of the applications of the biophase distribution model.

In actual practice, nonlinear regression is used to fit a suitable pharmacokinetic model described by the function $c(t)$ to time–concentration data. Then, the estimated parameters are used as constants in the pharmacodynamic model to estimate the pharmacodynamic parameters. Alternatively, simultaneous fitting of the model to the concentration–effect–time data can be performed. This is recommended as $c(t)$ and $E(t)$ time courses are simultaneously observed.

Example 10. Bolus Intravenous Injection

An example of the indirect-link model after bolus intravenous injection can be seen in Figure 12.3. The arrow indicates time flow. Each point represents a uniquely defined state and only one trajectory may pass from it. The state space has a point attractor, i.e., a steady state, which is obviously the point ($c = 0$, $E^* = 0$) reached at theoretically infinite time. Three different initial conditions of the form $c(0) =$

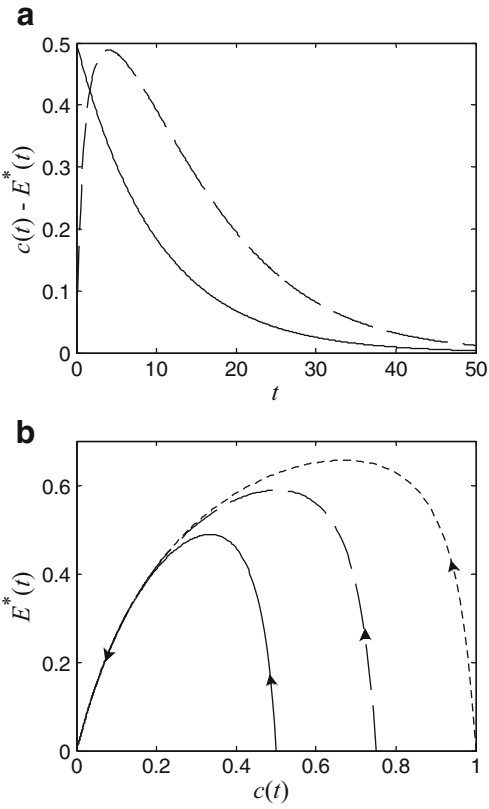


Fig. 12.3 Indirect-link model with bolus intravenous injection. (**a**) The classical time profiles of the two variables $c\,(t)$ (*solid line*) and $E^*\,(t)$ (*dashed line*) for dose $q_0 = 0.5$. (**b**) A two-dimensional phase space for the concentration $c\,(t)$ vs. effect $E^*\,(t)$ plot using three doses 0.5, 0.75, and 1 (*solid*, *dashed*, and *dotted lines*, respectively)

q_0/V_c, $E^*\,(0) = 0$, are used to generate three different trajectories, all of which end up at the point attractor. The integrated equations of the system are

$$c\,(t) = \frac{q_0}{V_c}\exp\left(-kt\right),$$

$$\tilde{y}\,(t) = \frac{q_0}{V_c}\frac{\left[\exp\left(-kt\right) - \exp\left(-k_y t\right)\right]}{k_y - k},$$

$$E^*\,(t) = \frac{E_{\max}\tilde{y}\,(t)}{\tilde{y}_{50} + \tilde{y}\,(t)},$$

where q_0 is the dose, V_c and k are the volume of distribution and the elimination rate constant for pharmacokinetics, k_y is the effect-site elimination rate constant, $E_{\max}$ is the maximum effect, and $\tilde{y}_{50}$ is the concentration at which 50% of the maximum effect is observed. Parameter values were set at

$$V_c = 1, \quad k = 0.1,$$
$$k_y = 0.5, E_{\max} = 1, \tilde{y}_{50} = 0.7.$$

where all units are arbitrary. ■

12.3.2 Response Models

Time is not an independent variable in the presented models. Dynamic behavior is either a consequence of the pharmacokinetics or the observed lag time by means of the effect compartment. Dynamic models from the occupancy theory and described by differential equations, such as (12.4), are scarce [595, 596].

Neglecting dynamic models in pharmacodynamics [597] is perhaps due to the fact in that instant equilibrium relationships between concentration and effect appear to occur for most drugs. For some drugs, such as cytotoxic agents, this delay is often extremely long, and attempts to model it are seldom made. One can describe these relationships as time-dissociated or nondynamic because the temporal aspects of the effect are not linked to the time–concentration profile.

In recent years, new models overcoming these defaults have been developed as the *indirect physiological models* introduced by Jusko and associates [598]. According to this last type of model, an endogenous substance or a receptor protein is formed at a constant rate and lost with a first-order rate constant. The drug concentration in plasma produces an effect by either stimulating or inhibiting the synthesis or removal of the endogenous substance leading to a change in the observed pharmacodynamic effect described by a suitable pharmacodynamic model.

12.3.2.1 Direct Response

The standard effect-compartment model, usually characterized as an atypical indirect-link model, also constitutes an example of what we will call a *direct-response* model in contrast to the *indirect-response* models. Globally, the standard direct-response models are models in which $c\,(t)$ affects all dynamic processes only linearly.

12.3.2.2 Indirect Response

Ariens [599] was the first to describe drug action through indirect mechanisms. Later on, Nagashima et al. [600] introduced the indirect-response concept to pharmacokinetic-dynamic modeling with their work on the kinetics of the anticoagulant effect of warfarin, which is controlled by the change in the prothrombin

complex synthesis rate. Today, indirect-response modeling finds extensive applications especially when endogenous substances are involved in the expression of the observed response.

From a modeling point of view, the last equilibrium assumption that can be relaxed, for the processes depicted in Figure 12.1, is **H4**, between the activated receptors (υ variable in the occupancy model) and the response E. Instead of the activated receptors directly producing the response, they interfere with some other process, which in turn produces the response E. This mechanism is usually described mathematically with a transducer function $\mathcal{T}$ which is no longer linear (cf. Section 12.4.1). This type of pharmacodynamic model is called *indirect response* and includes modeling of the response process usually through a linear differential equation of the form

$$\dot{E}(t) = k_i g_i(t) - k_o g_o(t) E(t), \quad E(0) = k_i/k_o, \tag{12.14}$$

where k_o is a first-order rate constant, and k_i represents an apparent zero-order production rate of the response. Stationarity conditions set the initial response value $E(0)$ at the ratio k_i/k_o. Functions $g_i(t)$ and $g_o(t)$ depend on the drug concentration through $E_{\max}$ functions and can produce either stimulation or inhibition, respectively:

$$g(t) = 1 + \frac{S_{\max} c(t)}{Sc_{50} + c(t)} \quad \text{or} \quad g(t) = 1 - \frac{I_{\max} c(t)}{Ic_{50} + c(t)}. \tag{12.15}$$

In these expressions, $g(t)$ is either $g_i(t)$ or $g_o(t)$, $S_{\max}$ is maximum stimulation rate, $I_{\max}$ is maximum inhibition rate with $I_{\max} < 1$, Sc_{50} and Ic_{50} are the drug concentrations at which $g(t) = 1 + (S_{\max}/2)$ and $g(t) = 1 - (I_{\max}/2)$, respectively. Consequently, four basic models are formulated: inhibition of k_i, inhibition of k_o, stimulation of k_i, and stimulation of k_o, Figure 12.1.

This family of the four basic indirect-response models has been proven to characterize diverse types of pharmacodynamic effects and it constitutes the current approach for pharmacokinetic-dynamic modeling of responses generated by indirect mechanisms. Thus, indirect-response models have been used to interpret the anticoagulant effect of warfarin, adrenal suppression by corticosteroids, cell trafficking effects of corticosteroids, the antipyretic effect of ibuprofen, aldose reductase inhibition, etc. [601]. Basically, the indirect-response concept is appropriate for modeling the pharmacodynamics of drugs that act through inhibition or stimulation of the production or loss of endogenous substances or mediators.

However, the general model described above is considered to be mechanistic. It is opposed to the completely empirical approach, since it is based on a general physiological process like receptor activation. But it is too general and abstract to describe complicated drug processes. Stimulation and inhibition of a baseline through the saturable $E_{\max}$ function is often not enough, since drugs interplay with complicated physiological processes. Thus, during the last ten years Jusko's

group and other investigators have expanded the application of indirect-response mechanisms to real mechanistic pharmacodynamic modeling and have included detailed modeling of the underlying physiology and then modeled the effect of drugs on it. These models are called *extended indirect-response models* [602] and they have been used to describe tolerance and rebound phenomena [603], time-dependency of the initial response [604, 605], cell trafficking dynamics [606], etc.

It is rather obvious that an indirect-response mechanism, whatever the detailed processes involved, results in a counterclockwise hysteresis loop for the effect–concentration relationship, Figure 12.2. Here, however, the elaboration of the observed response is usually secondary to a previous time-consuming synthesis or degradation of an endogenous substance(s) or mediator(s). Since both the indirect-link and indirect-response models have counterclockwise hysteresis effect–concentration plots, an approach based on the time of the maximum effect has been applied to furosemide data [607] for indirect (link or response) model selection.

When one looks into the basic functions of the link and indirect-response models, it is clear that one of the differences resides in the input functions to the effect and the receptor protein site, respectively. For the link model a linear input operates in contrast to the indirect model, where a nonlinear function operates. For the link model the time is not directly present and the pharmacological time course is exclusively dictated by the pharmacokinetic time, whereas the indirect model has its own time expressed by the differential equation describing the dynamics of the integrated response.

12.4 Other Pharmacodynamic Models

A number of other pharmacodynamic approaches focusing either on prereceptor or postreceptor events have been proposed in the literature and are discussed below.

12.4.1 The Receptor–Transducer Model

First, mention can be made of cases in which the measured effect, instead of being proportional to the activated receptors, follows a more general function $E = \mathcal{T}(\upsilon)$. This model is called *receptor–transducer* and was introduced by Black and Leff [576]. The function $\mathcal{T}$ is called a *transducer function* and its most common form is yet again the $E_{\max}$ function, which when replaced in (12.5) results in an $E_{\max}$ model but with different shape parameters called an *operational model* [608].

12.4.2 Irreversible Models

All the above-mentioned pharmacokinetic-dynamic models are characterized by reversibility of the drug–receptor interaction. In several cases, however, drug action relies on an irreversible bimolecular interaction; thus, enzyme inhibitors and chemotherapeutic agents exert their action through irreversible bimolecular interactions with enzymes and cells (bacteria, parasites, viruses), respectively.

The irreversible inactivation of endogenous enzymes caused by drugs, e.g., the antiplatelet effect of aspirin after oral administration [609], the 5α-reductase inhibition by a new nonsteroidal inhibitor [610], and the H^+, K^+-ATPase inactivation by proton pump inhibitors [611], is modeled with turnover models. The simplest model [609] includes terms for the production rate k_i and loss rate k_o of the response E, coupled with a function $g(c)$ representing the change of plasma or effect-compartment drug concentration:

$$\dot{E}(t) = k_i - [k_o + g(c)]\, E(t),$$

where k_i and k_o have the same meaning as defined for (12.14) while the function $g(c)$ is either linear or of Michaelian type.

The models used for the irreversible effects of chemotherapeutic agents quantify the response $E(t)$ in terms of the cell number since irreversible inactivation leads to cell killing. In these models, the function of the natural proliferation of cells $r(E)$ is combined with the cell-killing function $g(c)$, which again represents the change of plasma or effect-compartment drug concentration:

$$\dot{E}(t) = r(E) - g(c)\, E(t).$$

The function $r(E)$ can take various forms describing the natural growth of the cell population in the absence of drug [539, 612], while $g(c)$ can be either linear or nonlinear [602, 613, 614]. Due to the competitive character of the functions $r(E)$ and $g(c)$, the cell number vs. time plots are usually biphasic with the minimum effective concentration of drug being the major determinant for the killing or regrowth phases of the plot.

12.4.3 Time-Variant Models

Contrary to the already mentioned models, which include constant parameters, pharmacodynamic models may include time-varying parameters as well. Typical examples include models of drug tolerance or sensitization, where the parameters vary as a function of the dosing history. Other examples concern modeling of circadian rhythms where parameters depend explicitly on time through biological clocks, e.g., the baseline of a pharmacological response, and it is necessary to

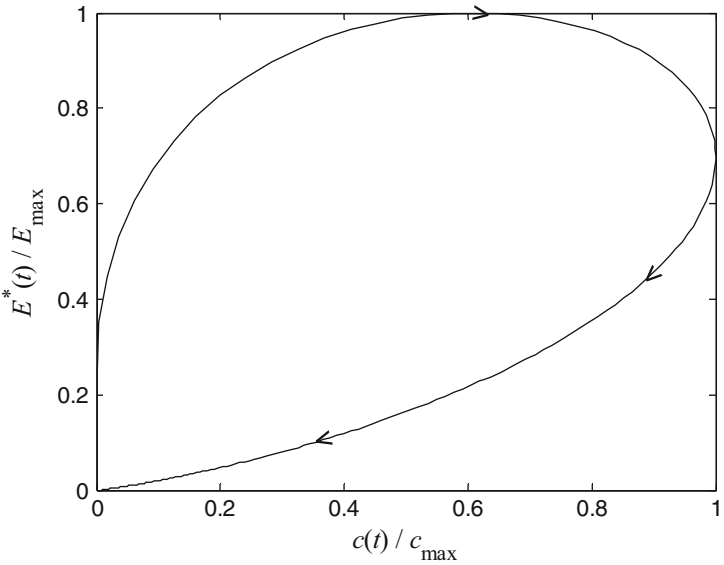


Fig. 12.4 Normalized effect–plasma drug concentration state space for tolerance phenomena. As time flows (indicated by arrows) a clockwise hysteresis loop is formed

include periodicity in the pharmacokinetic-dynamic modeling. This is usually done by empirical periodic functions directly on the baseline, such as trigonometric functions, for example. An example is the effect of fluticasone propionate on cortisol [605]. All models associated with these phenomena are called time-variant.

12.4.3.1 Drug Tolerance

This phenomenon is characterized by a reduction in effect intensity after repeated drug administration. The explanation for the diminution of the effect as a function of time is attributed either to a decrease in receptor affinity or a decrease in the number of receptors. These changes result in a clockwise hysteresis loop when the effect is plotted vs. the plasma concentration, Figure 12.4. Usually, tolerance phenomena are discussed with respect to the E_{max} model. In this case, tolerance is associated with either a decrease in E_{max} over time or an increase in Ec_{50}^{γ} over time (12.8). An example of this kind of time dependency is the work of Meibohm et al. [615] on the suppression of cortisol by triamcinolone acetonide during prolonged therapy.

Apart from the decrease in the number or affinity of the receptors, more complex mechanisms have been proposed for tolerance phenomena. In the so-called *counterregulation* models, the development of tolerance is driven by the primary effect of drug perhaps via an intermediary transduction step. This mechanism was postulated by Bauer and Fung [616] for hemodynamic tolerance to nitroglycerine.

According to these authors, initial nitroglycerin-induced vasodilatation controls the counterregulatory vasoconstrictive effect. Moreover, the *desensitization of receptors* can reduce the drug effects on prolonged exposure. The receptor-inactivation theory [617] can be used to model this mechanism.

12.4.3.2 Drug Sensitization

This term is used to describe the increase in pharmacological response with time to the same drug concentration. The up-regulation of receptors is considered to be the primary cause for sensitization. This phenomenon is observed when the negative feedback of an agonist is removed. A clinical example of sensitization is the chronic administration of beta-blockers, which induces up-regulation of beta-adrenoreceptors. This leads to increased adenyl cyclase activity and hypersensitivity to catecholamines after sudden withdrawal of the antagonist [618]. Due to the increase of the effect over time in sensitization phenomena, the effect–plasma concentration plots have a counterclockwise hysteresis loop, Figure 12.2.

12.4.4 Dynamic Nonlinear Models

Using the approach of Sheiner and Verotta [619], a large number of pharmacodynamic models can be considered as hierarchical models composed of a series of submodels. These submodels are linear or nonlinear, static or dynamic input–output, elementary models. Several possible combinations of such submodels have been considered, but they have systematically associated the linear with dynamic features, and the nonlinear with static ones. Is there hesitation or fear of using nonlinear dynamics in the traditional pharmacokinetic-dynamic modeling context?

The most interesting case arises by removing assumption **H1**, i.e., when the reaction between drug and receptor is not at equilibrium [595]. This happens when relatively slow rates of association and dissociation of the complex are observed. Under these conditions, a slow dynamic receptor-binding model is most applicable. By maintaining the proportionality between the effect and the concentration of the drug–receptor complex, (12.4) can be written in terms of the effect

$$\dot{E}(t) = k_{+1} c^{\gamma}(t) [E_{\max} - E(t)] - k_{-1} E(t), \quad E(0) = 0. \tag{12.16}$$

This equation is a nonlinear differential equation describing the time course of the effect and using an intrinsic pharmacodynamic time. An application of this model can be found in the work of Shimada et al. [596], who applied the drug–receptor nonequilibrium assumption to model the pharmacodynamics of eight calcium channel-blocking agents in hypertensive patients on the basis of their in vitro binding data. This model is rarely used because it produces profiles similar to the

Table 12.1 Assumptions and operable equations for the pharmacokinetic-dynamic models. The hysteresis column "Hyster" refers to the graph of the effect–concentration plot.

Model	Prereceptor	Receptor	Postreceptor	Hyster
Empirical E_{max}	None	12.8	None	No
Indirect link	12.11	12.13	None	Yes
Indirect response	Equilibrium	12.15	12.14	Yes
Transducer	None	None	$E(t) = \mathcal{T}(\upsilon)$	-
Nonlinear	Equilibrium	12.16	None	Yes

indirect-link model described above. However, the drug–receptor nonequilibrium model has more theoretical and practical interest since more complex solutions are also feasible by adding a feedback component to the effect of the drug [620]. The resulting model has a very rich dynamic behavior and is the essence of Chapter 13.

Table 12.1 summarizes assumptions and equations for the above-presented pharmacodynamic models.

12.5 Unification of Pharmacodynamic Models

Historically, delays between drug exposure and effect have been described with the so-called effect-compartment model, first described by Segre [582] and popularized by Sheiner and coworkers [583, 584]. Recently, Dayneka [598] focused attention on a set of indirect-effect models to introduce intrinsic pharmacodynamic time. The relevance of combined pharmacokinetic-dynamic modeling has been largely recognized [621, 622]. The discussion in Section 12.3 indicates that the development of the various pharmacokinetic-dynamic models was based on the dominating assumption for one of the drug processes depicted in Figure 12.1. Thus, the pharmacokinetic-dynamic models can be classified kinetically on the basis of the assumptions associated with:

- the prereceptor equilibrium,
- the drug–receptor interaction, and
- the postreceptor equilibrium.

A very general scheme for relating effects to concentration, of which both the effect-compartment and the indirect-effect models are special cases, was outlined by Sheiner and Verotta [619]. The models presented in the study can be considered to be a special case of that unified scheme. As judiciously presented by these authors, both direct-response and indirect-response models are composed of one nonlinear static submodel and one dynamic submodel, but the placement of the submodels in the global model differs:

- In a direct-response model, the output of a linear dynamic model (the link model) with input $c(t)$ drives a nonlinear static model (usually the E_{max} model) to produce the observed response.
- In an indirect-response model, the above order of models is reversed and now the static model precedes the dynamic one. The dynamic model describes the formation and loss of the response variable through a linear differential equation whose parameters are nonlinear saturable forms of the driving concentration $c(t)$.

All these models introducing the prereceptor and postreceptor events have an interesting appeal with respect to physiologically implied mechanisms. Sheiner and Verotta [619] pointed out the importance of knowing where the rate-limiting step is located in a series of events from pre- to postreceptor drug interactions.

The fundamental assumption and equations governing the effect–concentration relationship for each one of the models considered are listed in Table 12.1. The presence or not of a hysteresis loop in the effect–plasma concentration plot of each model is also quoted in Table 12.1. At present, the methodology for performing efficient pharmacokinetic-dynamic modeling is well established [572, 623, 624].

Chapter 13
Nonclassical Pharmacodynamics

The whole is more than its parts.

Aristotle (384–322 BC)
Metaphysics

Whereas the concentration of a drug depends on the administration protocol and on intrinsic pharmacokinetic characteristics of the process, endogenous substances are certainly regulated by appropriate internal control mechanisms. For example, the neurotransmitter norepinephrine is released from sympathetic nerve endings and its concentration is regulated by enzymes and by a mechanism for reuptake of this catecholamine into nerve endings. Deficiencies in the control of such important chemicals may result in vasospasm, spasticity, and a variety of behavioral abnormalities. Such observations strongly suggest the existence of control systems represented by negative feedback mechanisms. By means of those mechanisms, the dynamic system controls the local concentration of critical endogenous chemicals that interact with receptors according to the mass-action law. Indeed, the biomedical literature, particularly that of functional and biochemical pharmacology, is rich with detailed descriptive mechanisms of control and its modification induced by an extensive list of drugs and chemicals. However, mathematical analysis of such control is virtually nonexistent in the pharmacological literature. In contrast, there has been a steady evolution of concepts of control theory and dynamic modeling in many areas of physiology, elegantly traced by Glass and Mackey [32], with an extension by these authors to physiopathological states [49].

As stated in Chapter 12, when the drug–receptor interaction involves feedback, the system becomes more complex. Hence, we will first present modeling and associated mathematical analysis of three typical processes. This will be followed by several applications involving drug pharmacodynamics organized around pharmacotherapy with drugs affecting the endocrine, central nervous, and cardiovascular systems.

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_13

13.1 Nonlinear Concepts in Pharmacodynamics

First, we provide a typical stability analysis for a model involving a nonlinear mechanism, which links pharmacodynamics with pharmacokinetics. Also, we provide new insights that may aid in understanding the variety of oscillations displayed in biological systems and how they may be related to the maintenance or loss of control in such systems. Examples of periodic phenomena abound in biological systems, in many cases due to fluctuations of ligand interacting with a receptor.

Homeostatic regulation will be studied as represented by negative feedback mechanisms. First, it will be shown how the properties of negative feedback are related to the geometric properties of the binding and control curves in a ligand–receptor interaction and, further, how changes in their geometry affect the system's response to variations of the ligand release. Second, in the analysis of the hemopoietic chain, negative feedback is supplemented by a lag time that leads to bifurcations and oscillatory, chaotic behavior.

For these analyses, the procedure is to use dimensionless parameters in the set of differential equations describing the model, look for the steady state, investigate the linear stability, and determine the conditions for instability. Near the bifurcation values of the parameters, which initiate an oscillatory growing solution, perturbation analysis provides an estimate for the period of the ensuing limit cycle behavior.

13.1.1 Nonlinear Mechanism

As an indirect-link model, the effect-compartment model has been frequently revealed not flexible enough to fit experimental data, nor sensitive enough with respect to dosage modifications in order to adjust regimens in clinical pharmacokinetics. Some others were subsequently introduced to offer the required properties. The *interface model* introduced first by Meille et al. [625] is an alternative of the biophase effect-compartment model.

The interface model receives as input the pharmacokinetic concentrations $c(t)$ and elaborates the *exposition variable* $y(t)$ which drives the pharmacodynamic effect. The interface model is described by the following differential equation

$$\dot{y}(t) = -\kappa \exp[-\alpha y(t)]\, y(t) + \beta [c(t) - \gamma] H[c(t) - \gamma] \tag{13.1}$$

where $H[.]$ is the Heaviside step function. The model parameter

- α controls the extent of nonlinearity; when $\alpha = 0$, the model becomes the traditional biophase compartment model;
- β quantifies the influence of pharmacokinetics on the interface model; and
- γ is a threshold on the circulating drug concentrations.

13.1.1.1 Stability Analysis

The section presents successively the reparametrization of the model to obtain dimensionless parameters, the search of equilibrium point, the investigation of the linear stability, and the determination of the conditions for instability.

The proposed reparametrizations are

$$\tau = \kappa t \quad \eta(\tau) = c(\tau) - \gamma \quad \psi(\tau) = \alpha y(\tau)$$

and the model becomes

$$\dot{\psi}(\tau) = -\exp[-\psi(\tau)]\,\psi(\tau) + \xi\eta(\tau) \tag{13.2}$$

with $\xi = \alpha\beta/\kappa$.

The equilibrium points ψ^* depend on the extent of $\xi\eta(\tau)$ term and they are roots of the following equation

$$\exp\left(-\psi^*\right)\psi^* = \xi\eta(\tau) \tag{13.3}$$

obtained from the differential equation 13.2 by setting $\dot{\psi}(\tau) = 0$. The maximum of $\psi\exp(-\psi)$ occurs for $\psi = 1$. Equation 13.3 may have no solution, one solution, or two solutions when $\xi\eta(\tau)$ is higher, equal, or less than $\exp(-1)$, respectively.

For linear stability, the local behavior of the differential equation 13.2 near ψ^* is determined by linearizing at ψ^*

$$\frac{\mathrm{d}}{\mathrm{d}\psi(\tau)}\left[-\exp[-\psi(\tau)]\,\psi(\tau) + \xi\eta(\tau)\right]\Big|_{\psi(\tau)=\psi^*} = \exp\left(-\psi^*\right)\left(\psi^* - 1\right) \tag{13.4}$$

The above expression is the eigenvalue ζ of the system evaluated at the equilibrium points and it states about the stability conditions. So, the conditions for instability are connected with the kind of roots ψ^* of the equation 13.3:

- if $\xi\eta(\tau) > \exp(-1)$, no real solutions and the system is unstable;
- if $\xi\eta(\tau) = \exp(-1)$, one saddle-point solution $\psi^* = 1$ and the system is unstable;
- if $\xi\eta(\tau) < \exp(-1)$, two real solutions and the sign of $\psi^* - 1$ determines the stability conditions. If $\psi^* < 1$, the model is stable; otherwise, it is unstable. For the stable solution ψ_1^*, the eigenvalue 13.4 is negative and for the unstable solution ψ_2^*, it is positive. The positive eigenvalue is the Lyapunov exponent that is connected with the sensitivity with respect to the initial conditions.

13.1.1.2 Initial Conditions

The role of initial conditions was not mentioned in the above analysis. In fact, the stability was studied after a long time, sufficient enough to drive the system toward the equilibrium points. Now, we propose to examine the trajectories starting from different initial conditions to end at the equilibrium points. The behavior of the differential equation 13.2 is examined when the initial condition is chosen near the equilibrium points ψ^*, i.e., the sign of $\dot{\psi}(\tau)$ is studied near $\psi(0) = \psi^* + \delta$. Here, δ is a small quantity and $\xi\eta(\tau)$ is assumed constant $\xi^* \triangleq \xi\eta(\tau)$. From the above

$$\dot{\psi}(\tau)\Big|_{\tau=0} = \exp\left[-\left(\psi^* + \delta\right)\right]\left(\psi^* + \delta\right) + \xi^*$$

and by expanding the above expression, by using equation 13.3, and by approximating $\exp(\delta) \approx 1 + \delta$, we obtain

$$\dot{\psi}(\tau)\Big|_{\tau=0} = \exp\left(-\psi^*\right)\left(\psi^* - 1\right)\delta.$$

- If $\psi^* < 1$ (i.e., $\psi(0)$ near to ψ_1^*), $\dot{\psi}(\tau)\Big|_{\tau=0} < 0$ for $\delta > 0$ and $\dot{\psi}(\tau)\Big|_{\tau=0} > 0$ for $\delta < 0$, i.e., the trajectory is attracted toward the stable equilibrium point ψ_1^*.
- If $\psi^* > 1$ (i.e., $\psi(0)$ near to ψ_2^*), $\dot{\psi}(\tau)\Big|_{\tau=0} < 0$ for $\delta < 0$, i.e., the trajectory is leaving the instable ψ_2^* neighborhood and it reaches lower $\psi(\tau)$ values.
- If $\psi^* > 1$ (i.e., $\psi(0)$ near to ψ_2^*), $\dot{\psi}(\tau)\Big|_{\tau=0} > 0$ for $\delta > 0$, i.e., the trajectory reaches higher to $\psi(\tau)$ values, it is diverging and becomes unstable.

In summary, ψ_2^* is a critical point for the initial conditions, if $\psi(0) < \psi_2^*$, the trajectory ends up to ψ_1^*, otherwise the trajectory becomes unstable. This comments the sensitivity of the system trajectories with respect to the initial conditions.

For instance, with $\xi^* = 0.23$, Figure 13.1 presents graphically the position of the two real solutions $\psi_1^* = 0.3152$ and $\psi_2^* = 2.3046$. Also, Figure 13.2 presents the trajectories starting near $\psi_2^* = 2.3046$ for different values of δ.

The Lyapunov exponent is the eigenvalue ζ evaluated at the unstable equilibrium point ψ_2^*, $\zeta = \exp\left(-\psi_2^*\right)\left(\psi_2^* - 1\right) = 0.1302$. It is named 'exponent' because the same value ζ is the λ parameter in $\exp(\lambda t)$. This increasing exponential term quantifies the time profile of the difference between two trajectories starting from nearly located initial conditions. Based on this property, the Lyapunov exponent was calculated at $\lambda = 0.13014$ close to $\zeta = 0.1302$ by fitting the simulated trajectories in the time interval up to 115 h.

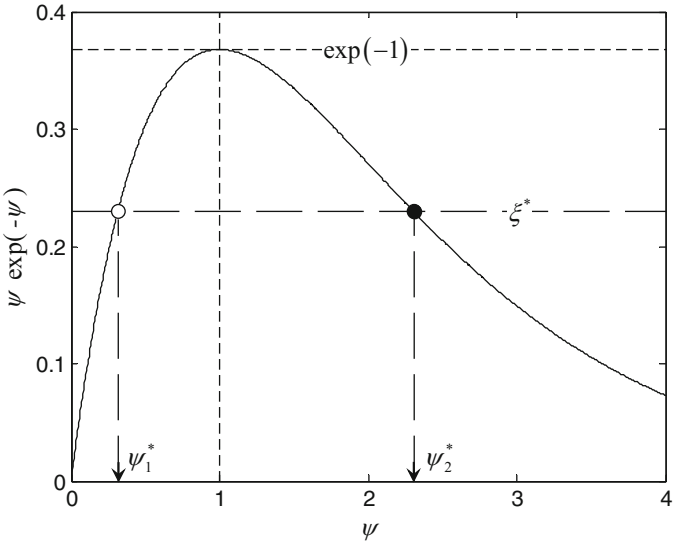


Fig. 13.1 Stability analysis for the interface model for $\xi^* = 0.23$. Open and full circles represent the stable ψ_1^* and the unstable ψ_2^* solutions, respectively

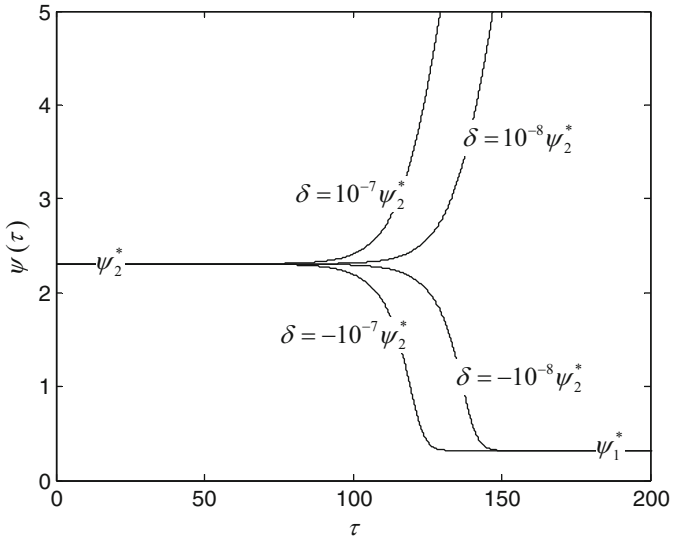


Fig. 13.2 Sensitivity of the system trajectories to the initial conditions. The initial conditions start near the critical point ψ_2^*

13.1.1.3 Unexpected Toxicity

In the above analyses, the input function in the interface model was assumed constant, $\xi^* \triangleq \xi\eta(\tau)$. In this section the input will be considered variable in time, as it is driven by the pharmacokinetic concentration time profile. The model involves the pharmacokinetic, interface, and pharmacodynamic components as follows

$$\begin{aligned} \dot{c}(t) &= -kc(t) + \tfrac{u(t)}{V} & c(0) &= 0 \\ \dot{\psi}(t) &= -\exp[-\psi(t)]\,\psi(t) + \xi c(t) & \psi(0) &= 0 \end{aligned}$$

- The pharmacokinetic component is a simple one-compartment model describing the concentration profile $c(t)$ with parameters, the volume of distribution $V = 100$ l and elimination rate $k = 0.05\ \mathrm{h}^{-1}$. The drug is assumed to be given in repeated administrations by oral route with absorption rate $k_a = 1\ \mathrm{h}^{-1}$. This administration protocol is adequately described by the $u(t)$ function.
- The interface model uses the $\xi = 0.23$ parameter and it generates the normalized exposure variable $\psi(t)$. The differential equation 13.1 was reparametrized by selecting $\alpha = \kappa = 1$, $\beta = 0.23$ and the concentration threshold was set at 0, $\gamma = 0$.
- The pharmacodynamic component is a simple sigmoidal function that quantifies the risk of toxicity

$$\rho(t) = \frac{\psi(t)}{\psi_{50} + \psi(t)}.$$

  According to this model, $\rho(t)$ is driven by the exposure variable $\psi(t)$ generated by the interface model and involves the parameter $\psi_{50} = 10$, which is the exposure required to induce a risk of 50%.

The dose used was fixed at 180 mg daily. Two administration schedules were studied, twice in day (b.i.d.) and once in day (o.i.d.). For the two schedules, the average concentration remains the same, $c_{ave} = 1.5\ \mathrm{mg\,l}^{-1}$, but the fluctuations

$$\frac{c_{\max} - c_{\min}}{c_{ave}}$$

are 41.62 and 95.02% for the b.i.d. and o.i.d. schedules, respectively.

By using the model, simulations were made to show that, when instability conditions are met within the interface model, spontaneous toxicity could be observed as pharmacodynamic outcome. The stability graphics are first presented in Figure 13.3. In this graphics, the input function of the interface model $\xi c(t)$ is plotted by solid line; the $\psi(t)\exp[-\psi(t)]$ function and the $\exp(-1)$ threshold are plotted by dashed lines. It is important to mention that considering the overall pharmacokinetic behavior materialized by the average concentration c_{ave}, the input function $\xi c_{ave} = 0.345$ leads to levels lower than the $\exp(-1) = 0.3679$ threshold, thus preventing the instability of the system.

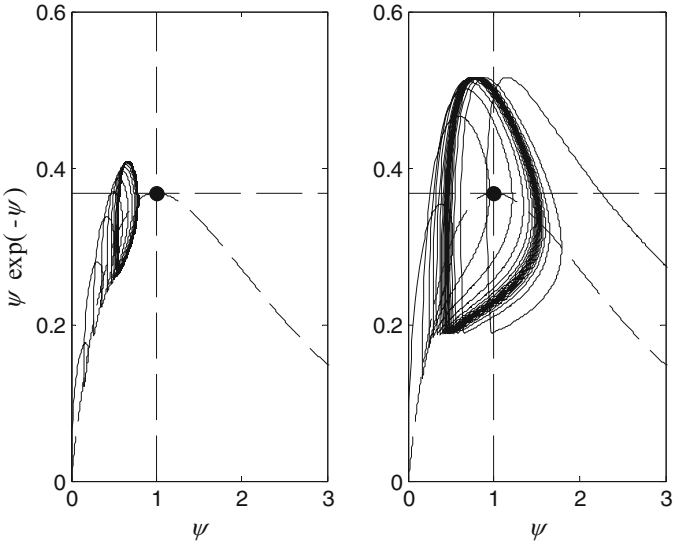


Fig. 13.3 The stability graphics for the b.i.d. (left panel) and the o.i.d. schedules (right panel) for a daily dose of 180 mg

Because of the repeated administrations, the pharmacokinetic profile is varying periodically and the trajectories of the input are represented in a spiral form. For the two administration schedules, the input function using high concentration levels exceeds the threshold of $\exp(-1)$, and this could induce some instabilities. Since the concentration profile fluctuations for the o.i.d. are two-fold larger than that for the b.i.d. schedule, this behavior is more pronounced for the o.i.d. (Figure 13.3, right panel) than that for the b.i.d. schedule (Figure 13.3, left panel). Moreover, important segments of the trajectories of the input function in the o.i.d. schedule are projected in the $\psi(t) > 1$ domain, where the equilibrium points become unstable. This should result in an unstable behavior of the o.i.d. as compared with the b.i.d. administration schedule.

The time profiles of the exposure variable $\psi(t)$ for both administration schedules are depicted in Figure 13.4. In fact, the b.i.d. schedule (low-level profile) is not sufficiently powerful to trigger instability; in opposite, the o.i.d. schedule (high-level profile) induces uncontrolled exposure after 10 d of repeated drug administrations. Again in the phase plane concentration–exposure (Figure 13.5), the hysteresis curves obtained with the o.i.d. schedule are confined in the low dashed-line-box for the early times (first 10 d) and, for the longer times, the hysteresis curves jump spontaneously to 100-fold higher levels (Figure 13.5). This instability has severe consequences regarding the risk for toxicity $\rho(t)$ as described by the sigmoidal function (Figure 13.6). The b.i.d. schedule can induce toxicity risks up to 7%, but the o.i.d. schedule after 10 d of daily repeated doses induces a risk that can exceed 90% in a spontaneous rather than gradual manner.

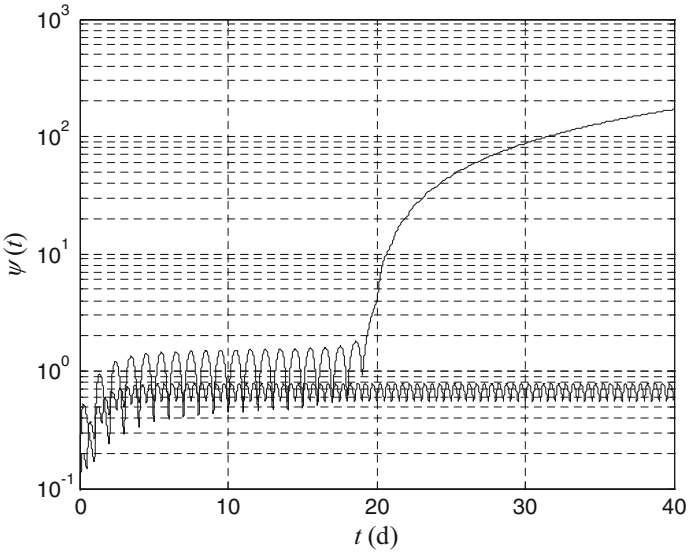


Fig. 13.4 The time-exposure profiles for the b.i.d. (low-level profile) and o.i.d. schedules (high-level profile)

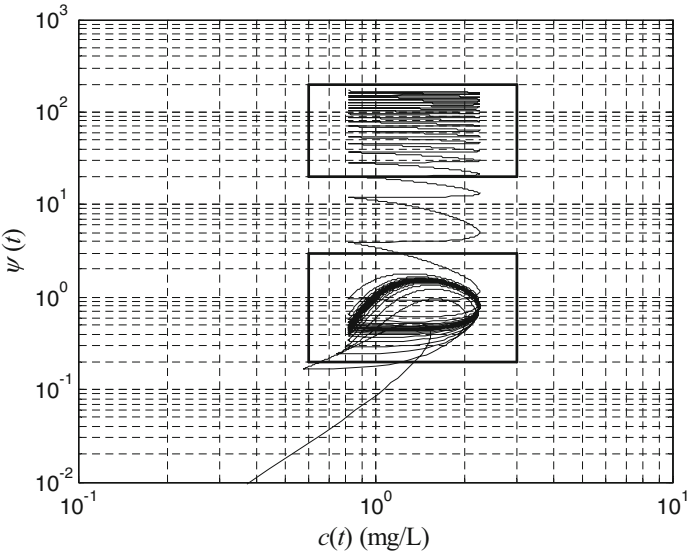


Fig. 13.5 The hysteresis curves depicted into the $\psi(t)$ vs. $c(t)$ phase plane

The interface model is a nonlinear effect-compartment model, in which the elimination rate $\kappa \exp[-\alpha y(t)]$ is controlled by the state $y(t)$. This nonlinear feedback mechanism induces a highly versatile exposure variable controlling pharmacodynamics. The potential instability results from the saturation of this

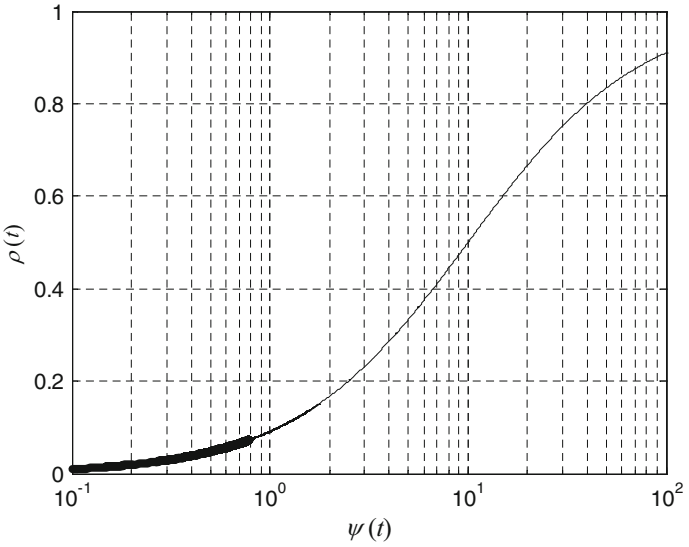


Fig. 13.6 The sigmoidal function toxicity risk $\rho(t)$ vs. exposure variable $\psi(t)$ for the b.i.d. (*thin line*) and for the o.i.d. (*thick line*) schedules.

nonlinear mechanism triggered by high drug levels $c(t)$ although for short time period and periodically. In fact, several references support the instability and the spontaneous toxicity in pharmacodynamics [626].

13.1.2 Negative Feedback

The interaction of a drug or an endogenous ligand with a specific receptor is most often modeled as a bimolecular reversible reaction

$$[\text{ligand}] + [\text{receptor}] \underset{k_{-1}}{\overset{k_{+1}}{\rightleftarrows}} [\text{ligand-receptor complex}],$$

to which the mass-action law applies. This is the classic model presented elsewhere (8.8), (12.1). It is the basis of most studies aimed at quantitatively characterizing receptors with specific radioligands as well as in functional studies where the effect is related to receptor occupancy [617].

For a concentration $c(t)$ of the drug or ligand and a total receptor concentration r_0, we thus have

$$\dot{\upsilon}(t) = k_{+1} c(t) [r_0 - \upsilon(t)] - k_{-1} \upsilon(t), \quad \upsilon(0) = 0, \tag{13.5}$$

where $\upsilon(t)$ is the concentration of the ligand–receptor complex, and k_{+1} and k_{-1} are the forward and reverse rate constants, respectively, of the reaction. The main features for the ligand model are that ligand is continuously released at rate $u(t)$, and eliminated exponentially with a rate constant k, and that there exists a negative feedback control function $\Phi(\upsilon)$ that depends on the concentration of occupied receptor $\upsilon(t)$ that modulates the release; thus,

$$\dot{c}(t) = -kc(t) + \Phi(\upsilon) + u(t), \quad c(0) = 0, \tag{13.6}$$

with $\Phi(\upsilon) \geq 0$ and $\mathrm{d}\Phi(\upsilon)/\mathrm{d}\upsilon \triangleq \Phi'(\upsilon) \leq 0$. The model is based on evidence, obtained largely from studies of the release of neurotransmitters, that the quantity of ligand released per unit time is modulated by the nerve terminal itself as a result of stimulation by the neurotransmitter of a subset of the receptors termed "autoreceptors" [617]. Thus, receptor stimulation not only produces effects but also inhibits or augments release, thereby maintaining a basal level of the ligand. The feedback signal may originate at a site other than the occupied receptor; however, it is functionally related to $\upsilon(t)$.

We make the variables of the above equations dimensionless:

$$\lambda = \frac{k_{+1}}{k}, \qquad \mu = \frac{k_{-1}}{k}, \qquad \kappa = \frac{k_{-1}}{k_{+1}},$$
$$y(\tau) = \frac{\upsilon(\tau)}{r_0}, \qquad \phi(y) = \frac{\Phi(y)}{k}, \qquad \rho(\tau) = \frac{u(\tau)}{k},$$

with $\tau = kt$. The set of differential equations becomes

$$\begin{aligned} \dot{y}(\tau) &= \lambda c(\tau)[1 - y(\tau)] - \mu y(\tau), \quad y(\tau) = 0, \\ \dot{c}(\tau) &= -c(\tau) + \phi(y) + \rho(\tau), \qquad c(\tau) = 0. \end{aligned} \tag{13.7}$$

Equilibrium points (y^*, c^*) of the system are those for which $\dot{y}(\tau) = \dot{c}(\tau) = 0$.

- Equating $\dot{y}(\tau)$ to zero, we have the *binding curve*. The binding curve and its slope are given by

$$c_B = \kappa y/(1 - y) \text{ and } \mathrm{d}c_B/\mathrm{d}y = \kappa/(1 - y)^2 \tag{13.8}$$

respectively.

- Equating $\dot{c}(\tau)$ to zero, we have the *feedback curve*. The feedback curve and its slope are given by

$$c_F = \phi(y) + \rho(\tau) \text{ and } \mathrm{d}c_F/\mathrm{d}y = \phi'(y) \tag{13.9}$$

respectively.

Equilibrium points of the system are the intersections of the binding curve with the feedback curve, i.e., $c^* = c_B = c_F$. Their location in the state space depends on κ and $\rho(\tau)$ (equations 13.8 and 13.9).

The stability of equilibrium points is determined by standard stability analysis (cf. Appendix A). The Jacobian matrix of the linearized system,

$$\mathbf{A}(y) = \begin{bmatrix} -(\lambda c + \mu) & \lambda(1-y) \\ \phi'(y) & -1 \end{bmatrix},$$

supplies the eigenvalues ζ_1 and ζ_2. Given (13.8) and (13.9), these eigenvalues are

$$\zeta_{1,2} = \frac{1}{2}\left[-(\nu+1) \pm \sqrt{(\nu+1)^2 + 4\lambda(1-y)\left(\phi'(y) - \frac{\mathrm{d}c_B}{\mathrm{d}y}\right)}\right], \qquad (13.10)$$

where

$$\nu = \lambda c + \mu = \frac{\mu}{1-y}.$$

Since $y < 1$ at any equilibrium point, it follows that a negative feedback $\phi'(y) \leq 0$ ensures that the second term under the radical is negative, so that the eigenvalues are real and negative or complex with a negative real part; hence, such an equilibrium point is stable. For a large negative value of $\phi'(y)$ the eigenvalues are complex and the point is a stable focus. A shallow negative slope gives two real negative eigenvalues and thus a stable node. In the previous equation it is seen that the eigenvalues do not depend on the normalized ligand input rate $\rho(\tau)$.

In simple negative feedback, $\phi(y)$ is a monotone decreasing function and the equilibrium point is unique. However, due to a variety of factors, it is expected that at very low ligand–receptor numbers, $\phi(y)$ becomes an increasing function, implying a positive feedback. The feedback becomes a mixture of positive and negative feedbacks, called *mixed feedback*, and it has been reported elsewhere [627]. Positive slopes in $\phi(y)$ generate other equilibrium points [620, 628, 629]. The characterization of the eigenvalues in these new equilibrium points, and hence the stability of the system, follows from the application of the following theorem (cf. Appendix G): "The derivatives of two successive intersection points between two continuous functions, one of which is monotone, have opposite signs."

Application of this theorem permits analysis of the equilibrium points of the system with a monotone binding curve. If in the equilibrium point $P_1 = (y_1^*, c_1^*)$ we have $\phi'(y_1^*) < 0$, the system is stable. If the feedback curve is assumed to be continuous over a domain of permissible values of receptor occupancy $y(\tau)$, in the nearest equilibrium point $P_2 = (y_2^*, c_2^*)$, we will have $\phi'(y_2^*) > 0$. This condition is necessary but not sufficient for the instability of the system. But if moreover $\phi'(y_2^*) > \mathrm{d}c_B^*/\mathrm{d}y^*$, one eigenvalue from (13.10) is positive and the system becomes unstable at P_2, which is an unstable saddle point or repellor.

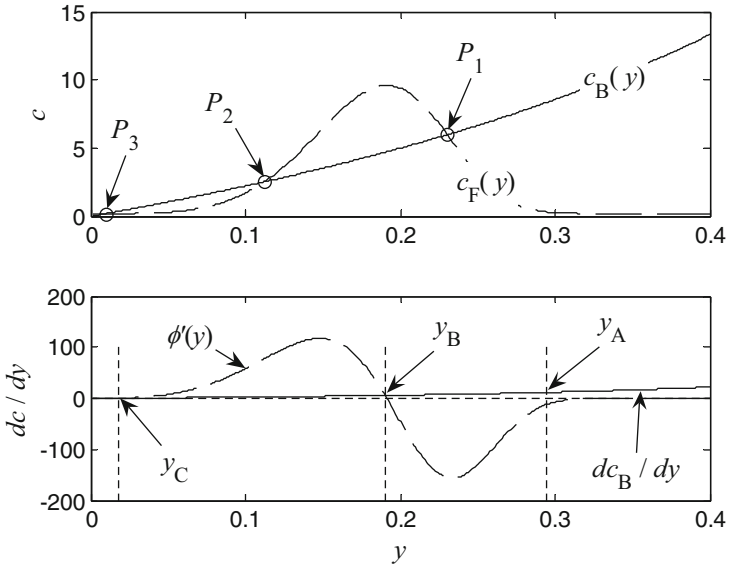


Fig. 13.7 Upper panel: Binding curve (*solid line*) and the intersections P_1, P_2, and P_3, with the feedback curve (*dashed line*) to give the equilibrium points. Lower panel: the slope graphical analysis determines y_A, y_B, and y_C, which are intersections of slopes $\phi'(y)$ (*dashed line*) and dc_B/dy (*solid line*) helping us to analyze the stability graphically

Example 11. Stability with Feedback

We use a feedback curve similar to the Weibull distribution with both scale and shape parameters equal to 5; for the binding curve, we set $\kappa = 20$ and $\rho(\tau) = 0.2$. In the state space, Figure 13.7 illustrates the position of the three equilibrium points, $P_1 = (0.2304, 5.9878)$, $P_2 = (0.1119, 2.5204)$, and $P_3 = (0.0099, 0.2002)$ (Figure 13.7A). The graphic slope analysis determines $y_A = 0.2947$, $y_B = 0.1901$, and $y_C = 0.0182$ (Figure 13.7B), which are threshold values for $\phi'(y)$ and $\mathrm{d}c_B/\mathrm{d}y$ comparison. Thus, according to (13.10), we note that for:

- $y^* < y_C$, the equilibrium is stable since $\phi'(y) < \mathrm{d}c_B/\mathrm{d}y$; since $y_3^* < y_C$, P_3 is stable;
- $y_C \leq y^* < y_B$, the equilibrium is unstable since $\phi'(y) > \mathrm{d}c_B/\mathrm{d}y$; since $y_C < y_2^* < y_B$, P_2 is unstable;
- $y_B \leq y^* < y_A$, the equilibrium is a stable focus since $\left|\phi'(y)\right| \gg \mathrm{d}c_B/\mathrm{d}y$; since $y_B \leq y_1^* < y_A$, P_1 is a stable focus; and
- $y_A \leq y^*$, the equilibrium is stable since $\left|\phi'(y)\right| < \mathrm{d}c_B/\mathrm{d}y$.

These results are supported by the standard stability analysis of Figure 13.8, where λ is set to 0.1 and $\mu = 2$ ($\mu = \kappa\lambda$). The eigenvalues computed by (13.10) are plotted as functions of y. In this figure, unstable and stable equilibrium points are

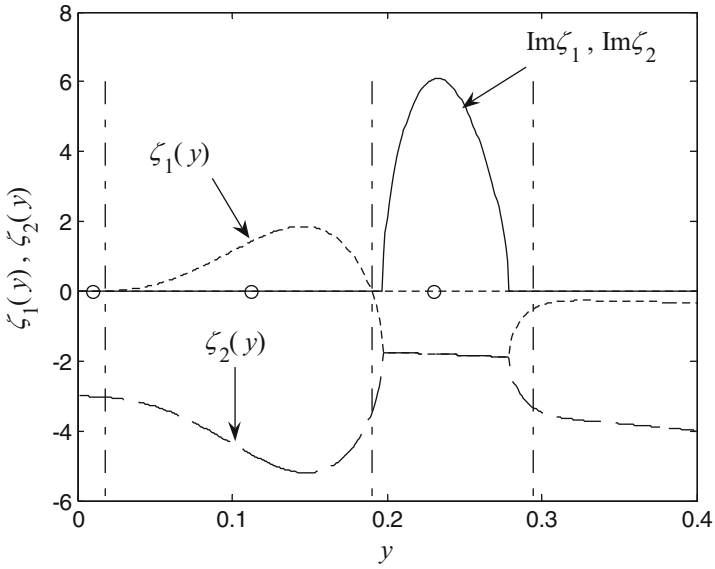


Fig. 13.8 Eigenvalues computed via (13.10). (∘) indicates the positions of equilibrium points P_1, P_2, and P_3, and vertical *dash-dotted lines*, the positions of y_A, y_B, and y_C. The *solid line* represents the imaginary parts, and *dashed* and *dotted lines* represent the real parts of eigenvalues

clearly separated by an interval, $[0.1974; 0.2790]$, where eigenvalues are complex, leading to a stable focus. With increasing λ, this interval becomes narrower and for $\lambda > 0.65$, the eigenvalues have only real parts.

Finally, Figure 13.9 illustrates the dynamics of the model described by (13.7) when a different initialization is used. The unstable P_2 point is actually a repellor, P_1 is a stable focus, and P_3 is stable. Depending on the magnitude of the release of ligand from intracellular storage sites or exogenous administration, and the location of the unstable point, the state vector will either return to the operating point or be propelled to extreme states. Such propulsion means loss of control in the effector system governed by occupation of the receptor. The distance between the stable and the unstable points is crucial in this regard. Equally important are the eigenvalues at the control point. When these are complex, the recovery follows a path that is closer to the unstable point.

It is also interesting to comment on the influence of the parameters on the system behavior:

1. The elimination rate constant k is a time-scale parameter and it does not affect the present stability analysis.
2. The forward and reverse rate constants, k_{+1} and k_{-1}, respectively, govern the affinity. A change in affinity, with the same receptor density, also affects stability. A decrease in affinity (decrease in λ or increase in κ expressing the dissociation constant) results in an increase in distance between P_1, P_2, and P_3. The parameter κ affects exclusively the binding curve.

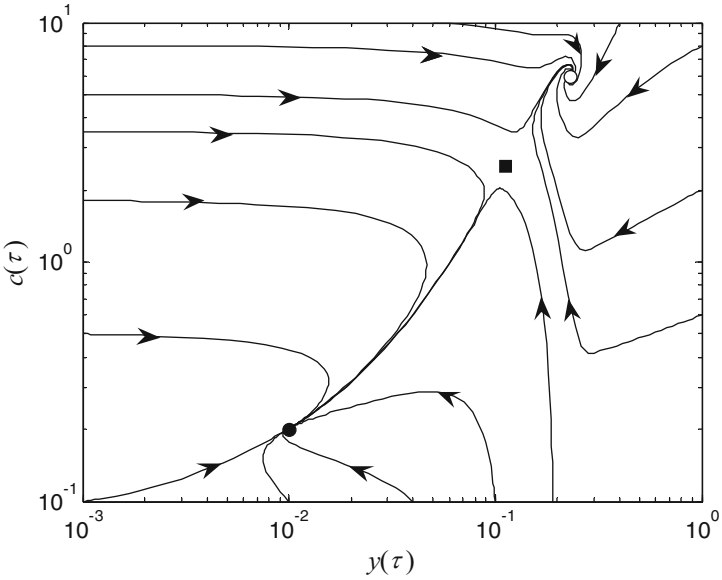


Fig. 13.9 State space for different initial conditions. The equilibrium point P_1 (∘) is a stable focus, P_2 (■) is an unstable saddle point, and P_3 (●) is stable

3. The steep negative slope $\phi'(y) \leq 0$ results in complex eigenvalues. The frequency of the oscillation increases with the steepness. The operating point in such cases is a stable focus. In contrast, shallow negative slopes are indicative of a nonoscillating operating point or stable node. Since eigenvalues do not depend on $\rho(\tau)$, figures like Figure 13.8 are useful for following the positions of equilibrium points in the state space when $\rho(\tau)$ varies.
4. When fixed, the number of receptors r_0 is a scale parameter and it does not affect the stability. When r_0 is not fixed but changes as a result of pharmacological interventions or pathological states, the operating point will, of course, change.

When $u(t)$ and r_0 are varying and the other parameters are fixed, simulations (not presented here) with (13.5) and (13.6) reveal that a decrease in $u(t)$ results in a decrease in distance between P_2 and P_3. Conversely, an increase in r_0 results in a decrease in distance between P_1 and P_2. ■

The above-mentioned theorem allows speculation about a monotone feedback curve and a nonlinear binding curve. Their intersections will have derivatives with alternate signs, and therefore, they lead to stable and unstable equilibrium points. In this sense, Tallarida [629] used a U-shaped feedback curve to analyze experiments involving neurotransmitter norepinephrine systems [630].

Analysis on the state space proved to be very useful and demonstrated how possible changes in receptor affinity or receptor number affect the distance between the operating point and the unstable equilibrium point, and thus the ability of the system to return to the operating point after a perturbation such as endogenous

release. The new information reported here pertains to the geometry on the state space, which allows us to predict both the stability of equilibrium points and the characteristic frequency from the slopes of both curves at their intersection. The relationship between the slope and the frequency of the system is especially important in the further development of models for particular receptor systems, since examples of rhythmic phenomena abound in biological systems.

As we proceed it will be seen that the most important results do not depend on a particular assumption regarding the form of the feedback function. Thus far we have not located the equilibrium points for the system under study because the function $\Phi(\upsilon)$ was kept general. The model we have used is applicable to both endogenous and exogenous substances.

In a series of contributions, Tallarida also studied the control of an endogenous ligand in the presence of a second compound (agonist or antagonist) that interacts with the same receptor [630] or under periodic release of the ligand [631]. That author showed by computer simulation how the parameters of the model affect the time course of released ligand resulting from administration of an antagonist and the suppression of such release when the second compound is an agonist.

A new quantitative concept that describes the feedback control of the dopaminergic system was also introduced, the control curve. Once known, the ligand's control curve has predictive value that may be useful in the design of efficient drug tests. These theoretical results were confirmed experimentally on numerous cases as for neurotransmitters, hormones, peptides, etc., whose concentrations in the various organs and tissues remain bounded. For example, the control of dopamine release by negative feedback was confirmed in the rat striatum [632]. A consequence of this model is that competitive antagonists augment dopamine release, whereas competing agonists reduce such release.

These findings may be of general importance since baseline parameters are crucial in determining pharmacodynamic responses [633], while feedback mechanisms are frequently involved in physiological processes, e.g., the secretion of hormones and the recurrent inhibitory pathway for γ-aminobutyric acid in the hippocampus, which has been described in almost every type of neural tissue ranging from the lowest invertebrates through humans [634], and the production of biotech products in humans [635].

13.1.3 Delayed Negative Feedback

An interesting case is that of a delay mechanism inserted in a closed loop process with negative feedback. The typical process is the hemopoietic process that incorporates some control elements that regulate homeostatically the rates of release of marrow cells to proliferation, maturation, and to the blood.

The dynamic response of the process of neutrophil granulocyte production to perturbations has been studied in a number of ways. In leukopheresis, neutrophils are removed from the blood artificially over a short span of time. Following

such an acute depletion of the neutrophil blood count, referred to as a state of *neutropenia*, neutrophils rapidly enter the blood from the marrow and produce an abnormally large number of neutrophils in the blood, or a state of *neutrophilia* [636]. The magnitude of the neutrophil blood count seen in such a state is about $2-3$ times normal. Such observations suggested the presence of some mechanisms for regulating the release of marrow neutrophils in response to the number of neutrophils circulating in the blood [637].

The most notable feature of the dynamic response of the process to large perturbations in the number of blood cells is that the system "rings," displaying an oscillatory behavior in the number of cells in the blood and other compartments of the system, as a function of time. Such large perturbations are produced by leukopheresis or exposure of the system to disease, which depletes the number of blood cells, and in total body irradiation experiments or some drug treatments such as chemotherapy, which deplete the total number of cells in the production process of neutrophil granulocytes.

Besides these perturbations on the hemopoietic process, there are some diseases, collectively referred to as the *periodic diseases*, in which symptoms recur on a regular basis of days to months. The most common of these disorders are cyclic neutropenia (also known as periodic hemopoiesis) [638] and cyclic thrombocytopenia [639]. It has long been suspected that periodic hematologic diseases arise because of abnormalities in the feedback mechanisms that regulate blood cell number [640–642]. But in a dynamic feedback process such as hemopoiesis it is difficult to distinguish between cause and effect. Oscillations occurring in one cell stage may induce cycling in other stages via feedback regulation. The mechanisms regulating neutrophil production are not as well understood. The important role of the cytokine granulocyte colony-stimulating factor (G-CSF) for the in vivo control of neutrophil production was demonstrated by Lieschke [643, 644]. Several studies have shown an inverse relation between circulating neutrophil density and serum levels of G-CSF [645]. Coupled with the in vivo dependency of neutrophil production on G-CSF, this inverse relationship suggests that the neutrophils would regulate their own production through negative feedback, in which an increase (decrease) in the number of circulating neutrophils would induce a decrease (increase) in the production of neutrophils through the adjustment of G-CSF levels. G-CSF has synergetic effects on the entry into cycling of dormant hemopoietic stem cells.

These observations have provided impetus for mathematicians to determine the conditions for the observed oscillations. Thus far, there have been two surprising discoveries [48, 49]:

- qualitative changes can occur in blood cell dynamics as quantitative changes are made in feedback control and
- under appropriate conditions, these feedback mechanisms can produce aperiodic, irregular fluctuations, which could easily be mistaken for noise and/or experimental error [32, 646, 647].

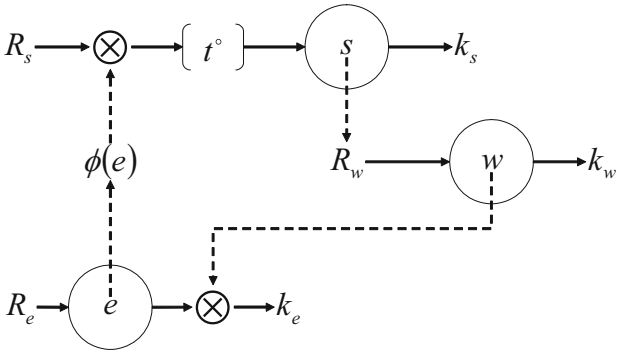


Fig. 13.10 The organization of normal hemopoiesis. Symbols are defined in the text

In the following, we will examine some theoretical developments and discuss their implications in a pharmacodynamic context. A simple, physiologically realistic mathematical model of neutrophil lineage is first proposed, including homeostatic regulation by means of cytokine G-CSF. Next, investigation of the properties of the model by stability analysis shows that this variety of clinical outcome can be described mainly from the dynamics of neutrophil counts governed by feedbacks. The sharpness of the feedback signals is essentially determined by the stability of the oscillatory behavior.

13.1.3.1 Modeling of Neutrophil Regulation

The organization of normal hemopoiesis is shown in Figure 13.10. It is generally believed that there exists a self-maintaining pluripotent stem cell population capable of producing committed stem cells specialized for the erythroid, myeloid, or thromboid cell lines [648]. The lineage studied here is the myeloid, ending with the neutrophils in the bloodstream. Three differential equations describe the mechanisms implied in this hemopoietic scheme:

1. The influx R_s of cells from the pluripotent stem cell population to the committed stem cell lines is assumed mainly regulated by long-range humoral mechanisms $\phi\,(e)$, implicating the cytokine G-CSF, $e\,(t)$. An intrinsic property of the hemopoietic chain is the presence of a time delay t° that arises because of finite cell maturation times and cell replication times for the neutrophil myelocytes, $s\,(t)$. In fact, it is important to remember that once a cell from the pluripotent stem cell population is committed to the neutrophil lineage, it undergoes a series of nuclear divisions and enters a maturational phase for a period of time ($t^\circ \approx 5 - 7\,\mathrm{d}$) before release into circulation. The production function R_s has not only to be amplified, but also to be delayed as described by the feedback component $\phi\,(e\,(t - t^\circ))$, because a change in the blood neutrophil numbers can only augment or decrease the influx into the circulation after a

period of time t° has elapsed. Thus, changes that occur at time t were actually initiated at a time $t - t^\circ$ in the past. For the sake of simplicity, we will use the notation $e^\circ(t)$ instead of $e(t - t^\circ)$. We can describe these dynamics by

$$\dot{s}(t) = R_s \phi\left(e^\circ(t)\right) - k_s s(t),$$

where k_s is the loss rate for $s(t)$. The previous equation is a differential–delay equation. In contrast to ordinary differential equations, for which we need only to specify an initial condition as a particular point in the state space for a given time, for differential–delay equations we have to specify an initial condition in the form of a function, usually called the *history function*, $\psi(t)$ and defined for a period of time equal to the duration of the time delay. Thus, we will select

$$s(t) = \psi(t), \quad -t^\circ \leq t \leq 0.$$

We will consider only initial functions that are constant, i.e., $s(t) = s_0$.

2. Mature neutrophil myelocytes $s(t)$ are now controlling the input rate R_w of neutrophils $w(t)$ that disappear from the blood with a rate constant k_w. The input function R_w in its simplest form can be considered as an amplification of $s(t)$, expressing the proliferation of neutrophil myelocytes, i.e., $R_w = \alpha s(t)$:

$$\dot{w}(t) = \alpha s(t) - k_w w(t), \quad w(0) = w_0.$$

3. At the physiological equilibrium state, cytokine G-CSF, $e(t)$, is delivered at the rate R_e and cleared by mechanisms characterized by rate constant k_e. A fall in circulating neutrophil numbers $w(t)$ leads to an acceleration of the G-CSF clearance, which has as consequence a decrease in $e(t)$ levels. This decrease in turn triggers the production of committed stem cells, which increases cellular efflux of neutrophil precursors, and ultimately augments $w(t)$ (i.e., negative feedback). This regulated behavior can be implemented by means of the $e(t)$ clearance depending on $w(t)$ levels and the $\phi(e^\circ(t))$ function [649]. The differential equation for $e(t)$ is expressed by

$$\dot{e}(t) = R_e - k_e w(t)\, e(t), \quad e(0) = e_0.$$

4. Of primary importance is the form of feedback mechanism implying the previous differential equation, where the $w(t)$ level controls the $e(t)$ clearance, and the function $\phi(e^\circ(t))$ that modulates the committed stem cell production. More specifically

$$\phi\left(e^\circ(t)\right) = \vartheta\left\{1 - \exp\left[\log\left(\frac{\vartheta - 1}{\vartheta}\right)\left(\frac{e^\circ(t)}{e_0}\right)^\theta\right]\right\},$$

a monotone increasing Weibull-like function with $\phi(e_0) = 1$. In the present model, parameters ϑ and θ express the amplification and the sharpness of the G-CSF effect.

In the above equations, s_0, w_0, and e_0 denote history and initial conditions corresponding to the undisturbed state of the process at equilibrium.

Thus, we propose to study a model with delayed feedback, as done by several investigators [48, 49, 640–642, 650–654].

Normally, the equilibrium behavior of the process requires that the production rate equal the disappearance rate. These conditions and the introduction of the variable transformation $g(t) = \alpha s(t)$ allow us to determine the input rates R_s and R_e:

$$\begin{aligned} \dot{g}(t) &= k_s k_w w_0 \phi(e^{\circ}(t)) - k_s g(t), & g(-t^{\circ} \le t \le 0) &= k_w w_0, \\ \dot{w}(t) &= g(t) - k_w w(t), & w(0) &= w_0, \\ \dot{e}(t) &= k_e [w_0 e_0 - w(t) e(t)], & e(0) &= e_0. \end{aligned} \tag{13.11}$$

Figure 13.11 illustrates the function regulating neutrophil production depending on the circulating neutrophil numbers.

We make the variables of the above equations dimensionless:

$$g(\tau) = k_w w_0 x(\tau), \quad w(\tau) = w_0 y(\tau), \quad e(\tau) = e_0 z(\tau),$$

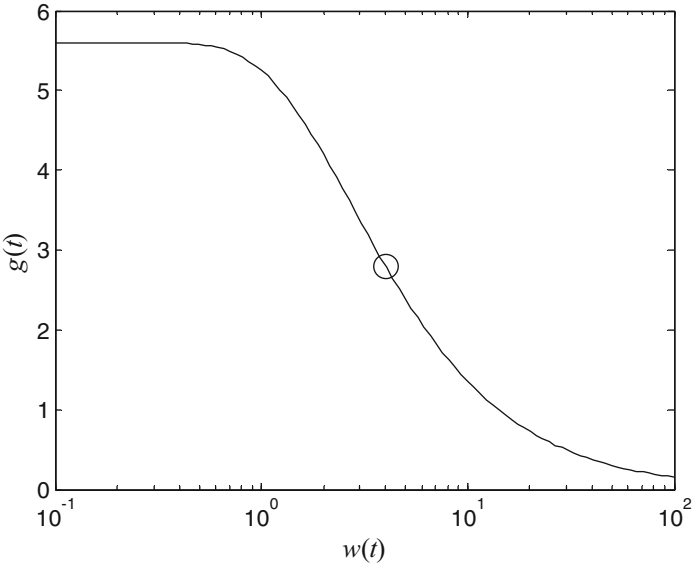


Fig. 13.11 Homeostatic control for regulation of neutrophil production. Parameters are $w_0 = 4$ and $e_0 = 1$ cells$\times 10^6$ ml^{-1}, $\vartheta = 2$ and $\theta = 1$, and $k_w = 0.7$ d^{-1}. ∘ indicates the position of the equilibrium point

with $\tau = k_e w_0 t$, and we set

$$\tau^\circ = k_e w_0 t^\circ, \;\; \gamma = k_s / (k_e w_0), \;\; \lambda = k_w / (k_e w_0).$$

The set of differential equations becomes

$$\begin{aligned} \dot{x}(\tau) &= \gamma [\phi(z^\circ(\tau)) - x(\tau)], \;\; x(-\tau^\circ \le \tau \le 0) = 1, \\ \dot{y}(\tau) &= \lambda [x(\tau) - y(\tau)], \qquad y(0) = 1, \\ \dot{z}(\tau) &= 1 - y(\tau) z(\tau), \qquad z(0) = 1, \end{aligned} \tag{13.12}$$

with $z^\circ(\tau) = z(\tau - \tau^\circ)$ and

$$\phi\left(z^\circ(\tau)\right) = \vartheta \left\{ 1 - \exp\left[\log\left(\frac{\vartheta - 1}{\vartheta} \right) \left(z^\circ(\tau)\right)^\theta \right] \right\}. \tag{13.13}$$

Current analytic and numerical work determine the time-dependent changes in blood cell number as certain quantities, referred to as control parameters, are varied. Examples of control parameters in the regulation of hemopoiesis are the dimensionless maturation time τ° and the peripheral destruction rates γ and λ.

It is well established that under appropriate circumstances, delayed negative feedback mechanisms can produce oscillations. To illustrate this point we continue with the stability analysis.

13.1.3.2 Stability Analysis

Equilibrium points (x^*, y^*, z^*) of the system are those for which $\dot{x}(\tau) = \dot{y}(\tau) = \dot{z}(\tau) = 0$. As previously defined, the equilibrium state of the process leads to the single equilibrium point $(x^*, y^*, z^*) = (1, 1, 1)$. Since $z(\tau)$ is not changing with time, we have also $(z^\circ)^* = z^* = 1$. We would now like to know what conditions on the parameters of our model are required to warrant stability, and even further, what happens in the case of instability.

Because the model (13.12) that describes this physiological process is nonlinear, we cannot answer these questions in total generality. Rather, we must be content with understanding what happens when we make a small perturbation on the states x, y, and z away from the equilibrium. The fact that we are assuming that the perturbation is small allows us to carry out what is known as linear stability analysis of the equilibrium state.

The nonlinearity of (13.12) comes from the terms $\phi(z^\circ(\tau))$ and $y(\tau) z(\tau)$ involved in the nonlinear negative feedback regulation. What we want to do is replace these nonlinear terms by a linear function in the vicinity of the equilibrium state (x^*, y^*, z^*). This involves writing the Jacobian matrix of the linearized system (cf. Appendix A):

$$\mathbf{A} = \begin{bmatrix} -\gamma & 0 & \gamma\phi'(1)\,(\mathrm{d}z^\circ/\mathrm{d}z) \\ \lambda & -\lambda & 0 \\ 0 & -1 & -1 \end{bmatrix}.$$

To analyze stability of the linearized model, we have to examine the eigenvalues that are solutions of the characteristic equation of $\mathbf{A}$. Usually the eigenvalue is a complex number $\zeta = \mu + \mathrm{i}\omega$. If $\mu = \mathrm{Re}\,\zeta < 0$, then the solution is a decaying oscillating function of time, so we have a stable situation. If $\mu = \mathrm{Re}\,\zeta > 0$ on the other hand, then the solution diverges in an oscillatory fashion and the solution is unstable. The boundary between these two situations, where $\mu = \mathrm{Re}\,\zeta = 0$, defines a Hopf bifurcation in which an eigenvalue crosses from the left-hand to the right-hand complex plane.

The usual procedure to obtain solutions of the characteristic equation of $\mathbf{A}$ is to assume that the solution of $z(\tau)$ has the form

$$z(\tau) \propto \exp(\zeta\tau)$$

and find out the requirements on the parameters of the equation so that there is an eigenvalue ζ allowing $z(\tau)$ to be written in this form. Under this assumption,

$$\mathrm{d}z^\circ/\mathrm{d}z = \exp\left(-\zeta\tau^\circ\right),$$

and the eigenvalues are given as solutions of the characteristic equation

$$(\zeta+\gamma)(\zeta+\lambda)(\zeta+1) + \gamma\lambda\phi'(1)\exp\left(-\zeta\tau^\circ\right) = 0. \tag{13.14}$$

In contrast to systems without delay, the previous equation has generally an infinite number of roots. Nevertheless, there are only a finite number of roots with real parts [655]. Figure 13.12 illustrates solutions of (13.14) with $\gamma = 0.025$, $\lambda = 0.175$, $\phi'(1) = 2$, and $\tau^\circ = 0, 10, 100$. We note that:

- all roots are complex conjugates, except for $\tau^\circ = 0$, where one root is real,
- only for $\tau^\circ = 100$ do we have one pair of roots with positive real part, $\mu = \mathrm{Re}\,\zeta \approx 0.002947$,
- in the complex plane, the roots' density near the origin is higher for $\tau^\circ = 100$ than the density for the other τ° values.

Let us find the critical delay value $\tau^\bullet$ at which the characteristic roots intersect the stability boundary, i.e., the imaginary axis $\mu = 0$, thus rendering the system unstable with $\zeta = \mathrm{i}\omega^\bullet$. We substitute this into the previous equation, and after separating real and imaginary parts, we have

$$\begin{aligned} \gamma\lambda\left[1+\phi'(1)\cos\omega^\bullet\tau^\bullet\right] &= (\gamma+\lambda+1)\,\omega^{\bullet 2}, \\ \gamma\lambda\phi'(1)\sin\omega^\bullet\tau^\bullet &= (\gamma\lambda+\gamma+\lambda)\,\omega^\bullet - \omega^{\bullet 3}. \end{aligned} \tag{13.15}$$

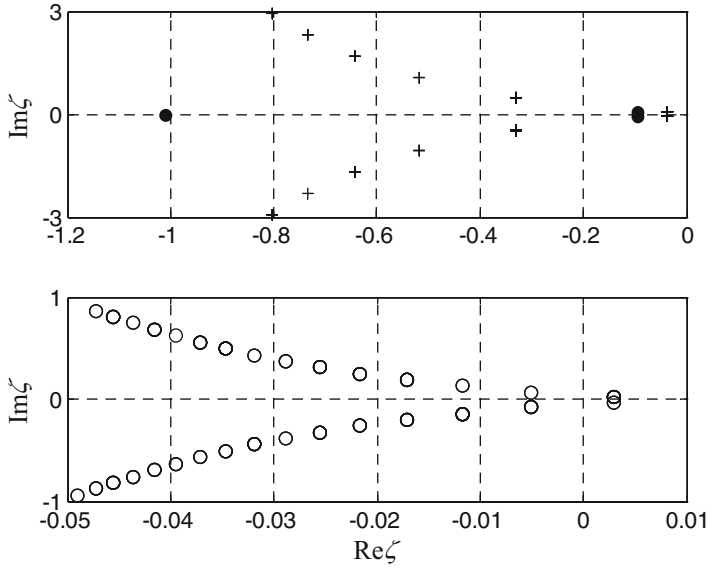


Fig. 13.12 Characteristic roots for different τ° values: $\tau^\circ = 0$ (•), $\tau^\circ = 10$ (+), and below for $\tau^\circ = 100$ (∘)

Note that if $i\omega^\bullet$ is a characteristic root of (13.14), then $-i\omega^\bullet$ is also a characteristic root. Then, we can assume that $\omega^\bullet > 0$. Squaring and adding the above two equations defines a polynomial equation in $\omega^{\bullet 6}$ with only even powers. If we set $\chi = \omega^{\bullet 2} > 0$, this equation becomes a third-order polynomial equation in χ:

$$\chi^3 + \beta_2\chi^2 + \beta_1\chi + \beta_0 = 0$$

with

$$\begin{aligned}\beta_2 &= \gamma^2 + \lambda^2 + \gamma + \lambda + 1,\\ \beta_1 &= \gamma^2\lambda^2 + \gamma^2 + \lambda^2,\\ \beta_0 &= (\gamma\lambda)^2\left[1 - \phi'(1)^2\right].\end{aligned}$$

According to the Descartes rule of signs and since β_1 and β_2 are positive, the inequality $\beta_0 < 0$ or $\phi'(1) > 1$ is a necessary condition to have the unique positive solution $\chi > 0$.

After evaluating this solution, we obtain $\omega^\bullet$, and then, from one of (13.15), we calculate the critical value of $\tau^\bullet$. Figure 13.13 shows the frequency $\omega^\bullet$ (upper panel) and the critical value of $\tau^\bullet$ (lower panel) as functions of $\phi'(1)$. We can say that the real parts of ζ will be positive, and thus (13.12) will be unstable, if and only if the actual delay τ° is greater than $\tau^\bullet$. For example, for $\phi'(1) = 2$ and the set γ and λ values, $\tau^\bullet \approx 44.049$ and any $\tau^\circ > \tau^\bullet$ triggers periodic oscillations following some perturbation in the system. Otherwise, the system is locally stable.

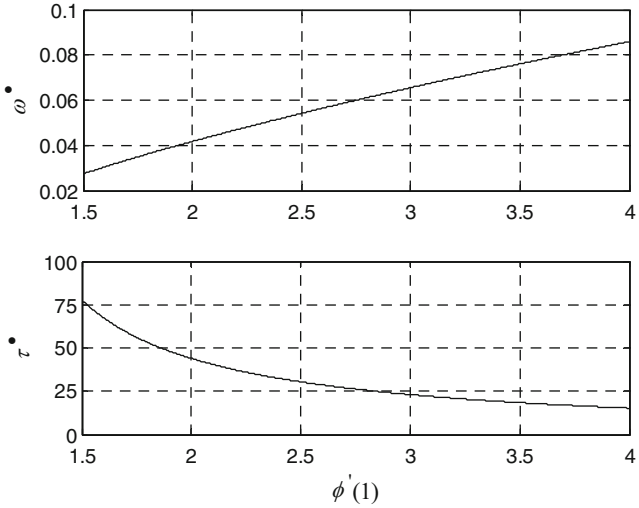


Fig. 13.13 Critical frequency $\omega^\bullet$ and delay $\tau^\bullet$ as functions of $\phi'(1)$

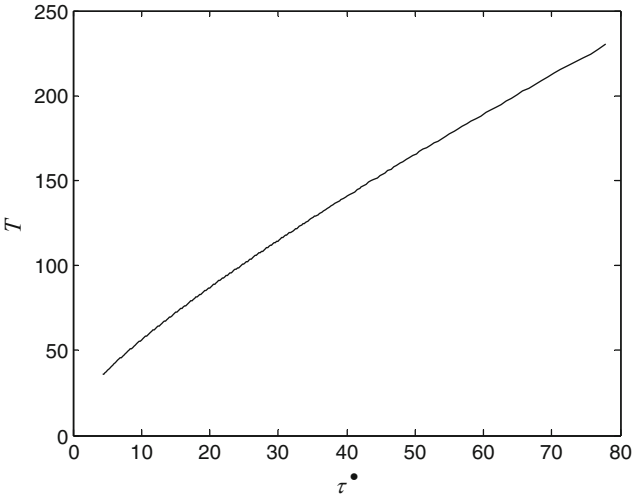


Fig. 13.14 The period T of oscillations as a function of $\tau^\bullet$

The period T of the periodic solution can be obtained by noting that $\omega = 2\pi/T$. In general, the period of an oscillation produced by a delayed negative feedback mechanism is at least twice the delay [640]. For our model of neutrophil production, the functional relationship between $\tau^\bullet$ and T is shown in Figure 13.14. We note that the period of the oscillation should be four times the delay as reported by Mackey [641] and as also concluded by another simpler model without the neutrophil

myelocytes $s(t)$ (results not shown). Since the maturational delay for neutrophil production is $\tau \approx 5-7$ d, we would expect to see oscillations in neutrophil numbers with periods of about $3-4$ weeks.

13.1.3.3 Chemotherapy

Neutropenic episodes place patients at increased risk for infective processes (e.g., abscesses, pneumonia, septicemia), and up to 20% of patients may die during these episodes. For example, in patients undergoing cancer chemotherapy, neutropenia is frequently a dose-limiting side effect. The importance of the time profile of hematologic effects in analyzing properties of anticancer agents has been recently recognized [656, 657]. Different models of the entire time course of responses have been proposed. They can be classified as either mechanistic or empirical. The latter models postulate explicit relationships between the effect and pharmacodynamic and pharmacokinetic parameters [597, 656], whereas the mechanistic models describe the biological processes controlling the change of the affected cells [657, 658].

Here, a new model of hematologic toxicity of anticancer agents is introduced. The postulated mechanisms that influence the response variable (e.g., neutrophil count) are:

- an indirect mechanism, where by means of a logistic function, the cell production rate of neutrophil myelocytes $s(t)$ is modulated by the blood concentrations of the anticancer drug and
- direct toxicity of the anticancer drug levels, according to which the killing rate of neutrophil myelocytes is proportional to $s(t) \times$ [drug levels] [659].

This model was identified from data gathered in a clinical study [660] aiming to define a regular and tolerable dose of the epirubicin–docetaxel combination in first-line chemotherapy on 65 patients with metastatic breast cancer. Following the analysis of these data, parameters were set to

$$\begin{array}{lll} k_w = 0.7\,\mathrm{d}^{-1}, & k_e = 1\,\mathrm{d}^{-1}, & k_s = 0.1\,\mathrm{d}^{-1}, \\ w_0 = 4000\ \text{cells} \times \mathrm{mm}^{-3}, & e_0 = 1000\ \text{cells} \times \mathrm{mm}^{-3}, & \\ \vartheta = 2, & \theta = 2.8854, & \end{array}$$

leading to the dimensionless parameters

$$\gamma = 0.025,\ \lambda = 0.175,\ \phi'(1) = 2.$$

Two different delays were also assessed, $t^\circ = 5$ and 15 d, corresponding to $\tau^\circ = 20$ and 60, respectively. The simulation of the neutrophil count is presented in the Figure 13.15. We observe that:

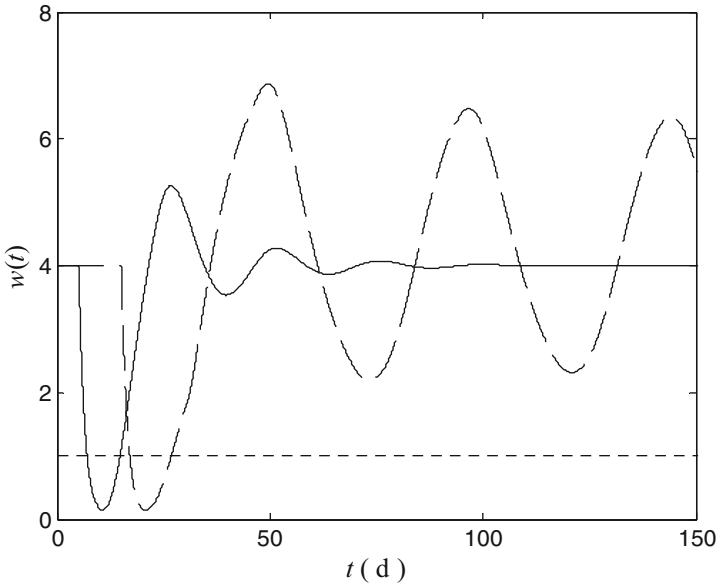


Fig. 13.15 Simulation of the neutrophil count kinetics for $t^\circ = 5$ (*solid line*) and 15 d (*dashed line*). The *dotted line* indicates the minimum allowed neutrophil level

- while in the first month following initial delay the two kinetic patterns look the same, their behavior has been differentiated for the subsequent time leading either to oscillatory or dampening behavior corresponding to $t^\circ = 15$ and $t^\circ = 5$ d delays, respectively, and
- the periods of oscillations in both cases are 4-fold higher than the initial delay.

The clinical significance of the previous analysis is that it may be possible to develop new therapeutic strategies with agents shortening the period of chemotherapy-associated neutropenia such as lenograstim [661]. Such agents may reduce incidence or duration of serious infections and enable greater dose-intensification. In the long run, quantitative modeling may support the design of chemotherapy or growth factor drug regimens based on manipulation of feedback [32, 48–50]. Alternatively, the model can be used to identify more specifically the effects of drugs in the hemopoietic system.

13.1.3.4 Mixed Feedback

Recently, Bernard et al. [654] studied oscillations in cyclical neutropenia, a rare disorder characterized by oscillatory production of blood cells. As above, they developed a physiologically realistic model including a second homeostatic control on the production of the committed stem cells that undergo apoptosis at their

proliferative phase. By using the same approach, they found a local supercritical Hopf bifurcation and a saddle-node bifurcation of limit cycles as critical parameters (i.e., the amplification parameter) are varied. Numerical simulations are consistent with experimental data and they indicate that regulated apoptosis may be a powerful control mechanism for the production of blood cells. The loss of control over apoptosis can have significant negative effects on the dynamic properties of hemopoiesis.

In the previous analysis, delayed negative feedback mechanisms were considered only for neutrophil regulation. However, if over a wide range of circulating neutrophil levels, the neutrophil production rate decreases as the number of neutrophils increases (i.e., negative feedback), in the range of low neutrophil numbers the production rate must increase as neutrophil number increases (i.e., positive feedback). This type of feedback was reported as *mixed feedback* [627].

In order to contrast the dynamics that arise in delayed negative and mixed feedback mechanisms, Mackey et al. [662] considered periodic chronic myelogenous leukemia in which peripheral neutrophil numbers oscillate around elevated levels with a period of $30 - 70$ d even in the absence of clinical interventions [663, 664]. On closer inspection it can be seen that the number of days between successive maximum numbers of neutrophils is not constant, but varies by a few days. Moreover, the morphology of each waveform differs slightly and there are shoulders on some of them. Mackey et al. [48, 49] have explored the possibility that these irregularities are intrinsic properties of the underlying control mechanism. These studies indicate that the dynamics of mixed feedback are much richer than for the simple negative feedback model. Increases in t° are of particular interest since a prolongation of the neutrophil maturation time is inferred in patients with chronic myelogenous leukemia [665]. As t° is increased an initially stable equilibrium becomes unstable and stable periodic solutions appear. Further increases in t° lead to a sequence of period-doubling bifurcations, which ultimately culminates in an apparently chaotic or aperiodic regime. Here, the model predicts that levels of circulating neutrophils are random simply as a consequence of their own deterministic evolution.

The observations in these notes emphasize that an intact control mechanism for the regulation of blood cell numbers is capable of producing behaviors ranging from no oscillation to periodic oscillations to more complex irregular fluctuations, i.e., chaos. The type of behavior produced depends on the nature of the feedback, i.e., negative or mixed, and on the value of certain underlying control parameters , e.g., peripheral destruction rates γ and λ or maturation times τ°. Pathological alterations in these parameters can lead to periodic hematologic disorders. The observation that periodic hematologic diseases have periods that are multiples of 7 may simply be a consequence of the combination of delayed feedback mechanisms with maturation times that are on the order of $5 - 7$ d. Thus it is not necessary to search for elusive and mystical entities [666], such as circadian rhythms, to explain the periodicity of these disorders.

The realization that physiological control mechanisms can generate exceedingly complex oscillations, such as chaos, is a subject of great interest [32, 48, 49, 646,

647]. It is quite possible that both interesting and relevant dynamic changes are often observed, but their significance is wrongly ascribed to environmental noise and/or experimental error. Careful attention to these dynamic behaviors may eventually provide important insights into the properties of the underlying control mechanisms.

13.1.3.5 Periodic and Dynamical Diseases

The first explicit description of the concept of periodic diseases, where the disease process itself may flare or recur on a regular basis of days to months, was provided over 40 years ago by H. Reimann [667]. That author described and catalogued a number of periodic disease states ranging from certain forms of arthritis to some mental illnesses and hereditary diseases such as familial Mediterranean fever. As an extension to the concept of periodic diseases introduced by Reimann and to encompass irregular physiologic dynamics thought possibly to represent deterministic chaos, the term *dynamical disease* has been introduced [32, 48–50]. A dynamical disease is defined as a disease that occurs in an otherwise intact physiological control system but operates within a range of control parameters that leads to abnormal dynamics. Clearly the hope is that it may eventually be possible to identify these altered parameters and then readjust them to values associated with healthy behaviors.

13.2 Pharmacodynamic Applications

During the last fifteen years many investigators have expanded traditional pharmacodynamic modeling (law of mass action at equilibrium) to mechanistic pharmacodynamic modeling including detailed modeling of the underlying physiology and then modeling the effect of drugs on it. On the other hand, as just pointed out, deterministic chaos is typically the recorded behavior of complex physiological systems implicating feedback regulations and nonlinear elements. In the next paragraphs, three major fields of physiological systems with great importance in pharmacotherapy, namely cardiovascular, central nervous, and endocrine systems, where tools and concepts from nonlinear dynamics have been applied, will be discussed.

13.2.1 Drugs Affecting Endocrine Function

It is widely appreciated that hormone secretion is characterized by pulsatility. The first experimental studies of the pulsatile nature of hormone secretion started more than thirty years ago. Hellman et al. reported in 1970 [668] that “cortisol is secreted episodically by normal man.” It was also realized that this pulsatility

was not due to noise, but was actually associated with physiological processes. Indeed, the circadian clock, the interaction between hormones through feedback mechanisms, and the interaction of hormones with central and autonomic nervous systems are some of the reasons for this behavior. It has been apparent that the theory of dynamic systems is the right field in which to find useful tools for the study of hormonal systems. This has been done along two directions:

- experimental studies using tools from time series analysis and
- modeling with differential equations.

13.2.1.1 A Dynamic System for Cortisol Kinetics

Although the detailed features of the interactions involved in cortisol secretion are still unknown, some observations indicate that the irregular behavior of cortisol levels originates from the underlying dynamics of the hypothalamic–pituitary–adrenal process. Indeed, Ilias et al. [669], using time series analysis, have shown that the reconstructed phase space of cortisol concentrations of healthy individuals has an attractor of fractal dimension $d_f = 2.65 \pm 0.03$. This value indicates that at least three state variables control cortisol secretion [670]. A nonlinear model of cortisol secretion with three state variables that take into account the simultaneous changes of adrenocorticotropic hormone and corticotropin-releasing hormone has been proposed [671].

The Model These observations prompted us to model cortisol plasma levels [672] relying on the well-established erratic secretion rate [673] and the circadian rhythm, while other factors controlling cortisol secretion are also considered but not expressed explicitly:

- Cortisol concentration is described by a nonlinear time-delay differential equation [48, 674] with two terms, i.e., a secretion rate term that adheres to the negative feedback mechanism [675, 676] and drives the pulsatile secretion, and a first-order output term with rate constant k_o:

$$\dot{c}\,(t) = k_i \frac{g^{\gamma} c^{\circ}\,(t)}{g^{\gamma} + [c^{\circ}\,(t)]^{\gamma}} - k_o c\,(t)\,, \tag{13.16}$$

  where $c\,(t)$ is the cortisol concentration, $c^{\circ}\,(t)$ is the value of $c\,(t)$ at time $t - t^{\circ}$, γ is an exponent, and k_i and k_o are the input and output rate constants, respectively.
- The circadian rhythm of cortisol secretion is implemented phenomenologically by considering the parameter of the model as a simple cosine function of 24-h period:

$$g\,(t) = \alpha \cos\left[(t - \varphi)\,\frac{2\pi}{1440}\right] + \beta, \tag{13.17}$$

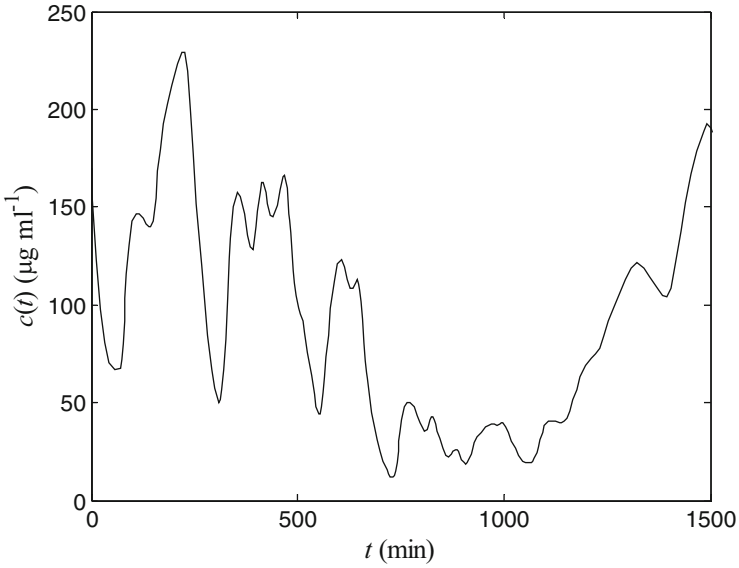


Fig. 13.16 A 24-h simulated profile generated by the model of cortisol kinetics

where α and β are constants with concentration units, φ is a constant with time units, and t is time in minutes. Similar approaches relying on simple periodic functions were used by Rohatagi et al. [677] to describe the secretion rate of cortisol.

Our dynamic model consists of (13.16) and (13.17). The physical meaning of the time delay in (13.16) is that the cortisol concentration $c(t)$ affects other physiological parameters of the hypothalamic–pituitary–adrenal process (not present in equation 13.16), which in turn affect, via the feedback mechanism, cortisol concentration; thereby, cortisol controls its own secretion [605]. This cycle is postulated to last time t°, and that is how the concentration $c^\circ(t)$ at time $t - t^\circ$ arises.

The simulated profile generated by (13.16) and (13.17) is shown in Figure 13.16. Model parameters take the values

$$k_i = 0.0666\,\text{min}^{-1},\ k_o = 0.0333\,\text{min}^{-1},\ c(0) = 170\,\mu\text{g ml}^{-1},\ \gamma = 10,$$
$$\alpha = 70\,\mu\text{g ml}^{-1},\quad \beta = 100\,\mu\text{g ml}^{-1},\quad \varphi = 250\,\text{min},\qquad t^\circ = 70\,\text{min}.$$

The value assigned to t° corresponds to about one cortisol secretion burst per hour in accordance with experimental observations [673]. The simulations were performed by a numerical solution of (13.16) and (13.17). This simulation exhibits the circadian rhythm, as well as the pulsatile nature of the cortisol secretion system.

Since (13.16) has an infinite number of degrees of freedom [678], we constructed a pseudophase space [4, 33] for the system of (13.16) and (13.17) using the model

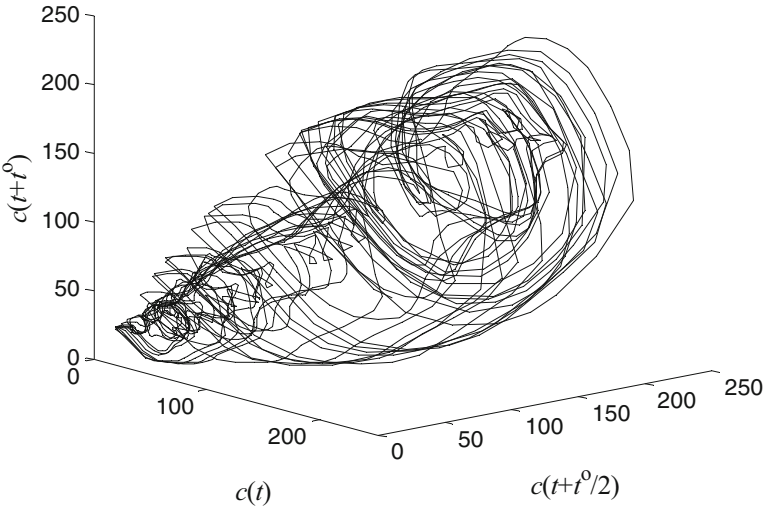


Fig. 13.17 A pseudophase space for the model of cortisol kinetics. Variables $c\,(t)$, $c\,(t + t^{\circ}/2)$, $c\,(t + t^{\circ})$ are expressed in $\mathrm{g\,ml^{-1}}$

variables $c\,(t)$, $c\,(t + t^{\circ}/2)$, $c\,(t + t^{\circ})$, Figure 13.17. The use of three dimensions is in accordance with the embedding dimension that Ilias et al. [669] have found. The attractor of our system is quite complicated geometrically, i.e., it is a strange attractor. The real phase space is of infinite dimension. However, trajectories may be considered to lie in a low-dimensional space (attractor). The model parameters take the same values as in Figure 13.16 and time runs for 10 days.

A Dynamic Perspective of Variability The model under study here offers an opportunity to refer to some implications of the existence of nonlinear dynamics. Apart from the jagged cortisol concentration profile, elements such as the sensitive dependence on the initial conditions (expressed by the positive Lyapunov exponent), as well as the system's parameters, play an important role and may explain the inter- and intraindividual variability observed in the secretion of cortisol. These implications, together with other features absent from classical models, are demonstrated in Figure 13.18.

In all plots the dashed line is generated from (13.16) and (13.17) using the above parameter values, while the sampling interval is fixed to 30 min. The solid lines correspond to the same set of parameter values applying a change only in one of them. This change, however, is enough to produce significant visual change in the profile: (A): k_o is set to $0.03\,\mathrm{min^{-1}}$; (B): $c\,(0)$ is set to $160\,\mu\mathrm{g\,ml^{-1}}$; (C): the second-day profile is compared to the first-day profile; (D): sampling is performed every 80 min instead of 30 min. The dashed and solid lines of plots C and D have identical values for the model parameters. Thus, a change in the initial conditions or the parameter values of (13.16) and (13.17) may be depicted in a relatively large change of the final profile, Figure 13.18A and B. Also, the profiles corresponding to two

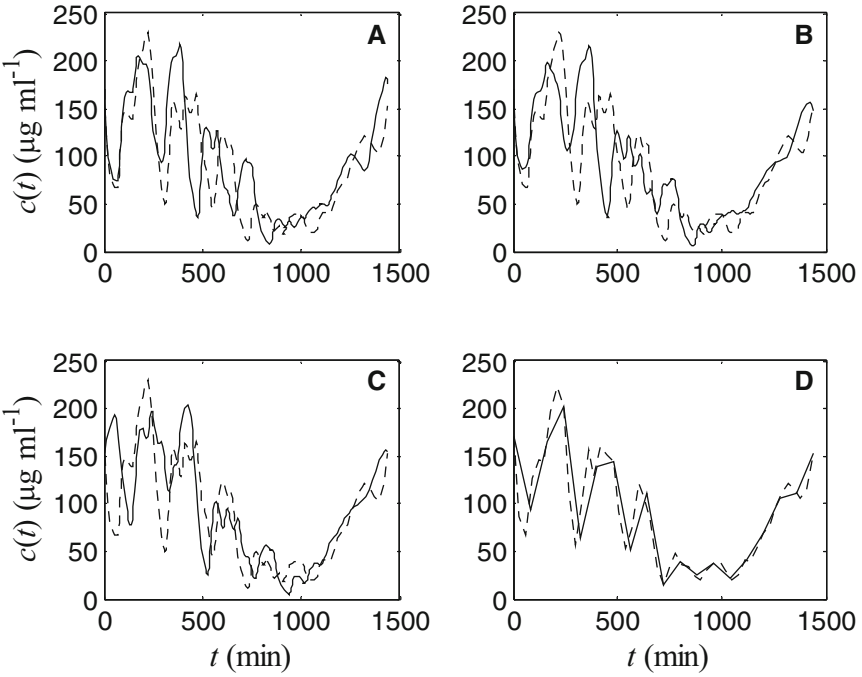


Fig. 13.18 The *dotted lines* are generated from the model of cortisol kinetics using the same parameter values as for Fig. 13.16. The *solid lines* correspond to the same set of parameter values applying a change only in one of them: k_o (**A**), $c\,(0)$ (**B**), and observation sampling (**D**). In (**C**), the second-day profile is compared to the first-day profile

successive days (Figure 13.18C), or two different sampling designs (Figure 13.18D), may differ remarkably, even though the exact same set of parameter values is used. Overall, our analysis based on nonlinear dynamics offers an alternative explanation for the fluctuation of cortisol levels. However, the most important implication of the presence of nonlinear dynamics in cortisol secretion processes is the limitation for long-term prediction, which makes practical application of the classical models questionable.

As we have already mentioned (cf. Chapter 3), one of the most important features of nonlinear dynamics is the sensitivity to initial conditions. A measure to verify the chaotic nature of a dynamic system is the Lyapunov exponent [33], which quantifies the sensitive dependence on initial conditions. In the present model we found [679] the largest Lyapunov exponent to have a positive value of around $0.00011\,\mathrm{min}^{-1}$, which is a clear indication for chaotic behavior.

Cortisol Suppression by Corticosteroids The model presented here allows the consideration of external corticosteroid administration as a perturbation of the cortisol secretion system. As a matter of fact, corticosteroids cause a temporary diminution of plasma cortisol levels [677]. Assuming that the drug follows one-

compartment model disposition with first-order input and output, the effect-site [680] concentration is described by the following equation [584]:

$$y(t) = \frac{F_a q_0}{V} \frac{k_a k_y}{k_a - k_e} \left[\frac{\exp\left(-k_y t\right) - \exp\left(-k_e t\right)}{k_e - k_y} - \frac{\exp\left(-k_y t\right) - \exp\left(-k_a t\right)}{k_a - k_y} \right], \tag{13.18}$$

where F_a is the bioavailable fraction of dose q_0, V is the volume of distribution of the pharmacokinetic compartment, k_a, k_e are the input and elimination first-order rate constants from the pharmacokinetic compartment, respectively, and k_y is the elimination rate constant from the effect compartment.

The effect-site concentration of the corticosteroids can be considered to affect one or more parameters of the model described by (13.16). This must be implemented so that the presence of $y(t)$ suppresses the cortisol secretion in accordance with the experimental data. Instead of $g(t)$, the parameter describing the circadian rhythm, a new parameter $\tilde{g}(t)$ was introduced to include the effect of corticosteroid administration following a receptor reduction:

$$\tilde{g}(t) = g(t) \left[1 - \frac{y(t)}{Ec_{50} + y(t)} \right], \tag{13.19}$$

where Ec_{50} is a coefficient that expresses the concentration of the drug when $\tilde{g}(t) = g(t)/2$. In this simple way, and in the presence of external corticosteroid drug administration, realistic cortisol blood levels can be obtained as illustrated in Figure 13.19, for the case of fluticasone propionate [605].

The solid and dashed lines represent simulations for two "individuals" with significantly different profiles corresponding to different initial conditions, $c(0) = 90\,\mu\mathrm{g\,ml^{-1}}$ (solid) and $c(0) = 150\,\mu\mathrm{g\,ml^{-1}}$ (dashed). Parameter values were set to

$$\begin{array}{lll} k_i = 0.0666\,\mathrm{min^{-1}}, & k_o = 0.0333\,\mathrm{min^{-1}}, & \gamma = 10, \\ \alpha = 90\,\mu\mathrm{g\,ml^{-1}}, & \beta = 100\,\mu\mathrm{g\,ml^{-1}}, & \varphi = 200\,\mathrm{min}, \qquad t^{\circ} = 70\,\mathrm{min}. \\ k_a = 0.14\,\mathrm{min^{-1}}, & k_e = 0.002\,\mathrm{min^{-1}}, & k_y = 0.005\,\mathrm{min^{-1}}, \\ V = 22.2\,\mathrm{l}, & F_a q_0 = 1\,\mathrm{mg}, & Ec_{50} = 20\,\mu\mathrm{g\,ml^{-1}}, \end{array}$$

These values were selected in order to generate qualitatively similar profiles to the experimental data and were not optimized since fitting is not well established for chaotic systems. In parallel, the sensitive dependence of the detailed final profile from the exact values of the concentration $y(t)$ should be emphasized, since $y(t)$ directly affects one of the parameters of the chaotic oscillator (13.19).

Finally, experimental evidence indicates that fluctuations in cortisol secretion are not produced by random processes. In fact, the large inter- and intraindividual variability observed in studies dealing with the effect of fluticasone propionate on cortisol levels [681] may be partly explained with the erratic behavior of the system of (13.16) to (13.19).

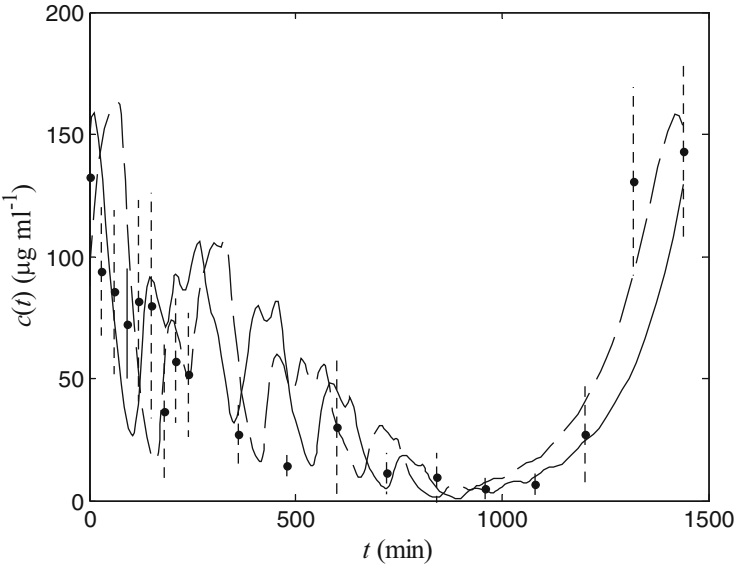


Fig. 13.19 Diminution of cortisol blood levels in the presence of fluticasone propionate. *Circles* represent averaged experimental data of four volunteers after the administration of 1 mg of inhaled drug [605], while the *solid* and *dashed lines*, generated by the model of cortisol kinetics, represent simulated data for two "individuals" with different initial conditions

13.2.1.2 Parametric Models

Numerous other experimental studies of hormonal systems utilize tools from nonlinear dynamic systems theory. Smith in 1980 [682] used a mathematical model of three interacting hormones, namely testosterone, luteinizing hormone, and luteinizing hormone-releasing hormone, to describe qualitatively their behavior. The initial model was improved later by Cartwright and Husain [683], introducing time-retarded terms of the three state variables to make the system more realistic, exhibiting limit cycle solutions. Further improvements of the model were studied by Liu and Deng [684] and also by Das et al. [685]. Apart from testosterone other efforts in the same context have been made to model the secretion of hormones. Examples include the work of Lenbury and Pacheenburawana [671] in the system of cortisol, adrenocorticotropic hormone, and corticotrophin-releasing hormone, the work of Topp et al. in the system of β-cell mass, insulin, and glucose [686], and also the work of Londergan and Peacock-Lopez [687]. The latter is a general model of hormone interaction description with negative feedback, exhibiting very rich dynamics and even chaotic behavior.

Many drugs affect normal hormonal secretion, either as their primary target of action or as a side effect. Many studies in recent years have considered models of hormonal secretion together with the dominant pharmacokinetic-dynamic concepts of drug action. Examples include the effect of corticosteroids on cortisol

by Chakraborty et al. [605]; the effect of the gonadotropin-releasing hormone antagonist on testosterone and luteinizing hormone by Fattinger et al. [688]; the effect of the dopaminomimetic drug DCN 203-922 on prolactin by Francheteau et al. [689]; the effect of the calcimimetic agent R-568 on parathyroid hormone by Lalonde et al. [690]; and the effect of ipamorelin on growth hormone by Gobburu et al. [691].

All the above studies share a common element. The hormone secretion modeling is kept to a minimum, usually consisting of a single differential equation or even an algebraic equation that gives a simple smooth hormone baseline. Then, the pharmacokinetic-dynamic models, such as direct or indirect link and response [572], relate the inhibition or stimulation of the baseline with the drug concentration. In order to set the baseline, only the most obvious characteristics of the hormone profile are integrated, as a periodic circadian rhythm. The dynamic structure of the underlying physiology is practically ignored and so is pulsatility, which is considered to be noise. The only studies in which pulsatility is considered as a feature of the profile are the works of Francheteau et al. [689] for the effect of dopaminomimetic drug DCN 203-922 on prolactin and Chakraborty et al. [605] for the effect of fluticasone propionate on cortisol. However, even in these studies the pulsatility is integrated phenomenologically through spline terms or Fourier harmonics, respectively, and not through modeling of the dynamic origin of the pulsatility. It must be noted though that there are studies in which the pulsatility does not play an important role, like the study of Gobburu et al. [691] for the effect of ipamorelin on growth hormone, where the baseline of the hormone is reasonably considered zero due to the multifold amplification of the growth hormone levels after the administration of the drug.

A mathematical model of the insulin–glucose feedback regulation in man was proposed by Tolic et al. [692] to examine the effects of an oscillatory supply of insulin compared to a constant supply at the same average rate. The model analysis allowed them to interpret seemingly conflicting results of clinical studies in terms of their different experimental conditions with respect to hepatic glucose release. If this release operates near an upper limit, an oscillatory insulin supply will be more efficient in lowering the blood glucose level than a constant supply. If the insulin level is high enough for the hepatic release of glucose to nearly vanish, the opposite effect is observed. For insulin concentrations close to the point of inflection of the insulin–glucose dose–response curve, oscillatory and constant insulin infusion produce similar effects.

13.2.1.3 Nonparametric Models

The phase space reconstruction approach, making use only of the hormone plasma profiles, was utilized in order to assess dimensionality and thus expose the chaotic nature of the underlying dynamics of various hormones. In all these studies, reconstruction of the phase space gave attractors of fractal dimension, evidence for

the presence of nonlinear dynamics. Such examples are the work of Prank et al. [693] on parathyroid hormone, Ilias et al. [669] on cortisol and growth hormone, and Papavasiliou et al. [694] on prolactin.

By using nonlinear dynamics methods, Papavasiliou et al. [694] analyzed the circadian profiles of prolactin, directly from the experimental data, by combining in a single time series (432 measurements), six individual 24-h prolactin profiles (72 measurements per profile, 20 min sampling interval), obtained from young healthy human volunteers, under basal conditions. These authors found that significant autocorrelation exists between any given point of the time series and a limited number of its successors [694]. Fourier analysis showed a dominant frequency of 1 cycle$\times$ d^{-1}, without sub-24-h harmonics. Poincaré section indicated the presence of a fractal attractor, and a sketch of the attractor revealed a highly convoluted geometric structure with a conical contour. The box-counting dimension was found to be fractional, namely $d_f = 1.66$, indicating that diurnal prolactin secretion is governed by nonlinear dynamics. *Information dimension* and *correlation dimension* confirmed the above value of the attractor. The two dimensions did not differ significantly from each other, and exhibited saturation at an embedding dimension of 2. The evidence taken together suggests that under basal conditions, the daily changes in the peripheral blood levels of prolactin are governed by nonlinear deterministic dynamics, with a dominant rhythm of 1 cycle $\times$ d^{-1} mixed with a higher-frequency, low-amplitude signal.

Pincus developed in 1991 a different method to quantify hormone pulsatility, which is referred to as the approximate entropy algorithm [695] and is based on the concept of Lyapunov exponents. This method has been applied for several hormones such as adrenocorticotropic hormone, cortisol, prolactin, insulin, growth hormone, testosterone, and luteinizing hormone, quantifying the observed pulsatility and comparing it between different groups such as sick vs. healthy, different age groups, etc. ([696] and references therein). The experimental evidence of the chaotic nature of hormonal underlying dynamics clarifies the origin of pulsatility and acts as a guide for proper modeling.

Serial data of glucose and insulin values of individual patients vary over short periods of time, since they are subject to biological variations. The classic homeostatic control model assumes that the physiological mechanisms maintaining the concentrations of glucose and insulin are linear. The only deviations over a short period of time one should observe are in relation to glucose load or major hormonal disturbance. Otherwise, the values of glucose and insulin should be constant and any variations should be due to random disturbances. Kroll [697] investigated previously published serial data (three for glucose and one for insulin) with both linear and nonlinear techniques to evaluate the presence of deterministic components hidden within the biological (intraindividual) variation. Within the linear techniques, the power spectra failed to show dominant frequencies, but the autocorrelation functions showed significant correlation, consistent with a deterministic process. Within the nonlinear techniques, the correlation dimension was finite, around 4.0, and the first Lyapunov exponent was positive, indicative of a deterministic chaotic process. Furthermore, the phase portraits showed directional flow. Therefore, the short-term

biological variation observed for glucose and insulin records arises from nonlinear, deterministic chaotic behavior instead of random variation.

From the above studies, it is evident that although significant progress has been made as far as the physiological modeling of hormonal systems is concerned, the relevant pharmacodynamic modeling, even in state-of-the-art studies dealing with the effect of drugs on hormonal levels, practically ignores these findings. It is a necessity to develop new pharmacodynamic models for drugs related to hormonal secretion, compatible with the physiological modeling and the experimental findings that suggest low-dimensional nonlinear dynamic behavior. This kind of modeling is not only more realistic: it integrates a new rationale as well. The notions of sensitivity with respect to the initial conditions and qualitatively different behavior for different, even slightly, values of the control parameters surely play an important role and must be taken into account in modeling since their presence is suggested by experiments.

An important outcome of these studies is the opportunity that it offers to discuss the implications of the presence of nonlinear dynamics in processes such as the secretion of cortisol. Based on the aforementioned discussion it is evident that the concepts of deterministic nonlinear dynamics should be adopted in pharmacodynamic modeling when supported by experimental and physiologic data. This is valid not only for the sake of more detailed study, but mainly because nonlinear dynamics suggest a whole new rationale fundamentally different from the classical approach. Moreover, the clinical pharmacologist should be aware of the limitations of chaotic models for long-term prediction, which is contrary to the routine use of classical models.

If chaotic dynamics are present, experimental errors do not originate exclusively from classical randomness. Thus, the measures of central tendency used to describe or treat experimental data are questionable, since averaging is inappropriate and masks important information in chaotic systems [259].

13.2.2 Central Nervous System Drugs

The application of nonlinear dynamics to brain electrical activity offered new information about the dynamics of the underlying neuronal networks and formulated the brain disorders on the basis of qualitatively different dynamics [634].

13.2.2.1 Parametric Models

Serotonin plays an active role in temperature regulation and in particular in the maintenance of the body's set-point [698–700]. More recently, numerous pharmacological studies have suggested the involvement of homeostatic control mechanisms [699, 701] that are achieved through interplay between the 5-hydroxytryptamine (HT)1A and 5-HT2A/C receptor systems [700, 702, 703]. Administration of a

5-HT1A receptor agonist that is used therapeutically as an antidepressant and antianxiety drug causes hypothermia [704, 705].

So far, only very few of these models incorporate complex regulatory behavior [689, 706, 707]. Specifically, no mathematical models have been developed to characterize the complex time behavior of the hypothermic response in a strict quantitative manner, and neither attempts to link existing temperature regulation models [708] to pharmacokinetic models describing the time course of the drug concentration in the body.

To characterize 5-HT1A-agonist-induced hypothermia, Zuideveld et al. [709] developed a mathematical model that describes the hypothermic effect on the basis of the concept of a set-point and a general physiological response model [598, 710]. The model was applied to characterize hypothermic response vs. time profiles after administration of different doses of the reference 5-HT1A receptor agonists R- and S-8-OH-DPAT.

Example 12. Temperature Regulation

The classical three-compartment model describes pharmacokinetics of 5-HT1A receptor agonists. By means of a sigmoidal function $E(c)$, the 5-HT1A agonist concentration $c(t)$ influences the set-point signal that dynamically interacts with the body temperature. By using $x(t)$ and $y(t)$ as dimensionless state variables for the set-point and temperature, respectively, the model is expressed by the set of two nonlinear differential equations:

$$\begin{aligned} E(c) &= S_{\max} c^n(t) \left[S c_{50}^n + c^n(t) \right]^{-1}, \\ \dot{x}(t) &= A\left[1 - E(c) - y(t)\right], \\ \dot{y}(t) &= B\left[1 - x^{-\gamma}(t)\, y(t)\right], \end{aligned}$$

where the initial conditions are those at equilibrium $(x^*, y^*) = (1, 1)$. The symbols and the parameter values are as reported in [709]. Figures 13.20 and 13.21 simulate the dynamic behavior of the model for two dose levels, 200 and 1000 mg. We note that for the low dose, damped oscillations appear in the temperature $y(t)$ variable, whereas for the larger dose the perturbed temperature slowly reaches the reference value. These behaviors result from the type of the eigenvalues of the linearized model, complex conjugated for the low dose and real with negative part for the higher dose. ■

The developed model is able to reproduce the observed complex effect vs. time profile:

- When the model is not fully "pushed" into the maximal effect, a plateau phase appears. This plateau originates from damped oscillations that occur around the equilibrium point on returning to baseline. Hence, the observed plateau phase is an intrinsic part of the regulatory mechanism related to the oscillatory behavior found in many regulatory systems [711, 712].

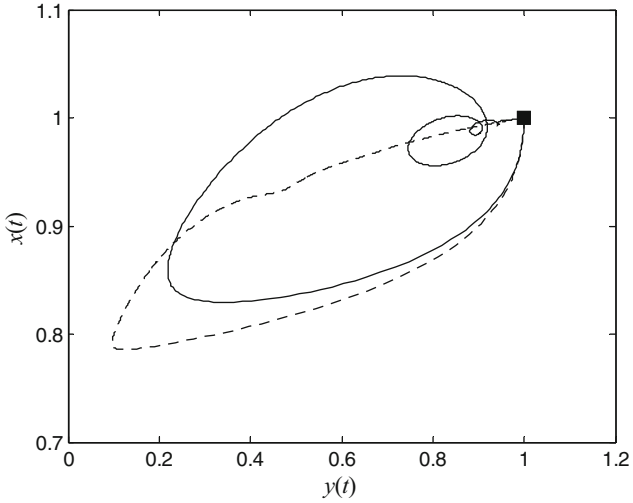


Fig. 13.20 The state space of the dimensionless set-point and temperature variables $x(t)$ and $y(t)$, respectively. *Solid* and *dashed lines* correspond to the low and high doses, respectively. (■) represents the stable equilibrium point

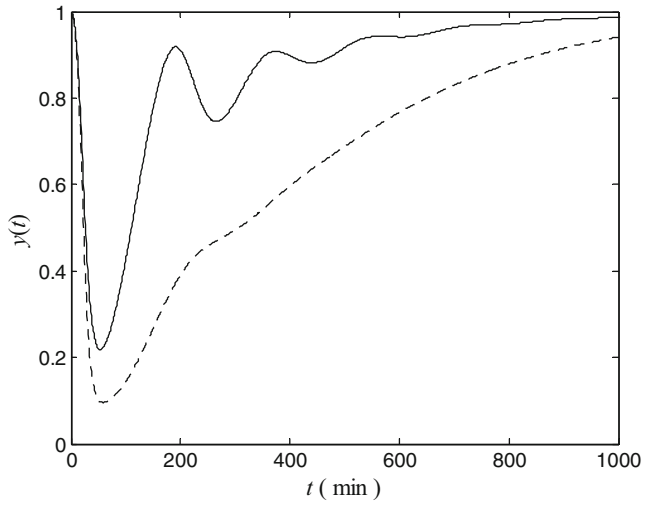


Fig. 13.21 The perturbed dynamics of the dimensionless temperature variable $y(t)$. *Solid* and *dashed lines* correspond to the low and high doses, respectively

- When the model is fully "pushed" into its maximal effect, such as in the case of a relatively high dose of a full agonist, the system becomes overdamped, thereby losing its oscillatory behavior.

The model described above has been successfully applied to characterize the in vivo concentration effect relationships of several 5-HT1A agonists including flesinoxan and buspirone [713, 714]. This model has also linked with the operational model of agonism into a full mechanism-based pharmacokinetic-dynamic model [715].

13.2.2.2 Nonparametric Models

Once again, most studies applying nonlinear tools in this field are based on experimental electroencephalogram recordings and demonstrate the irregular behavior of the brain electrical activity. Various metrics have been used to assess the electroencephalogram variability, using phase space reconstruction techniques or even calculating the fractality of the electroencephalogram recording in real time [716]. These tools, apart from pointing out the obvious complexity of the brain electrical signals, offer supplemental information to the classical techniques, such as Fourier analysis, in order to distinguish qualitatively different electroencephalogram recordings, e.g., in epileptic seizures [717], in Parkinson's disease [718], or in schizophrenia [719]. In the same context, low doses of ethanol have been found to reduce the nonlinear structure of brain activity [720]. Most of the pharmacokinetic-dynamic studies of centrally acting drugs rely on quantitative measures of electroencephalogram parameters [721]. However, an ideal electroencephalogram parameter to characterize the central nervous system effect of drugs has not been found as yet. To the best of our knowledge, time series analysis of electroencephalogram data of pharmacodynamic studies with central nervous system drugs using techniques of nonlinear dynamics are limited. Examples include investigations of the influence of anticonvulsive [722] and antiepileptic [723] drugs in epilepsy, the study of sleep electroencephalogram under lorazepam medication [724], the study of the effects of pregnenolone sulfate and ethylestrenol on rat behavior [725], the investigation of the electrophysiological effects of the neurotoxin 5, 7-dihydroxytryptamine [726], and the study of epileptiform bursts in rats after administration of penicillin and K^+ ions [727].

However, the pharmacodynamic mixed-effects model for the effect of temazepam on sleep [728] requires special mention. The model is based on hypnogram recordings and describes the probability of changes in sleep stage as a function of time after drug intake. The model predictions were found to be consistent with the observations of the effect of temazepam on sleep electroencephalogram patterns. Also, the effect of temazepam on the sleep–wake status was interpreted in terms of known mechanisms for sleep generation and benzodiazepine pharmacology.

Modeling in the brain is mainly targeted to the general qualitative principles underlying various phenomena such as epileptic seizures [729], and not to quantitative assessment and forecasting as one would expect to achieve in simpler systems. For example, in [634], recurrent inhibition and epilepsy are studied and also penicillin is considered as a γ-aminobutyric acid inhibitor.

The analysis of brain activity using tools from chaos theory can provide important information regarding the underlying dynamics if one takes into consideration that the qualitative electroencephalogram changes, induced by centrally acting drugs, e.g., ketamine, thiopental, etomidate, propofol, fentanyl, alfentanil, sufentanil, and benzodiazepines, differ considerably [721]. This exercise can also unmask the sources of extremely high variability (the coefficient of variation for model pharmacodynamic parameters of benzodiazepines in humans ranges from 30 to 100%) [721]. A plausible interpretation for the extremely high variability of pharmacodynamic parameters of benzodiazepines may be associated with the dynamic behavior of the underlying system, i.e., the recurrent inhibitory pathway of γ-aminobutyric acid [634].

It is also worthy of mention the work on the pharmacodynamics of midazolam in rats of Cleton et al. [730]. These authors found that the rate of change in plasma concentration is an important determinant of midazolam pharmacodynamics. In addition, the relationship found between the rate of change of blood concentration and the values of the different pharmacodynamic parameters is rather complex. These findings indicate that in vivo a homeostatic control mechanism is operative that may modify sensitivity to midazolam and whose activation is largely influenced by the rate of presentation of the drug in blood.

Keeping patients at a well-defined level of anesthesia is still a difficult problem in clinical practice. If anesthesia is too deep, a decompensation of the cardiovascular system is threatening. When anesthesia is too weak, the patient may wake up. Depth of anesthesia is expected to be reflected in the electroencephalogram. In current clinical practice, one or a few channels of the electroencephalogram are routinely displayed during difficult anesthesias. Since the attending personnel have to monitor several critical parameters (blood pressure, heart rate, etc.), the vast amount of information contained in the electroencephalogram must be severely condensed in order to be useful. Only a few numbers may be monitored at a typical intervention time scale. More pragmatically, a single number should be produced that indicates the instantaneous depth of anesthesia of the patient.

In that spirit, Widman et al. [731] adapted a prescription for an overall index of nonlinear coherence that has been found powerful for anticipating epileptic seizures from implanted electrode recordings. This index based on phase space reconstruction and correlation sums was called d^*, and it contains many ingredients familiar from the Grassberger–Procaccia algorithm for the correlation dimension [732].

Widman et al. [733] compared several indices measuring the depth of anesthesia from electroencephalogram data gathered from 17 patients undergoing elective surgery and anesthetized with sevoflurane. Two of these measures are based on the power spectrum, and the third is the bispectral index (BIS) [734]. The power spectrum measures are essentially useless and unreliable as indicators of depth of anesthesia in the investigated group of patients. While for both of the two nonlinear measures, bispectral index and d^*, such a relationship seems to exist, the correlation is strongest for d^*. Dimension d^* seems to be able to improve the quantification of depth of anesthesia from brain electrical activity, at least when sevoflurane is used as an anesthetic drug. To assess the depth of anesthesia of the patient, Bruhn et al. [735] recently proposed another index based on the Shannon entropy.

13.2.3 Cardiovascular Drugs

Numerous applications of nonlinear dynamics and chaos theory to cardiac physiology have been published [736]. Many techniques, either statistical, like spectral analysis, or dynamic, like phase space reconstruction, applied to electrocardiogram data clearly indicate that the frequency of the heartbeat is essentially irregular. The electrocardiogram was in fact one of the first biological signals studied with the tools of nonlinear dynamics. Studies applying concepts from chaos theory to electrocardiogram data, regarding the effects of drugs on the dynamics of cardiac physiology, have also been published. Examples include the effect of atropine on cardiac interbeat intervals [737], the induction of cellular chaos during quinidine toxicity [738], the attempt to control cardiac chaos using ouabain [739], and the effect of anticholinergic drugs on heart rate variability [740].

Another very successful application of nonlinear dynamics to the heart is through mathematical modeling. An example in which a simple model based on coupled oscillators describes the dynamics of agonist-induced vasomotion is in the work of de Brouwer et al. [741], where the route to chaos in the presence of verapamil, a class IV antiarrhythmic drug, is studied.

Undoubtedly, the most promising modeling of the cardiac dynamics is associated with the study of the spatial evolution of the cardiac electrical activity. The cardiac tissue is considered to be an excitable medium whose electrical activity is described both in time and space by reaction–diffusion partial differential equations [674]. This kind of system is able to produce spiral waves, which are the precursors of chaotic behavior. This consideration explains the transition from normal heart rate to tachycardia, which corresponds to the appearance of spiral waves, and the following transition to fibrillation, which corresponds to the chaotic regime after the breaking up of the spiral waves, Figure 13.22. The transition from the spiral waves to chaos is often characterized as electrical turbulence due to its resemblance to the equivalent hydrodynamic phenomenon.

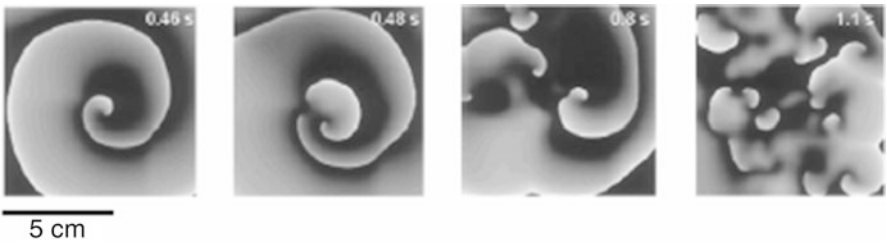


Fig. 13.22 The four snapshots show the evolution and breakup of a spiral wave pattern in 2-dimensional simulated cardiac tissue (300 × 300 cells). The chaotic regime shown in the final snapshot corresponds to fibrillation. Reprinted from [742] with permission from Lippincott Williams and Wilkins

These concepts have been successfully applied to the effect of antiarrhythmic drugs as well. It is widely known that although class II antiarrhythmic drugs, like isoproterenol, have shown satisfactory results [743], class I and III agents, such as encainide, flecainide, and moricizine, have been shown even to increase sudden death rate caused by ventricular fibrillation [744]. Although it is unclear how to integrate the drug action in the excitable media models, successful attempts have been made to simulate, mainly, two-dimensional cardiac tissue [745, 746]. Three-dimensional cardiac tissue has been simulated as well [747], where the three-dimensional equivalent of spiral waves, the scroll waves, appear. These models explain how a drug can exhibit antiarrhythmic action in a single-cell system, which ignores the spatial evolution, while acting as proarrhythmic in a system of a whole cardiac tissue of spatial dimension 2 or 3. This has given rise to a new approach for antiarrhythmic drug evaluation based on the chaotic dynamics of transition from tachycardia to fibrillation [742, 746, 747], which is also supported by experimental evidence [747]. The results of these recent studies [742] indicate that the failure to predict long-term efficacy of class I and III antiarrhythmic agents in patients with ischemic heart disease [744] may be associated with the limitations of the classical approach, which is based only on the suppression of premature ventricular polarization on the electrocardiogram, i.e., the initiation of tachycardia. Sudden cardiac death resulting from ventricular fibrillation, however, is separated into two components: initiation of tachycardia and degeneration of tachycardia to fibrillation. These studies suggest that a new antiarrhythmic drug classification scheme must be adopted, which should incorporate the antifibrillatory profile based on results from excitable media modeling, together with the classical antitachycardiac profile (classes I to IV scheme). Also, the drug bretylium is proposed as a prototype for future development of antifibrillatory agents [747].

In the pharmaceutical literature [748] the pharmacodynamics of antiarrhythmic drugs are treated with the classical models, $E_{\max}$, indirect link with effect compartment, etc. Variability, wrong dosage scheme, narrow therapeutic index, and lack of individualization of treatment are the dominant interpretations for the failure of these drugs. Another factor held responsible for failure in treatment with antiarrhythmics is the possible nonbioequivalency of the generics used [749]. However, classical bioequivalence studies are based only on the comparison of pharmacokinetic parameters of the formulations ($c_{\max}$, area under curve (AUC)). Although testing for therapeutic equivalence is implied, pharmacodynamics are not taken into account at all. Thus, classical bioequivalence studies may be inappropriate for assessing the effects of antiarrhythmic drugs if their mechanism of action arises from nonlinear dynamic processes.

These studies show that it is possible to predict the time course of drug effects in vivo in situations in which complex homeostatic control mechanisms are operative. As such, they form the basis for the development of an entirely new class of pharmacokinetic-dynamic models. These models are important for the development of new drugs and the application of such drugs in clinical practice. For example, on

the basis of this kind of model, it becomes possible to predict whether withdrawal symptoms will occur on cessation of (chronic) drug treatment. Hence, these models may provide a scientific basis either for the selection of alternative drug candidates or the design of dosing regimens that show less-pronounced withdrawal phenomena. It is further anticipated that such models will provide a basis for pharmacokinetic-dynamic modeling with disease progression.

Chapter 14
Concluding Notes

I know one thing: that I know nothing (The Socratic paradox).
Sokrates (470/469-399 BC)
There is some disagreement about whether it accurately represents a Socratic view.

In various fields of research, scientists are finding that heterogeneity is everywhere: the heterogeneous conditions prevail in numerous physical, chemical, physiological, and biochemical processes. Indeed, many reactions and processes take place under dimensional or topological constraints that introduce structural heterogeneity. In parallel, drug molecules can differ in their kinetic behavior because of inherent variability in their characteristics such as molecular weight, chemical composition, or hepatic clearance involving a large number of metabolites. All these features introduce functional heterogeneity. Structural and functional heterogeneities can be described and understood with the concept of fractals.

On the other hand, series of measurements from many physiological processes appear randomly variable. The determinants of the observed variability cannot be known because of the multiplicity and interconnectivity of the factors affecting the phenomena. This idea relies on the classical view of randomness, which requires that a complex process have a large number of degrees of freedom that are not directly observed but whose presence is manifested through fluctuations. However, scientists from various fields of research have shown that the "randomness" generated by deterministic dynamic processes exhibits spectra practically indistinguishable from spectra of pure random processes. This is referred to as chaos, a specific behavior of nonlinear dynamic systems. Since most drugs are modifiers of physiological and biochemical states, it follows that many concepts of modern nonlinear dynamic theory have potential applications to pharmacology and drug development.

The aim of the book was to treat heterogeneity and analyze nonlinear behavior in the pharmaceutical processes (biopharmaceutics, pharmacokinetics, pharmacodynamics) by using modeling and simulation techniques.

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7_14

14.1 Heterogeneous Processes

In vivo drug dissolution, release, and uptake are heterogeneous processes since they take place at interfaces of different phases, e.g., liquid–solid and liquid–membrane boundaries, while diffusion, which is the principal mechanism of all processes, operates under topological constraints. In addition, all processes occur in heterogeneous environments, e.g., variable stirring conditions in the lumen. Also, the geometric constraints imposed by the heterogeneous fractal-like structure of the blood vessel network and the liver strongly modify drug dynamics.

Given the above considerations one can argue that drugs can be classified with respect to their gastrointestinal absorption characteristics into two broad categories, i.e., homogeneous and heterogeneous. Homogeneous drugs have satisfactory solubility and permeability, and are dissolved and absorbed mostly prior to their arrival in the large intestine. In contrast, drugs with low solubility and permeability can be termed heterogeneous, since they traverse the entire gastrointestinal tract, and are most likely to exhibit heterogeneous transit, dissolution, uptake, distribution, metabolism, and elimination.

The forecast of pharmacodynamic outcomes is more complex than that of biopharmaceutics and pharmacokinetics. This complexity first relies on the presence of nonlinear elements and multiple homeostatic regulations implying negative feedback mechanisms. In this context, physiological control mechanisms can generate exceedingly complex oscillations and generate chaos. Therefore, deterministic chaos is typically the recorded behavior of complex pharmacodynamics.

We deal with complex biological processes involving a large number of interrelated components exhibiting unpredictable behavior; this is the source of an important portion of the observed biological fluctuations. This information changes the view of biological variability. In the past, it was thought that the source of variability was external to the internal workings of the organism, that environmental factors, such as temperature, food ingestion, immobilization, and venous occlusion, were responsible for the short-term changes. Today, the observed biological variability is rather ascribed to the heterogeneity of the studied process.

14.2 Detecting Heterogeneity

Heterogeneous reactions taking place at interfaces, membrane boundaries, or within a complex medium like a fractal, when the reactants are spatially constrained on the microscopic level, culminate in deviant reaction rate coefficients that appear to have a sort of temporal memory. Fractal kinetic theory suggested the adoption of a time-dependent rate "constant," with power-law form, determined by the spectral dimension. This time dependency could be revealed from empirical models. When the observations are optimally fitted by the power-law and gamma empirical

models, the underlying processes are rather time-varying. The time-varying features of the observed processes are in fact the expression of functional or structural heterogeneities in the process.

Power-law expressions are found at all hierarchical levels of organization from the molecular level of elementary chemical reactions to the organismal level of growth and allometric morphogenesis. This recurrence of the power law at different levels of organization is reminiscent of fractal phenomena. The reverse is also true; if a power function of time describes the observed data, the reaction takes place in fractal physical support.

Useful results related to nonlinear dynamics can be obtained from observations gathered from real processes. Real-life observations, like biological signals, are usually time series of measured quantities. Instead of studying a time series statistically, the idea is to consider it as if it came out of a dynamic system. Then, one tries to reconstruct its phase space and see whether any structure is detectable, either visually or using certain mathematical and numerical tools. The absence of any structure in phase space means that the system is random. However, the presence of structure is evidence of the dynamic origin of the time series and the existence of an attractor. The dimension of the attractor can give us information about the dynamic behavior of the whole system. If, for example, the dimension of the attractor is not an integer, it corresponds to a strange attractor and the system exhibits chaotic behavior.

The observed time-varying features of a process are expressions of structural and functional heterogeneity. Thus, heterogeneity may be at the origin of fluctuations, which are the prelude of instability and chaotic behavior.

14.3 Modeling Heterogeneity

Empirical models helped us to recognize heterogeneity in the process and express it simply by mathematical models with time-varying parameters. The heterogeneous process operating in several circumstances and the resulting complexity of the kinetic behaviors require new techniques in order to comply with observations. In this way, two operational procedures may be retained:

- From a holistic point of view, the time-varying parameters fitting the observed data could represent the dynamic behavior of a complex system implying the states in feedback regulations and leading to nonlinear kinetics.
- From a molecular point of view, the stochastic formulation would be the most appropriate for capturing the structural and functional heterogeneity in the biological media that generates the observed uncertainty and fluctuations in the real process.

Stochastic compartmental analysis assumes probabilistic behavior of the molecules in order to describe the heterogeneous character of the processes. The huge amount of knowledge needed to describe heterogeneity could be summarized

only by the statistical concepts provided by stochastic modeling approaches. Stochastic modeling is able to (1°) generate process uncertainty, (2°) express process memory in the process, and (3°) supply tractable forms to fit the data involving time-varying parameters. The usual deterministic approach is incapable of describing all these features accurately.

Moreover, the time-varying parameters implied in stochastic models highlight the presence of feedback regulation mechanisms involving states of the process. Introduction of these states leads to nonlinear dynamic modeling associated with various levels of stability. Near state instabilities in such nonlinear systems, fluctuations and correlations can produce dramatic effects; for these systems the stochastic formulation would be the more appropriate choice. Thus, one must frequently expect chaotic-like behavior when the process is heterogeneous. In contrast, it is impossible to expect chaotic properties with homogeneous processes.

14.4 Estimation and Control

Usually a mathematical model simulates a process behavior, in what can be termed a forward problem. The inverse problem is the following: given the experimental measurements of behavior, what is the structure? This is a difficult problem, but an important one for the sciences. The inverse problem may be partitioned into the following stages: hypothesis formulation, i.e., model specification, definition of the experiments, identifiability, parameter estimation, experiment, and analysis and model checking.

Since this monograph is devoted only to the conception of mathematical models, the inverse problem of estimation is not fully detailed. Nevertheless, estimating parameters of the models is crucial for verification and applications. Any parameter in a deterministic model can be sensibly estimated from time series data only by embedding the model in a statistical framework. This is usually performed by assuming that instead of exact measurements on concentration, these values are blurred by observation errors that are independent and normally distributed. The parameters in the deterministic formulation are estimated by nonlinear least-squares or maximum likelihood methods.

The meaning of the estimation in the models with heterogeneous particles is qualitatively different from that described above because the constants have been replaced by random variables. Now, by fitting the model to the observed data, we obtain estimates of the parameters involved in the probability density function of the random variables. Moreover, the weighting scheme in the nonlinear regression is not the same as in the deterministic case. So we need to take into account the process uncertainty and the measurement error components that blur the observations. Special computational methods in nonlinear regression are available (unconditional and conditional generalized least squares, etc.), and the classical maximum likelihood approach is also possible based on the multinomial distribution of particles in the compartments.

Table 14.1 Process and model features associated with homogeneous and heterogeneous media.

	Media	
Models or features	**Homogeneous**	**Heterogeneous**
Empirical model	Sum of exponentials	Power law
Phenomenological model	Deterministic	Stochastic
Retention probability	Exponential	Weibull
Process memory	No	Yes
Process uncertainty	No	Yes

The presence of chaos may be a great advantage for control in a variety of situations. Therapeutic strategies should aim to invert the progress of the disease and restore normal physiological conditions by interfering with the control parameters. It is widely recognized that chaotic behavior dominates physiological systems. Moreover, periodic or other nonchaotic states are considered pathological, whereas chaotic behavior is considered to be the normal, healthy state. The reason for this must be associated with a fundamental advantage of nonlinear over classical systems. Indeed, small variations of the control parameters may offer finer, more rapid, and more energy-efficient controllability of the system compared to linear systems. Processes generating chaos are generally more controllable and offer better control performances than linear systems do. This may be the reason why nature prefers chaos to regularity, and of course the latter is a good enough reason for applied biological sciences such as biopharmaceutics, pharmacokinetics, and pharmacodynamics to adopt this rationale to a greater extent.

In short, the presented procedures show how to expect chaotic behaviors with processes revealing uncertainty and which are described by models involving time-varying parameters. All these considerations oriented us to complete the initial Table 1 referenced in the preface by Table 14.1.

Appendix A
Stability Analysis

Stability is determined by eigenvalue analysis at an equilibrium point for flows and by characteristic multiplier analysis of a periodic solution at a fixed point for maps [3].

- The *equilibrium point* $\underline{y}^*$ for flows is the solution of $\underline{g}\left(\underline{y}^*, t, \theta\right) = \underline{0}$. Local behavior of the flow near $\underline{y}^*$ is determined by linearizing $\underline{g}$ at $\underline{y}^*$; let $\mathbf{A}$ be the matrix formed by elements

$$a_{jk} = \left.\frac{\mathrm{d}g_j\left(\underline{y}\right)}{\mathrm{d}y_k}\right|_{\underline{y}=\underline{y}^*}.$$

  Let the eigenvalues of $\mathbf{A}$ be ζ_j with corresponding eigenvectors $\underline{\eta}^{(j)}$. If ζ_j is real, the eigenvalue is the rate of contraction (if $\zeta_j < 0$) or expansion (if $\zeta_j > 0$) near $\underline{y}^*$ in the direction of $\underline{\eta}^{(j)}$. If ζ_j are complex-conjugate pairs, the trajectory is a spiral in the phase space spanned by $\mathrm{Re}\left[\underline{\eta}^{(j)}\right]$ and $\mathrm{Im}\left[\underline{\eta}^{(j)}\right]$. The real part of ζ_j gives the rate of contraction (if $\mathrm{Re}\left[\zeta_j\right] < 0$) or expansion (if $\mathrm{Re}\left[\zeta_j\right] > 0$) of the spiral; the imaginary part of the eigenvalue is the frequency of rotation. Hence, one can conclude that if $\mathrm{Re}\left[\zeta_j\right] < 0$ for all ζ_j, then all sufficiently small perturbations tend toward 0 as $t \to \infty$, and $\underline{y}^*$ is asymptotically *stable*. If $\mathrm{Re}\left[\zeta_j\right] > 0$ for all ζ_j, then any perturbation grows with time, and $\underline{y}^*$ is *unstable*. If there exist j and k such that $\mathrm{Re}\left[\zeta_j\right] < 0$ and $\mathrm{Re}\left[\zeta_k\right] > 0$, then $\underline{y}^*$ is *unstable*. An unstable equilibrium point is often called a *saddle point*. A stable or unstable equilibrium point with no complex eigenvalues is often called a *node*.

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7

- The *fixed point* $\underline{y}^*$ for maps is the solution of $\underline{y}^* = \underline{g}\left(\underline{y}^*, \theta\right)$. The local behavior of the map near $\underline{y}^*$ is determined by linearizing the map at $\underline{y}^*$; let $\mathbf{A}$ be the matrix formed by elements

$$a_{jk} = \left.\frac{\mathrm{d}g_j\left(\underline{y}\right)}{\mathrm{d}y_k}\right|_{\underline{y}=\underline{y}^*}.$$

  Let the eigenvalues of $\mathbf{A}$ be ξ_j with corresponding eigenvectors $\underline{\eta}^{(j)}$. The eigenvalues ξ_j are called *characteristic multipliers* and they are a generalization of the eigenvalues at an equilibrium point. The characteristic multipliers' position in the complex plane determines the stability of the fixed point. If ξ_j is real, the characteristic multiplier is the amount of contraction (if $\xi_j < 1$) or expansion (if $\xi_j > 1$) near $\underline{y}^*$ in the direction of $\underline{\eta}^{(j)}$ for one iteration of the map. If ξ_j are complex-conjugate pairs, the orbit is a spiral in the phase space spanned by $\mathrm{Re}\left[\underline{\eta}^{(j)}\right]$ and $\mathrm{Im}\left[\underline{\eta}^{(j)}\right]$. The magnitude of ξ_j gives the amount of expansion (if $\left|\xi_j\right| > 1$) or contraction (if $\left|\xi_j\right| < 1$) of the spiral for one iteration of the map; the angle of the characteristic multiplier is the frequency of rotation. Hence, one can conclude that if $\left|\xi_j\right| < 1$ for all ξ_j, then all sufficiently small perturbations tend toward 0 as $i \to \infty$, and $\underline{y}^*$ is asymptotically *stable* and is said to be an *attracting equilibrium*. If $\left|\xi_j\right| > 1$ for all ξ_j, then any perturbation grows with iterations, and $\underline{y}^*$ is *unstable*. If there exist j and k such that $\left|\xi_j\right| < 1$ and $|\xi_k| > 1$, then $\underline{y}^*$ is *unstable*. An unstable fixed point is often called a *saddle point*. The critical values $\xi = \pm 1$ are where the fixed point $\underline{y}^*$ changes its behavioral character. The case $\xi = 1$ is called a *tangent bifurcation* and the case $\xi = -1$ is called a *pitchfork bifurcation*.

Both equilibrium and fixed points are simply referenced as *steady states*. The matrix $\mathbf{A}$ of the linearized system is called the *Jacobian* of the system.

Appendix B
Monte Carlo Simulations in Drug Release

Models that can be discretized either naturally or through approximation are suitable for study using Monte Carlo simulations. As an example, we have provided a brief outline below of drug release simulations from cylinders assuming Fickian drug diffusion and excluded volume interactions. This means that each molecule occupies a volume V where no other molecule can be at the same time.

First, a three-dimensional lattice in the form of a cube with L^3 sites is constructed. Next, a cylinder inside this cubic lattice is defined. The cylinder can leak from its side, but not from its top or bottom. A site is uniquely defined by its 3 indices i, j, k (coordinates). The sites are labeled as follows (R is the radius of the cylinder):

- When for a site $(R-1)^2 \leq i^2 + j^2 \leq R^2$, it is considered to be a leak site and it is marked as such.
- If $i^2 + j^2 \leq (R-1)^2$, then it belongs to the interior of the cylinder and it can host drug molecules.
- If, on the other hand, $i^2 + j^2 > R^2$, then it is outside the cylinder, and it is marked as a restricted area, so that particles are not allowed to go there; cf. Figure B.1 for a schematic.

When spherical matrices are constructed, the sites with indices $i^2 + j^2 + k^2 > R^2$ are considered outside of the sphere with radius R and marked as a restricted area, while leak sites are those whose indices satisfy the inequalities $(R-1)^2 \leq i^2 + j^2 + k^2 \leq R^2$.

The simulation method proceeds as follows: a number of particles are placed randomly on the sites of the cylinder, according to the initial concentration, avoiding double occupancy. The diffusion process is simulated by selecting a particle at random and moving it to a randomly selected nearest-neighbor site. If the new site is an empty site, then the move is allowed and the particle is moved to this new site. If the new site is already occupied, the move is rejected, since excluded volume interactions are assumed.

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7

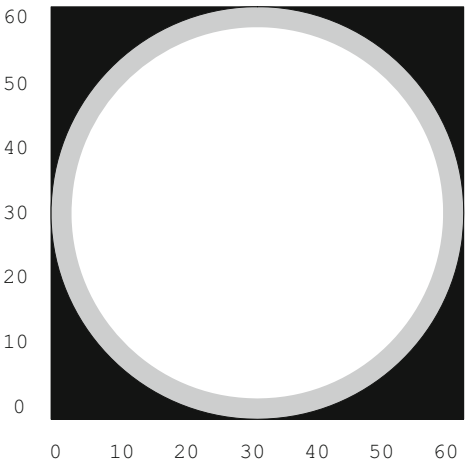


Fig. B.1 A cylindrical cross section with radius $R = 30$ sites. The dark area is restricted to particles. The gray area indicates the leaking sites. The white area is where the drug particles are initially located. Each site in the white area can be either occupied or empty

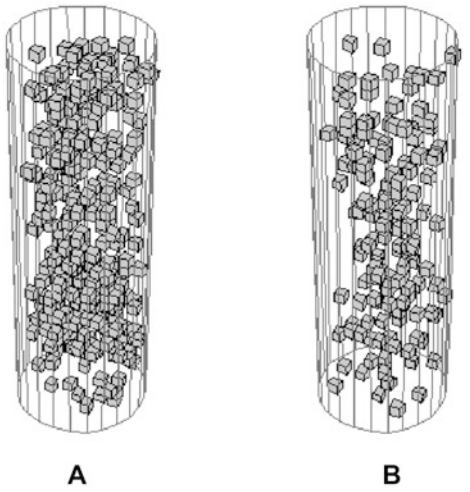


Fig. B.2 (**A**) A cylinder with radius 5 units and half-height 20 units initially contains 282 particles at completely random positions. Each particle is represented by a cuboid of volume 1 $(\text{unit})^3$. (**B**) A snapshot of the same cylinder during the release procedure. Now only 149 particles are left inside the cylinder. The positions of the particles are no longer completely random. On average a concentration gradient forms with fewer particles at the cylinder border

A particle is removed from the lattice as soon as it migrates to a site lying within the leak area. After each particle moves, time is incremented. The increment is chosen to be $1/n\,(t)$, where $n\,(t)$ is the number of particles remaining in the system. This is a typical approach in Monte Carlo simulations. The number of particles that are present inside the cylinder as a function of time is monitored until the cylinder is completely empty of particles. The results are averaged using different initial random configurations, but the same parameter. A pictorial view of particles in the cylinder at two different time points is presented in Figure B.2.

Appendix C
The Population Approach

The goal of pharmacokinetic and pharmacodynamic investigations is to establish a rational basis for the therapeutic use of a drug. Specifically, clinical trials aim at determining the dose and the dosage regimen of the new drug that will produce therapeutic benefit in patients while minimizing the inconvenience of side effects and risks of adverse drug reactions. This is particularly true in the clinical evaluation of new chemical and biological entities during drug development [750].

Data destined for pharmacokinetic analysis consist of one or more drug concentration vs. time observations, while pharmacodynamic data consist of specific concentration levels corresponding to a specific therapeutic effect or its validated biomarker. One distinguishes two types of data:

- *Experimental data* arise from studies carried out specifically for pharmacokinetic investigations, under controlled conditions of drug dosing and extensive blood sampling.
- *Observational data* are collected as a supplement in a study designed and carried out for another purpose. These data are characterized by lack of control and few design restrictions: the amount of kinetic data collected from each individual is variable, the timing of blood sampling differs, and the number of blood samples per patient is small, typically from 1 to 5.

It should be emphasized that in the collected data, several responses may be measured (e.g., drug plasma concentration, arterial blood pressure), and diverse administration schedules (single dose and chronic dosing) may be considered.

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7

C.1 Inter- and Intraindividual Variability

The population approach is a new point of view in clinical drug evaluation and therapy. It emphasizes the estimation of parameters describing the dose–concentration–response relationship both between and within patients (including average behavior and variability). The population approach recognizes variability as an important feature that should be identified and measured during drug evaluation [751]. Indeed:

- We need to know something about the distributions of the deviations of individual patient pharmacokinetic-dynamic model parameters from their population average values, and how these deviations correlate with one another. The deviations are population parameters of a different type: random individual effect parameters; random because individual deviations are regarded as occurring according to chance mechanisms.
- One may ask how much drug outcome (concentration/effect) varies across a modeling cycle within an individual. To answer this question, other random-effect population parameters are also needed: the variance of the combined random intraindividual and measurement error; random because outcome fluctuations and measurement errors are also regarded as occurring according to chance mechanisms.
- One may immediately imagine further subdividing the last type of variability. For example, one might wish to distinguish intraindividual variability due to different aspects of kinetics and separate all such variability from that due to measurement error. The problem with doing so is that most data are insufficiently detailed and not complete enough to allow these components of variance to be estimated separately. The two-way division we have proposed appears to suffice for most applications and data sets.

According to the population approach, the analysis of collected data requires an explicit mathematical model, including parameters quantifying population mean profiles, interindividual variability, and residual variability including intraindividual variability and measurement error [752].

C.2 Models and Software

Nonlinear mixed-effects modeling methods as applied to pharmacokinetic-dynamic data are operational tools able to perform population analyses [753]. In the basic formulation of the model, it is recognized that the overall variability in the measured response in a sample of individuals, which cannot be explained by the pharmacokinetic-dynamic model, reflects both interindividual dispersion in kinetics and residual variation, the latter including intraindividual variability and measurement error. The observed response of an individual within the framework of a population nonlinear mixed-effects regression model can be described as

$$y_{ij} = g\left(\underline{\theta}_i, t_{ij}\right) + \varepsilon_{ij},$$

where y_{ij} for $j = 1 : n_i$ are the observed data at time points t_{ij} of the i-th individual. An appropriate model of this type is defined for all $i = 1 : m$, where m is the number of individuals in the sample. The function $g\left(\underline{\theta}, t\right)$ is a specific function for predicting the response, $\underline{\theta}_i$ is the vector of unknown individual-specific parameters, and ε_{ij} accounts for the error between the unknown value and the corresponding measurement.

The sample of individuals is assumed to represent the patient population at large, sharing the same pathophysiological and pharmacokinetic-dynamic parameter distributions. The individual parameter $\underline{\theta}$ is assumed to arise from some multivariate probability distribution $\underline{\Theta} \sim f\left(\underline{\Psi}\right)$, where $\underline{\Psi}$ is the vector of so-called hyperparameters or "population characteristics." In the mixed-effects formulation, the collection of $\underline{\Psi}$ is composed of population "typical values" (generally the mean vector) and of population "variability values" (generally the variance–covariance matrix). Mean and variance characterize the location and dispersion of the probability distribution of $\underline{\Theta}$ in statistical terms.

Then, given a model for data from a specific drug in a sample from a population, mixed-effect modeling produces estimates for the complete statistical distribution of the pharmacokinetic-dynamic parameters in the population. Especially, the variance in the pharmacokinetic-dynamic parameter distributions is a measure of the extent of inherent interindividual variability for the particular drug in that population (adults, neonates, etc.). The distribution of residual errors in the observations, with respect to the "mean" pharmacokinetic or pharmacodynamic model, reflects measurement or assay error, model misspecification, and, more rarely, temporal dependence of the parameters.

Population modeling software varies in the number of assumptions made regarding the statistical distributions of the pharmacokinetic-dynamic parameters, the within-individual or residual error, and, particularly, the interindividual variability (random effect). They take either a parametric approach with strong assumptions, typically of a log-normal distribution [754–756] or Bayesian approaches [757], a semiparametric view with relaxed assumptions [758], or a nonparametric, no assumptions approach [759, 760]. NONMEM (NONlinear Mixed Effect Modeling, NONMEM Project Group, University of California at San Francisco, CA [754]) and NPEM2 (NonParametric Expectation Maximization, Laboratory of Applied Pharmacokinetics, University of Southern California, Los Angeles, CA [760]) are pioneers of parametric and nonparametric population modeling packages, respectively. Recently, Aarons [761] reviewed some of the software currently available for performing nonlinear mixed-effects modeling.

However, the Bayesian analysis using Gibbs sampling (BUGS) requires special mention since it is a general program for performing analysis for a wide range of statistical problems and is available on PC and Unix platforms and also in a PC Windows version. The Bayesian analysis is based on complex statistical models using Markov chain Monte Carlo methods [762–764].

C.3 Covariates

In the initial stage of the analysis, the pharmacokinetic or pharmacodynamic observations are blind with respect to the patients, i.e., no patient-specific demographic or physiological covariates are included, other than the dose. Both parametric and nonparametric population methods then, in this first stage, produce a base model for the centering and spread of the parameters in the population, which can then be used in subsequent steps in various ways.

However, the base model provides inadequate individualization, and to assist clinical decision-making, it is important to relate differences among individuals to readily identifiable and routinely measurable individual attributes or covariates, such as demographic (e.g., age), pathophysiological (e.g., serum creatinine, renal, or hepatic function), or genotypic (e.g., CYP2D6 polymorphism) data. Knowing the value of an influential covariate in a new patient before starting therapy increases the predictive power and therefore makes the choice of dose more reliable.

Explanation of parameter variability using covariates can be achieved:

- by simple regression of the individual empirical Bayes parameters from the base model with the covariates and
- within the population fitting process, estimating the covariate term coefficients jointly with the pharmacokinetic parameters.

Parametric population methods also obtain estimates of the standard error of the coefficients, providing consistent significance tests for all proposed models. A hierarchy of successive joint runs, improving an objective criterion, leads to a "final" covariate model for the pharmacokinetic parameters. The latter step reduces the unexplained interindividual randomness in the parameters, achieving an extension of the deterministic component of the pharmacokinetic model at the expense of the random effects. Recently used individual empirical Bayes estimations exhibit more success in targeting a specific individual concentration after the same dose.

C.4 Applications

The knowledge of population kinetic parameters has been proved important, and up to the present, the population approach has had a wide spectrum of applications:

- It is a currently accepted medical practice to measure a few drug levels after dosage has progressed for some time. In order to make the measured drug levels useful, one should estimate individual parameters using Bayesian estimation techniques. They consist in combining a few or even a single individual drug level measurement with the probability distribution function expressing interindividual variability. Once the individual parameters are obtained, the time-dependent pharmacokinetic model can be used for forecasting and predictive exploration of dosing regimens.

- Most decisions regarding drug regulation involve knowledge of the typical or average behavior of a drug in a population. To the extent that pharmacokinetic aspects of drugs are of interest to drug regulatory agencies, population pharmacokinetics will also be of interest.
- Although intraindividual kinetic variability has only been regarded as a nuisance, the typical degree of intraindividual kinetic variability from all causes can be used to fix rational limits on the increments for tablet dosage, and on permissible tablet-to-tablet and lot-to-lot variability.
- Finally, in drug development or evaluation phase studies, logistical trade-offs of pharmacokinetic-dynamic data may lead to reduced samples per patient (sparse data) and/or reduced patient group sizes, as well as noisy data (e.g., unknown variability in the dose strategy, noncompliance) (phase IV).

The ability to handle, in a statistically rigorous explanatory and predictive framework, large data sets of drug-related pharmacokinetic-dynamic clinical observations is of increasing importance to the industry, regulatory agencies, and patients, in order to reduce human and budgetary risks.

Beyond pharmacokinetics and pharmacodynamics, population modeling and parameter estimation are applications of a statistical model that has general validity, the nonlinear mixed-effects model. The model has wide applicability in all areas, in biomedical science and elsewhere, where a parametric functional relationship between some input and some response is studied and where random variability across individuals is of concern [750].

Appendix D
Probability

D.1 Basic Properties

- *Poincaré theorem*. Given n random events $A_1, \ldots, A_n$, the probability of their union is given by

$$\Pr\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} \Pr[A_i] - \sum_{j=2}^{n} \sum_{i=1}^{j-1} \Pr\left[A_i \bigcap A_j\right] + \sum_{j=3}^{n} \sum_{k=2}^{j-1} \sum_{i=1}^{k-1} \Pr\left[A_i \bigcap A_j \bigcap A_k\right] - \ldots + (-1)^n \Pr\left[\bigcap_{i=1}^{n} A_i\right].$$

  If the events are *mutually exclusive*, i.e., $\forall i, j \; A_i \cap A_j = \varnothing$, then

$$\Pr\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} \Pr[A_i].$$

- *Conditional probability*. Given the random events A and B, the conditional probability of A for observed B is defined by

$$\Pr[A \mid B] \triangleq \frac{\Pr[A \cap B]}{\Pr[B]}.$$

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7

Two events A and B are defined as *independent* if $\Pr[A \mid B] = \Pr[A]$, or $\Pr[A \cap B] = \Pr[A]\Pr[B]$. Given n random events $A_1, \ldots, A_n$, the probability of their intersection is given by

$$\Pr\left[\bigcap_{i=1}^{n} A_i\right] = \Pr\left[A_1 \mid \bigcap_{i=2}^{n} A_i\right] \Pr\left[A_2 \mid \bigcap_{i=3}^{n} A_i\right] \ldots \Pr[A_{n-1} \mid A_n] \Pr[A_n].$$

- *Total probability theorem*. Given n mutually exclusive events $A_1, \ldots, A_n$, whose probabilities sum to unity, then

$$\Pr[B] = \Pr[B \mid A_1]\Pr[A_1] + \ldots + \Pr[B \mid A_n]\Pr[A_n],$$

  where B is an arbitrary event, and $\Pr[B \mid A_i]$ is the conditional probability of B assuming A_i.
- *Bayes theorem*. For the same settings, the Bayes theorem gives the conditional probability

$$\Pr[A_i \mid B] = \frac{\Pr[B \mid A_i]\Pr[A_i]}{\sum_{k=1}^{n} \Pr[B \mid A_k]\Pr[A_k]}.$$

D.2 Expectation, Variance, and Covariance

For scalar continuous random variables X and Y with joint probability density $f(x, y)$, marginals and conditionals are defined as

$$f(x) = \int_y f(x, y)\,\mathrm{d}y,\ f(y) = \int_x f(x, y)\,\mathrm{d}x$$

and

$$f(x \mid y) = f(x, y)/f(y),\ f(y \mid x) = f(x, y)/f(x),$$

respectively. The statistical characteristics up to second order of X and Y are:

- *Expectation*. It can be interpreted as the center of gravity of random variables:

$$E[X] \triangleq \int_x xf(x)\,\mathrm{d}x,\ E[Y] \triangleq \int_y yf(y)\,\mathrm{d}y$$

  on X and Y axes, respectively, or

$$E[XY] \triangleq \int_{xy} xyf(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

  on X, Y plan. If X and Y are independent, $E[XY] = E[X]E[Y]$.

- *Variance*: It can be interpreted as the inertia about the centers of gravity $E[X]$ and $E[Y]$:

$$Var[X] \triangleq \int_x \{x - E[X]\}^2 f(x)\,\mathrm{d}x,\ Var[Y] \triangleq \int_y \{y - E[Y]\}^2 f(y)\,\mathrm{d}y.$$

- *Covariance*:

$$\begin{aligned} Cov[X,Y] &\triangleq \int_{x,y} \{x - E[X]\}\{y - E[Y]\} f(x,y)\,\mathrm{d}x\mathrm{d}y \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

and *correlation*:

$$Cor[X,Y] \triangleq \frac{Cov[X,Y]}{\sqrt{Var[X]\,Var[Y]}}.$$

D.3 Conditional Expectation and Variance

- *Conditional expectation*. It is defined as

$$E[X \mid y] \triangleq \int_x x f(x \mid y)\,\mathrm{d}x,\ \ E[Y \mid x] \triangleq \int_y y f(y \mid x)\,\mathrm{d}y,$$

and these are functions of y and x, respectively [765]. It follows that

$$E[X] = E_y E[X \mid y],\ \ E[Y] = E_x E[Y \mid x].$$

In these expressions, $E[X \mid y]$ and $E[Y \mid x]$ are considered as random variables and subscripts in E_x or E_y mean that expectation is taken with respect to x or y by using their respective marginals. The last two expressions are also known as *total expectations*.

- *Conditional variance*. It is defined as

$$Var[X \mid y] \triangleq \int_x \{x - E[X \mid y]\}^2 f(x \mid y)\,\mathrm{d}x$$

and

$$Var[Y \mid x] \triangleq \int_y \{y - E[Y \mid x]\}^2 f(y \mid x)\,\mathrm{d}y,$$

and they are functions of y and x, respectively [765]. It follows that

$$Var[X] = Var_y E[X \mid y] + E_y Var[X \mid y]$$

and

$$Var[Y] = Var_x E[Y \mid x] + E_x Var[Y \mid x]$$

As above, $Var[X \mid y]$ and $Var[Y \mid x]$ are considered as random variables and subscripts in Var_x or Var_y mean that variance is taken with respect to x or y using their respective marginals. The last two expressions are also known as *total variances*.

D.4 Generating Functions

Generating functions are coming into widespread use as methodological tools [552]. They may be used to obtain numerical summary measures of probability distributions in an analytical form by computing its moments and cumulants. For the nonnegative integer-valued random variable $X(t)$:

- The *probability generating function* $\mathcal{P}(s,t)$ is defined as

$$\mathcal{P}(s,t) \triangleq \sum s^x p_x(t),$$

where s is a "dummy variable" such that $|s| < 1$. It follows that one could obtain any probability, say $p_i(t)$, by differentiating $\mathcal{P}(s,t)$ with respect to s; specifically,

$$p_i(t) = \mathcal{P}^{(i)}(0,t),$$

where $\mathcal{P}^{(i)}(0,t)$ denotes the i-th derivative with respect to s evaluated at $s = 0$.

- The *moment generating function* $\mathcal{M}(\theta,t)$ is defined as

$$\mathcal{M}(\theta,t) \triangleq \sum \exp(\theta x) p_x(t),$$

where θ is a "dummy variable." Clearly, using the previous relation one has $\mathcal{M}(\theta,t) = \mathcal{P}(\exp(\theta), t)$. If $\mathcal{M}(\theta,t)$ is expressed as the power series

$$\mathcal{M}(\theta,t) = \sum_{i \geq 0} \frac{\mu_i(t)\,\theta^i}{i!},$$

the coefficients $\mu_i(t)$ in this series expansion are the i-th moments of $X(t)$, which are usually defined as $\mu_i(t) \triangleq \sum x^i p_x(t)$ with $\mu_0 = 1$. It follows that the i-th moment may be obtained from the moment generating function as

$$\mu_i(t) = \mathcal{M}^{(i)}(0,t).$$

- The *cumulant generating function* $\mathcal{K}(\theta, t)$ is defined as

$$\mathcal{K}(\theta, t) \triangleq \log \mathcal{M}(\theta, t),$$

with power series expansion

$$\mathcal{K}(\theta, t) = \sum_{i \geq 0} \frac{\kappa_i(t)\, \theta^i}{i!}.$$

This equation formally defines a cumulant $\kappa_i(t)$ as a coefficient in the series expansion of $\mathcal{K}(\theta, t)$. It too is easily found from its generating function as

$$\kappa_i(t) = \mathcal{K}^{(i)}(0, t).$$

The first three cumulants may be obtained as

$$\begin{aligned}
\kappa_1(t) &= \mu_1(t), \\
\kappa_2(t) &= \mu_2(t) - \mu_1^2(t), \\
\kappa_3(t) &= \mu_3(t) - 3\mu_1(t)\,\mu_2(t) + 2\mu_1^3(t),
\end{aligned}$$

which give the mean, variance, and skewness functions for $X(t)$ from the $\mu_i(t)$ moment functions.

Appendix E
Convolution in Probability Theory

A convolution is an integral that expresses the amount of overlap of one function g as it is shifted over another function f. It therefore "blends" one function with another. The convolution is sometimes also known by its German name, *Faltung* (*folding*). Abstractly, a convolution is defined as a product of functions f and g that are objects in the algebra of Schwartz functions in $\mathbb{R}^n$. Convolution of two functions $f(z)$ and $g(z)$ over a finite range $[0, t]$ is given by

$$f * g\,(t) \triangleq \int_0^t f\,(\tau - t)\, g\,(\tau)\, \mathrm{d}\tau,$$

where the symbol $f * g$ denotes convolution of f and g.

There is also a definition of the convolution that arises in probability theory and is given by

$$F * G\,(t) \triangleq \int_0^t F\,(t - z)\, \mathrm{d}G\,(z)\,,$$

where

$$\int_0^t F\,(t - z)\, \mathrm{d}G\,(z)$$

is a Stieltjes integral.

The Stieltjes integral is a generalization of the Riemann integral. Let $f(z)$ and $h\,(z)$ be real-valued bounded functions defined on a closed interval $[a, b]$. Take a partition of the interval $a = z_1 < z_2 < \ldots < z_{n-1} < z_n = b$ and consider the Riemann sum

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7

$$\sum_{i=1}^{n-1} f(\xi_i)\left[h(z_{i+1}) - h(z_i)\right]$$

with $\xi_i \in [z_i, z_{i+1}]$. If the sum tends to a fixed number $\mathcal{I}$ as $\max(z_{i+1} - z_i) \to 0$, then $\mathcal{I}$ is called the Stieltjes integral, or sometimes the Riemann–Stieltjes integral. The Stieltjes integral of f with respect to h is denoted by $\int f(z)\mathrm{d}h(z)$. If f and h have a common point of discontinuity, then the integral does not exist. However, if f is continuous and h' is Riemann integrable over the specified interval, then

$$\int f(z)\,\mathrm{d}h(z) = \int f(z)\,h'(z)\,\mathrm{d}z.$$

For enumeration of many of the Stieltjes integral properties, cf. [766] (p.105). In the following, we present some useful convolution relationships:

- $f * g(t) = g * f(t) \triangleq \int_0^t f(\tau - t)\,g(\tau)\mathrm{d}\tau = \int_0^t g(\tau - t) f(\tau)\mathrm{d}\tau$
- $f(t) * [k_1 g(t) + k_2 h(t)] = k_1 f * g(t) + k_2 f * h(t)$
- $f * g(t)|_{t=0} = 0$
- $\int_0^t f * g(\tau)\mathrm{d}\tau = f(t) * \int_0^t g(\tau)\mathrm{d}\tau = g(t) * \int_0^t f(\tau)\mathrm{d}\tau$
- $\int_0^\infty f * g(\tau)\mathrm{d}\tau = \left[\int_0^\infty f(\tau)\,\mathrm{d}\tau\right]\left[\int_0^\infty g(\tau)\,\mathrm{d}\tau\right]$
- $\frac{\mathrm{d}}{\mathrm{d}t}[f * g(t)] = f(t)\,g(0) + f * \frac{\mathrm{d}g(t)}{\mathrm{d}t} = g(t) f(0) + g * \frac{\mathrm{d}f(t)}{\mathrm{d}t}$
- $k * f(t) = k \int_0^t f(\tau)\mathrm{d}\tau$
- $\frac{\mathrm{d}}{\mathrm{d}t}[k * f(t)] = k f(t)$
- $\delta * f(t) = f(t)$

where k is a scalar constant and $\delta(t)$ is the Dirac delta function.

Appendix F
Laplace Transform

The Laplace transform $\tilde{f}(s)$ of the function $f(t)$ of the nonnegative variable t is defined by

$$\tilde{f}(s) \triangleq L\{f(t)\} = \int_0^\infty \exp(-st) f(t)\, \mathrm{d}t.$$

This transform is widely used to formulate semi-Markov stochastic models, where t and $f(t)$ are the random variable and its probability density function, respectively. In Table F.1, we briefly report some Laplace transform pairs.

The probability density functions Exp(κ), Erl(λ, ν), Gam(λ, μ), and Rec(α, β) are defined in Tables 11.1 and 11.2, and Chi(ν) is the χ^2 distribution with ν degrees of freedom. After modeling in frequency s-space, the solution in time t-space must be obtained by inverse Laplace transform. Nevertheless, given the complexity of the obtained model, the inverse transform may be rarely obtained from the above table. Usually, the numerical inverse Laplace transform is used [520, 527].

Table F.1 Some Laplace transform properties and pairs of functions used as probability density functions for semi-Markov modeling.

$f(t)$	$\tilde{f}(s)$
$\exp(-\alpha t) f(t)$	$\tilde{f}(s+\alpha)$
$f(t-\alpha) H(t-\alpha)$	$\exp(-\alpha s) \tilde{f}(s)$
$f(kt)$	$\tilde{f}(s/k)/k$
$f_1 * f_2(t)$	$\tilde{f}_1(s) \tilde{f}_2(s)$
Exp(κ)	$\kappa/(s+\kappa)$ with $s > -\kappa$
Erl(λ, ν)	$[1+(s/\lambda)]^{-\nu}$ with $s > -\lambda$
Chi(ν)	$(1+2s)^{-\nu/2}$ with $s > -\frac{1}{2}$
Gam(λ, μ)	$[1+(s/\lambda)]^{-\mu}$ with $s > -\lambda$
Rec(α, β)	$\{\exp(-\alpha s) - \exp[-(\alpha+\beta)s]\}/(\beta s)$

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7

Appendix G
Theorems

G.1 Continuous Functions

Lemma 13. *Let $h(z)$ be a derivable function of z over $[a, b]$ satisfying $h(z) \neq 0$ for all $z \in]a, b[$, $h(a) = h(b) = 0$, and $h'(a) < 0$. Then, for all $z \in]a, b[$, $h(z) < 0$ and $h'(b) \geq 0$.*

When the derivative $h'(a)$ is approximated by the quotient difference, we have

$$h'(a) \approx \frac{h(z) - h(a)}{z - a} = \frac{h(z)}{z - a} < 0,$$

and therefore, using continuity, $h(z) < 0$ for $z \in]a, a + \Delta a]$. Since $h(z)$ is continuous over $]a, b[$ and $h(z) \neq 0$, $h(z)$ preserves its sign for $z \in]a, b[$; consequently, $h(z) < 0$ for $z \in]a, b[$. Conversely, for all $z \in]a, b[$, we have

$$\frac{h(z) - h(b)}{z - b} = \frac{h(z)}{z - b} > 0$$

and

$$h'(b) = \lim_{z \to b} \frac{h(z) - h(b)}{z - b} \geq 0$$

■

Proposition 14. *Let $f(z)$ and $g(z)$ be derivable functions of z over $[a, b]$ satisfying: $g(z)$ is a monotone increasing function over $[a, b]$, $f(z) \neq g(z)$ for all $z \in]a, b[$, $f(a) = g(a)$ and $f(b) = g(b)$, and $f'(a) < 0$. Then $f(z) < g(z)$ and $f'(b) \geq 0$.*

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7

Let $h(z) = f(z) - g(z)$. Since $f'(a) < 0$ and $g'(a) \geq 0$ (g is monotone increasing), $h'(a) = f'(a) - g'(a) < 0$. Because of $f(z) \neq g(z)$, $h(z) \neq 0$ for all $z \in]a, b[$. According to the previous lemma, it follows that $h(z) < 0$, i.e., $f(z) < g(z)$ for all $z \in]a, b[$, and $h'(b) = f'(b) - g'(b) \geq 0$. Since $g'(b) \geq 0$ (g is monotone increasing), we have also $f'(b) \geq 0$. ■

A similar proof may be delineated for the dual proposition:

Proposition 15. *Let $f(z)$ and $g(z)$ be derivable functions of z over $[a, b]$ satisfying: $g(z)$ is a monotone increasing function over $[a, b]$, $f(z) \neq g(z)$ for all $z \in]a, b[$, $f(a) = g(a)$ and $f(b) = g(b)$, and $f'(a) > 0$. Then $f(z) > g(z)$ and $f'(b) \leq 0$.*

From the last two propositions, we can state the following result:

Theorem 16. *Let $f(z)$ and $g(z)$ be derivable functions of z over the interval I and let $g(z)$ be a monotone increasing function. Let also $a_1 < a_2 < \ldots < a_n$ be n reals over I satisfying: $f(a_i) = g(a_i)$ for $i = 1 : n$ and $f(z) \neq g(z)$ for all $z \in]a_i, a_{i+1}[$ with $i = 1 : n - 1$. Then $f(a_i)g(a_i) \leq 0$ for $i = 1 : n$. In other words, the derivatives on two successive intersection points between two continuous functions, one of which is monotone, have opposite signs.*

Simply, apply the previous propositions to the segment $[a_i, a_{i+1}]$. ■

G.2 Matrix Operations and Eigenvalues

Proposition 17. *Given a square matrix A with eigenvalues ζ_i, the eigenvalues $\overline{\zeta}_i$ of sA (with s scalar) are $\overline{\zeta}_i = s\lambda_i$ and the eigenvalues $\tilde{\zeta}_i$ of $\exp(A)$ are $\tilde{\zeta}_i = \exp(\zeta_i)$. The eigenvectors remain the same under the above matrix operations.*

The above results from the definition of eigenvalues and eigenvectors, and the Jordan factorization $A = GZG^{-1}$ leading to $\exp(A) = G \exp(Z) G^{-1}$. Here, G and Z denote the matrix of eigenvectors and the diagonal matrix of eigenvalues of A, respectively. ■

G.3 Matrix Inversion Lemma

Proposition 18. *The Matrix Inversion Lemma is the equation*

$$\left(A - BD^{-1}C\right)^{-1} = A^{-1} + A^{-1}B\left(D - CA^{-1}B\right)^{-1} CA^{-1}$$

where A and C are square invertible matrices, and B, D are matrices so that A and BCD have the same dimensions.

Appendix H
List of Symbols

The symbols in the following tables are classified in several lists according to their significance and form: symbols associated with functions and distributions (Table H.1), time-dependent variables (Table H.2), random variables (Table H.3), constants and parameters (Tables H.4, H.5, H.6), and Greek symbols (Table H.7). The list of tables is completed by a table of abbreviations (Table H.8).

In order to respect the initial writing in the literature of symbols, sometimes but in a different place the same symbol has been used for more than one purpose. For example, $s\,(t)$ denotes the substrate variable in Chapters 8 and 11, whereas it refers to the neutrophil myelocytes in Chapter 13. In such cases, we systematically report as reference for each use the number of the corresponding chapter. For random variables, a pair of symbols is used with the same character in uppercase and lowercase form to denote the name of a random variable and an element of that variable, respectively. For instance, A denotes the random variable "age" and a a given age. Underscored lowercase characters and bold uppercase denote vectors and matrices, respectively, e.g., $\underline{y}$ and $\mathbf{H}$. Usually, Greek letters κ, λ, μ, ν stand for the parameters of statistical distributions, and α, β, and γ are used as unspecified constants or parameters.

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7

Table H.1 Functions and distributions.

SYMBOL	LABEL
$B(t)$	Brownian motion
$\xi(t)$	Gaussian white noise, Chapt. 5
$\varphi(t)$	Fraction of dose dissolved, Chapts. 5, 6
$c_B(y)$	Binding curve
$c_F(y)$	Feedback curve
$\delta(\cdot)$	Dirac delta function
$f(a)$	Density function
$F(a)$	Distribution function
$\mathcal{F}(\cdot,\cdot)$	Transfer function
$\Phi(\cdot)$	Feedback control function
$\phi(\cdot)$	Dimensionless feedback function
$g(\cdot)$	Functional form
$I(\cdot)$	Intensity function
$J_0(\cdot)$	Zero-order Bessel function
$\mathcal{K}(\cdot,\cdot)$	Cumulant generating function
$\mathcal{M}(\cdot,\cdot)$	Moment generating function
$\mathcal{P}(\cdot,\cdot)$	Probability generating function
$\mathcal{S}(a)$	Survival function
$\mathcal{T}(\cdot)$	Transducer function
$u(t)$	Input function
$\rho(t)$	Dimensionless input function
$H(\cdot)$	Heaviside step function
$\Gamma(\cdot)$	Gamma function
$\psi(t)$	History function, Chapt. 13
$\mathrm{Bin}(\cdot,\cdot)$	Binomial distribution
$\mathrm{Chi}(\cdot)$	χ^2 distribution
$\mathrm{Erl}(\cdot,\cdot)$	Erlang distribution
$\mathrm{Exp}(\cdot,\cdot)$	Exponential distribution
$\mathrm{Gam}(\cdot,\cdot)$	Gamma distribution
$\mathrm{Rec}(\cdot,\cdot)$	Rectangular (uniform) distribution
$\mathrm{Wei}(\cdot,\cdot)$	Weibull distribution
$E[\cdot]$	Expectation
$Var[\cdot]$	Variance
$Cov[\cdot,\cdot]$	Covariance
$Cor[\cdot,\cdot]$	Correlation

Table H.2 Time-dependent variables.

SYMBOL	LABEL
$c(t)$	Drug concentration, Chapts. 2, 4, 6, 7, 8, 10, 12, 13
$e(t)$	Enzyme, Chapts. 8, 11
	Cytokines, Chapt. 13
$E(t)$	Pharmacological effect, Chapts. 12, 13
$q(t)$	Drug amount, Chapts. 2, 4, 5, 6, 7, 8, 12
$s(t)$	Substrate, Chapts. 8, 11
	Neutrophil myelocytes, Chapt. 13
$\upsilon(t)$	Enzyme–substrate complex, Chapt. 8
	Drug–receptor complex, Chapts. 12, 13
$w(t)$	Product of enzymatic reaction, Chapt. 8
	Blood neutrophils, Chapt. 13
$x(\tau), y(\tau), z(\tau)$	Dimensionless state variables
$y(t)$	State variable

Table H.3 Random variables.

SYMBOL	LABEL	DIMENSION
a, A	Age, Chapt. 11	Time
$c(t), C(t)$	Concentration, Chapt. 11	$\text{Mass}\times\text{Volume}^{-1}$
$n(t), N(t)$	No. of particles, Chapt. 11	
$q(t), Q(t)$	Amount, quantity, Chapt. 11	Mass
t, T	Time, Chapt. 5	Time
θ, Θ	Characteristic parameter, Chapts. 7, 8	

Table H.4 Constants, parameters (part 1). [apu] denotes arbitrary pharmacological units.

SYMBOL	LABEL	DIMENSION
$\mathcal{A}$	Area	Area
A_n	Absorption number	
AUC	Area under curve	Mass×Volume^{-1}×Time
$B_{\cdot\cdot}$	Coefficients in a sum of exponentials	Mass×Volume^{-1}
$b_{\cdot}$	Exponents in a sum of exponentials	Time^{-1}
$\mathbf{B}_{\cdot}, \underline{b}_{\cdot}$	Parameters in pseudocompartments	
c_s	Solubility	Mass×Volume^{-1}
c_0	Initial concentration	Mass×Volume^{-1}
$c_{\max}$	Peak drug concentration	Mass×Volume^{-1}
CL	Clearance	Volume×Time^{-1}
CV	Coefficient of variation	
d_t	Topological dimension	
d_f	Fractal dimension	
d_c	Capacity dimension	
d_e	Embedding dimension	
d_s	Spectral dimension	
d_w	Random walk dimension	
d_o	Cover dimension	
d^*	Index of nonlinear coherence	
$\mathcal{D}$	Diffusion coefficient	Area×Time^{-1}
$\mathcal{D}'$	Modified diffusion coefficient	Area×Time^{-1}
$\mathcal{D}_\gamma$	Fractional diffusion coefficient	Area×Time^{-1}
D	Dispersion coefficient	Area×Time^{-1}
D_n	Dissolution number	
e_0	Initial enzyme amount, Chapts. 8, 11	Mass
	Initial cytokine amount, Chapt. 13	Mass
E_0	Baseline in $E_{\max}$ model	[apu]
$E_{\max}$	Maximum pharmacological effect	[apu]
Ec_{50}	Concentration at half $E_{\max}$	Mass×Volume^{-1}
f_1	Difference factor	
f_2	Similarity factor	
f_u	Drug unbound fraction	
f_{un}	Fraction of unionized species	
F_a	Fraction of dose absorbed	
$h_{\cdot\cdot}$	Hazard rates	Time^{-1}
$\mathcal{H}$	Hypothesis for statistical test	
$\mathbf{H}$	Transfer intensity matrix	
$I_{\max}$	Maximum inhibition rate	
Ic_{50}	Concentration at half $I_{\max}$	Mass×Volume^{-1}
J	Net flux	Mass×Area^{-1}×Time^{-1}

Table H.5 Constants, parameters (part 2). [apu] denotes arbitrary pharmacological units.

SYMBOL	LABEL	DIMENSION
k	First-order rate const. (generic)	Time^{-1}
$k_\circ$	Reference rate const.	Time^{-1}
k_0	Case-II relaxation const.	$\text{Mass}\times\text{Area}^{-1}\times\text{Time}^{-1}$
k_{+1}	Forward enz. reaction rate const.	$\text{Mass}^{-1}\times\text{Time}^{-1}$
k_{-1}	Backward enz. reaction rate const.	Time^{-1}
k_{+2}	Enzymatic product formation rate const.	Time^{-1}
k_2	Pharmacological proport. const.	$[\text{apu}]\times\text{Mass}^{-1}\times\text{Volume}$
k_a	Macroscopic absorption rate const.	Time^{-1}
k'_a	Microscopic absorption rate const.	Time^{-1}
k_d	Dissolution rate const.	Time^{-1}
$k_{d,eff}$	Effective dissolution rate const.	Time^{-1}
k_c	Controlled dissolution rate const., Chapt. 6	Time^{-1}
k_y	Effect-compartment rate const., Chapt. 12	Time^{-1}
k_s	Surface area dissolution rate const.	
k_D	Dissociation const.	Mass
k_i, k_o	Input (orders 0 and 1), output rate const.	
k_e, k_s, k_w	Rate const. in hemopoiesis	
k_M	Michaelis–Menten const.	Mass
$k..$	Fractional flow rates	Time^{-1}
$\mathbf{K}$	Matrix of fractional flow rates	
L	Height of cylinder	Length
m	No. of objects (except particles): samples, compartments, individuals, vessels, administrations, sites	
$m_\circ$	No. of reaction channels	
n_a, n_b	No. of delayed outputs, inputs	
n_0	Initial no. of particles	
$n(t)$	No. of visited sites, Chapt. 2	
	No. of remained particles, Chapt. 4	
$\tilde{n}(t)$	No. of escaped particles	
N_{leak}	No. of leak sites	
N_{tot}	Total no. of sites	
N_{villi}	No. of villi	
p	Probability	
p_a	Probability for absorption by villi	
p_f	Forward probability to the output	
p_c	Critical probability	

Table H.6 Constants, parameters (part 3).

SYMBOL	LABEL	DIMENSION
P	Permeability	Length×Time^{-1}
P_{app}	Apparent permeability	Length×Time^{-1}
P_{eff}	Effective permeability	Length×Time^{-1}
P_c	Partition coefficient	
$\mathbf{P}$	State probability matrix	
q	the shift operator	
q_0	Initial drug amount (dose)	Mass
q_∞	Maximum cumulative amount	Mass
Q	Volumetric flow rate	Volume×Time^{-1}
R	Radius of tube	Length
$R_{..}$	Transfer rate	Mass×Time^{-1}
R^2	Coefficient of determination	
$R_{\max}$	Maximum biotransformation rate	Mass×Time^{-1}
r	Resolution, Chapt. 1	
	Random number, Chapt. 11	
r_0	Total no. of receptors	
R_i	Reference drug dissolved at i	Mass
$S^2_{..}$	Variance in the sample, Chapt. 10	
$S_{\max}$	Maximum stimulation rate	
Sc_{50}	Concentration at half $S_{\max}$	Mass×Volume^{-1}
t	Time	Time
t°	Time delay	Time
$t_\circ$	Time reference	Time
t_0	Initial time	Time
t_{sim}	Maximum simulation time	Time
t_{diff}	Diffusion time	Time
t_{reac}	Reaction time	Time
$t_{\max}$	Time to $c_{\max}$	Time
$t_{..}$	Observation times	Time
T	Infusion duration	Time
T_E	Infusion ending time	Time
T_S	Infusion starting time	Time
T_{si}	Small-intestinal transit time	Time
T_i	Test drug dissolved at i	Mass
u_i	Unit geometric vector in direction i	
v	Velocity	Length×Time^{-1}
V	Volume of distribution	Volume
V_c	Central compartment volume	Volume
V_y	Effect-compartment volume	Volume
$V_{\max}$	Maximum transport rate	Mass×Time^{-1}
$y_{..}$	Observations	Mass×Volume^{-1}
z	Spatial coordinates	Length

Table H.7 Greek symbols. [apu] denotes arbitrary pharmacological units.

SYMBOL	LABEL	DIMENSION
α, β, γ	Constants, parameters (usually)	
α	Significance level for statistical tests	
δ	Thickness, elementary distance	Length
Δ	Finite difference symbol	
ε	Intrinsic efficacy, Chapt. 12 Bias factor, Chapt. 6	[apu]×Mass^{-1}×Volume
	Subinterval length, Chapt. 11	Length
$\varepsilon_{..}$	Prediction error	
ζ	Eigenvalue	
η	Eigenvector	
θ	Characteristic parameter, Chapts. 1, 3, 11, 13 Solubility–dose ratio, Chapts. 5, 6	
$\kappa, \lambda, \mu, \nu$	Distribution parameters (usually)	
$\kappa_{.}$	Cumulants	
$\mu_{.}$	Moments	
ξ	Characteristic multiplier, Chapt. 3 Correlation length, Chapt. 1	
ϱ	Density of the drug	Mass×Volume^{-1}
ρ	Radius of particle	Length
ρ_c	Critical radius	Length
ρ_0	Initial radius	Length
σ^2	Variance, Chapts. 5, 8	
τ	Dimensionless time	
τ°	Dimensionless time delay	
$\underline{\varphi}$	Regression vector, Chapt. 7	
$\varphi_{..}$	Changes in population size by the reaction, Chapt. 11	
$\psi_{..}$	No. of particles implied in the reaction, Chapt. 11	
$\underline{\Psi}$	Vector of population parameters	
ω	Resolution, Chapt. 1	
$\omega_{..}$	Transition probabilities, Chapt. 11	

Table H.8 Abbreviations.

ABBREVIATION	EXPANSION
ABΓ	ABΓ system
ACAT	Advanced version of CAT
ARX	Auto-Regression with eXtra inputs (ARX) model
BCS	Biopharmaceutics Classification System
BDDCS	Biopharmaceutics Drug Disposition Classification System
CAT	Compartmental Absorption Transit
DLA	Diffusion-Limited Aggregation
DTPA	diethyltriamino-pentaacetic acid
EMA	European Medicines Agency
FDA	Food and Drug Administration
FDE	Fractional order ordinary Differential Equations
G-CSF	colony-stimulating factor
HPMC	hydroxypropyl methylcellulose
LTI	Linear Time-Invariant models
MCS	Monte Carlo microSteps
MSE	Mean Square Error
NONMEM	NONlinear Mixed Effect Modeling
NPEM2	NonParametric Expectation Maximization
NSAID	nonsteroidal anti-inflammatory drugs
NILT	Numerical Inverse Laplace Transform algorithm
PLGA	poly-lactic-co-glycolic acid
TOST	Two One-Sided Test procedure
TPGS	d-alpha-tocopheryl polyethylene glycol 1000 succinate
QBCS	Quantitative Biopharmaceutics Classification System
WHO	World Health Organization

Bibliography

1. Kac, M.: Some mathematical models in science. Science **166**, 695–699 (1969)
2. Mandelbrot, B.: Fractals: Form, Chance, and Dimension. W.H. Freeman, San Francisco (1977)
3. Parker, T., Chua, L.: Practical Numerical Algorithms for Chaotic Systems. Springer, New York (1989)
4. Bassingthwaighte, J., Liebovitch, L., West, B.: Fractal Physiology. Methods in Physiology Series. Oxford University Press, New York (1994)
5. Flory, P.: Molecular size distribution in three dimensional polymers. I. Gelation. J. Am. Chem. Soc. **63**(11), 3083–3090 (1941)
6. Flory, P.: Molecular size distribution in three dimensional polymers. II. Trifunctional branching units. J. Am. Chem. Soc. **63**(11), 3091–3096 (1941)
7. Flory, P.: Molecular size distribution in three dimensional polymers. III. Tetrafunctional branching units. J. Am. Chem. Soc. **63**(11), 3096–3100 (1941)
8. Stockmayer, W.: Theory of molecular size distribution and gel formation in branched-chain polymers. J. Chem. Phys. **11**, 45–55 (1943)
9. Kopelman, R.: Rate processes on fractals: theory, simulations, and experiments. J. Stat. Phys. **42**(1/2), 185–200 (1986)
10. Jacobs, M.: Diffusion Processes. Springer, Berlin (1967)
11. Pippa, N., Dokoumetzidis, A., Demetzos, C., Macheras, P.: On the ubiquitous presence of fractals and fractal concepts in pharmaceutical sciences: a review. Int. J. Pharm. **456**(2), 340–352 (2013)
12. Ovchinnikov, A., Zeldovich, Y.: Role of density fluctuations in bimolecular reaction kinetics. Chem. Phys. **28**, 215–218 (1978)
13. Toussaint, D., Wilczek, F.: Particle-antiparticle annihilation in diffusive motion. J. Chem. Phys. **78**(5), 2642–2647 (1983)
14. Lindenberg, K., West, B., Kopelman, R.: Steady-state segregation in diffusion limited reactions. Phys. Rev. Lett. **60**(18), 1777–1780 (1988)
15. Wenzel, C.: Lehre von der Verwandtschaft der Korper. Dresden 1777
16. Farin, D., Avnir, D.: Reactive fractal surfaces. J. Phys. Chem. **91**(22), 5517–5521 (1987)
17. Kopelman, R.: Fractal reaction kinetics. Science **241**, 1620–1626 (1988)
18. Clement, E., Kopelman, R., Sander, L.: The diffusion-limited reaction A+B –>0 on a fractal substrate. J. Stat. Phys. **65**(5–6), 919–924 (1991)
19. Klymko, P., Kopelman, R.: Fractal reaction-kinetics: exciton fusion on clusters. J. Phys. Chem. **87**(23), 4565–4567 (1983)

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7

20. Alexander, S., Orbach, R.: Density of states on fractals: Fractons. J. Phys. Lett. **43**, L625–631 (1982)
21. Zumofen, G., Blumen, A., Klafter, J.: Concentration fluctuations in reaction kinetics. J. Chem. Phys. **82**(7), 3198–3206 (1985)
22. Newhouse, J., Kopelman, R.: Steady-state chemical kinetics on surface clusters and islands: segregation of reactants. J. Phys. Chem. **92**(6), 1538–1541 (1988)
23. Jiang, Z., Ebner, C.: Simulation study of reaction fronts. Phys. Rev. A 42(12), 7483–7486 (1990)
24. DeGennes, P.: Kinetics of diffusion-controlled processes in dense polymer systems I. Nonentangled regimes. J. Chem. Phys. **76**(6), 3316–3321 (1982)
25. DeGennes, P.: Kinetics of diffusion-controlled processes in dense polymer systems II. Effects of entanglements. J. Chem. Phys. **76**(6), 3322–3326 (1982)
26. Savageau, M.: Michaelis-Menten mechanism reconsidered: implication of fractal kinetics. J. Theor. Biol. **176**, 115–124 (1995)
27. Srere, P., Jones, M., Mathews, C.: Structural and Organizational Aspects of Metabolic Regulation. Wiley-Liss, New York (1990)
28. Savageau, M.: Development of fractal kinetic theory for enzyme-catalysed reactions and implications for the design of biochemical pathways. Biosystems **47**, 9–36 (1998)
29. Schroeder, M.: Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise. W.H. Freeman, New York (1991)
30. Sokolov, I., Klafter, J., Blumen, A.: Fractional kinetics. Phys. Today **55**, 48–54 (2002)
31. Berding, C.: On the heterogeneity of reaction-diffusion generated patterns. Bull. Math. Biol. **49**(2), 233–252 (1987)
32. Glass, L., Mackey, M.: From Clock to Chaos. The Rhythms of Life. Princeton University Press, Princeton (1988)
33. Tsonis, A.: Chaos: From Theory to Applications. Kluwer Academic Publishers, Amsterdam (1992)
34. Dokoumetzidis, A., Iliadis, A., Macheras, P.: Nonlinear dynamics and chaos theory: concepts and applications relevant to pharmacodynamics. Pharm. Res. **18**(4), 415–426 (2001)
35. Farmer, J., Ott, E., Yorke, J.: The dimension of chaotic attractors. Phys. D **7**(1–3), 153–180 (1983)
36. Packard, N., Crutchfield, J., Farmer, J., Shaw, R.: Geometry from a time series. Phys. Rev. Lett. **45**(9), 712–716 (1980)
37. Froehling, H., Crutchfield, J., Farmer, D., Packard, N., Shaw, R.: On determining the dimension of chaotic flows. Physica D **3**(3), 605–617 (1981)
38. Nayfeh, A., Balachandran, B.: Applied Nonlinear Dynamics: Analytical, Computational, and Experimental Methods. Wiley, New York (1995)
39. Baake, E., Baake, M., Bock, H., Briggs, K.: Fitting ordinary differential equations to chaotic data. Phys. Rev. A **45**(8), 5524–5529 (1992)
40. Grassberger, P., Procaccia, I.: Characterization of strange attractors. Phys. Rev. Lett. **50**(5), 346–349 (1983)
41. Thompson, J., Stewart, H.: Nonlinear Dymanics and Chaos. Wiley, New York (1986)
42. Albano, A., Mess, A., DeGuzman, G., Rapp, P.: Data requirements for reliable estimation of correlation dimensions. In: Degn, H., Holden, A., Olsen, L. (eds.) Chaos in Biological Systems, pp. 207–220. Plenum, New York (1987)
43. Kaplan, D., Glass, L.: Direct test for determinism in a time series. Phys. Rev. Lett. **68**(4), 427–430 (1992)
44. Kaplan, D., Glass, L.: Understanding Nonlinear Dynamics. Textbooks in Mathematical Sciences. Springer, Berlin (1995)
45. Lorenz, E.: Deterministic nonperiodic flow. J. Atmos. Sci. **20**(2), 130–148 (1963)
46. Shinbrot, T., Grebogi, C., Ott, E., Yorke, J.: Using small perturbations to control chaos. Nature **363**, 411–417 (1993)
47. Kapitaniak, T.: Controlling Chaos: Theoretical and Practical Methods in Non-linear Dynamics. Academic, London (1996)

48. Mackey, M., Glass, L.: Oscillation and chaos in physiological control systems. Science **197**(4300), 287–289 (1977)
49. Glass, L., Mackey, M.: Pathological conditions resulting from instabilities in physiological control systems. Ann. N. Y. Acad. Sci. **316**, 214–235 (1979)
50. Mackey, M., Milton, J.: Dynamical diseases. Ann. N. Y. Acad. Sci. **504**, 16–32 (1987)
51. Skinner, J.: Low-dimensional chaos in biological systems. Biotechnology **12**(6), 596–600 (1994)
52. Dighe, S., Development of dissolution tests for immediate release oral dosage forms. In: Blume, H., Midha, K.K. (eds.) Bio-lnternational 2, Bioavailabilily, Bioequivalence and Pharmacokinetic Studies, pp. 247–255. Medpharm, Stuttgart (1995)
53. Fagerholm, U., Lennernas, H.: Experimental estimation of the effective unstirred water layer thickness in the human jejunum, and its importance in oral drug absorption. Eur. J. Pharm. Sci. **3**(5), 247–253 (1995)
54. Dressman, J., Fleisher, D.: Mixing tank model for predicting dissolution rate control of oral absorption. J. Pharm. Sci. **75**(2), 109–116 (1986)
55. Sinko, P., Leesman, G.D., Amidon, G.L.: Predicting fraction dose absorbed in humans using a macroscopic mass balance approach. Pharm. Res. **8**(8), 979–988 (1991)
56. Oh, D., Curl, R., Amidon, G.: Estimating the fraction dose absorbed from suspensions of poorly soluble compounds in humans: a mathematical model. Pharm. Res. **10**(2), 264–270 (1993)
57. Higuchi, T.: Rate of release of medicaments from ointment bases containing drugs in suspension. J. Pharm. Sci. **50**, 874–875 (1961)
58. Higuchi, T.: Mechanism of sustained-action medication. Theoretical analysis of rate of release of solid drugs dispersed in solid satrices. J. Pharm. Sci. **52**, 1145–1149 (1963)
59. Lapidus, H., Lordi, N.: Some factors affecting the release of a water-soluble drug from a compressed hydrophilic matrix. J. Pharm. Sci. **55**(8), 840–843 (1966)
60. Lapidus, H., Lordi, N.: Drug release from compressed hydrophilic matrices. J. Pharm. Sci. **57**(8), 1292–1301 (1968)
61. Siepmann, J., Peppas, N.: Modeling of drug release from delivery systems based on hydroxypropyl methylcellulose (HPMC). Adv. Drug Deliv. Rev. **48**(2–3), 139–157 (2001)
62. Crank, J.: The Mathematics of Diffusion, 2nd edn. Oxford University Press, Oxford (1975)
63. Wang, T., Kwei, T., Frisch, H.: Diffusion in glassy polymers. III. J. Polym. Sci. **7**(12), 2019–2028 (1969)
64. Enscore, D., Hopfenberg, H., Stannett, V.: Effect of particle size on the mechanism controlling n-hexane sorption in glassy polystyrene microspheres. Polymer **18**(8), 793–800 (1977)
65. Frisch, H.: Sorption and transport in glassy polymers - a review. Polym. Eng. Sci. **20**, 2–13 (1980)
66. Peppas, N.: Analysis of Fickian and non-Fickian drug release from polymers. Pharm. Acta Helv. **60**(4), 110–111 (1985)
67. Ritger, P., Peppas, N.: A simple equation for description of solute release I. Fickian and non-Fickian release from non-swellable devices in the form of slabs, spheres, cylinders or discs. J. Control. Release **5**(1), 23–36 (1987)
68. Ritger, P., Peppas, N.: A simple equation for description of solute release II. Fickian and anomalous release from swellable devices. J. Control. Release **5**(1), 37–42 (1987)
69. Kosmidis, K., Rinaki, E., Argyrakis, P., Macheras, P.: Analysis of Case II drug transport with radial and axial release from cylinders. Int. J. Pharm. **254**(2), 183–188 (2003)
70. Peppas, N., Korsmeyer, R.: Dynamically swelling hydrogels in controlled applications. In: Peppas, N. (ed.) Hydrogels in Medicine and Pharmacy. Properties and Applications, vol. 3, pp. 109–136. CRC, Boca Raton (1987)
71. Rinaki, E., Valsami, G., Macheras, P.: The power law can describe the 'entire' drug release curve from HPMC-based matrix tablets: a hypothesis. Int. J. Pharm. **255**(1–2), 199–207 (2003)
72. Fu, J., Hagemeir, C., Moyer, D.: A unified mathematical model for diffusion from drug-polymer composite tablets. J. Biomed. Mater. Res. **10**(5), 743–758 (1976)

73. Gao, P., Fagerness, P.: Diffusion in HPMC gels. I. Determination of drug and water diffusivity by pulsed-field-gradient spin-echo NMR. Pharm. Res. **12**(7), 955–964 (1995)
74. Gao, P., Nixon, P., Skoug, J.: Diffusion in HPMC gels. II. Prediction of drug release rates from hydrophilic matrix extended-release dosage forms. Pharm. Res. **12**(7), 965–971 (1995)
75. Narasimhan, B., Peppas, N.: Disentanglement and reptation during dissolution of rubbery polymers. J. Polym. Sci. **34**(5), 947–961 (1996)
76. Narasimhan, B., Peppas, N.: On the importance of chain reptation in models of dissolution of glassy polymers. Macromolecules **29**(9), 3283–3291 (1996)
77. Narasimhan, B., Peppas, N.: Molecular analysis of drug delivery systems controlled by dissolution of the polymer carrier. J. Pharm. Sci. **86**(3), 297–304 (1997)
78. Siepmann, J., Kranz, H., Bodmeier, R., Peppas, N.: HPMC-matrices for controlled drug delivery: a new model combining diffusion, swelling, and dissolution mechanisms and predicting the release kinetics. Pharm. Res. **16**(11), 1748–1756 (1999)
79. Siepmann, J., Podual, K., Sriwongjanya, M., Peppas, N., Bodmeier, R.: A new model describing the swelling and drug release kinetics from hydroxypropyl methylcellulose tablets. J. Pharm. Sci. **88**(1), 65–72 (1999)
80. Siepmann, J., Kranz, H., Peppas, N., Bodmeier, R.: Calculation of the required size and shape of hydroxypropyl methylcellulose matrices to achieve desired drug release profiles. Int. J. Pharm. **201**(2), 151–164 (2000)
81. Siepmann, J., Peppas, N.: Hydrophilic matrices for controlled drug delivery: an improved mathematical model to predict the resulting drug release kinetics (the 'sequential layer' model). Pharm. Res. **17**(10), 1290–1298 (2000)
82. Streubel, A., Siepmann, J., Peppas, N., Bodmeier, R.: Bimodal drug release achieved with multi-layer matrix tablets: transport mechanisms and device design. J. Control. Release **69**(3), 455–468 (2000)
83. Siepmann, J., Streubel, A., Peppas, N.: Understanding and predicting drug delivery from hydrophilic matrix tablets using the 'sequential layer' model. Pharm. Res. **19**(3), 306–314 (2002)
84. Kosmidis, K., Argyrakis, P., Macheras, P.: A reappraisal of drug release laws using Monte Carlo simulations: the prevalence of the Weibull function. Pharm. Res. **20**(7), 988–995 (2003)
85. Bonny, J., Leuenberger, H.: Matrix type controlled release systems: I. Effect of percolation on drug dissolution kinetics. Pharm. Acta Helv. **66**(5–6), 160–164 (1991)
86. Bunde, A., Havlin, S., Nossal, R., Stanley, H., Weiss, G., On controlled diffusion-limited drug release from a leaky matrix. J. Chem. Phys. **83**(11), 5909–5913 (1985)
87. Pham, A., Lee, P.: Probing the mechanisms of drug release from hydroxypropylmethyl cellulose matrices. Pharm. Res. **11**(10), 1379–1384 (1994)
88. Bettini, R., Catellani, P., Santi, P., Massimo, G., Peppas, N., Colombo, P.: Translocation of drug particles in HPMC matrix gel layer: effect of drug solubility and influence on release rate. J. Control. Release **70**(3), 383–391 (2001)
89. Kosmidis, K., Argyrakis, P., Macheras, P.: Fractal kinetics in drug release from finite fractal matrices. J. Chem. Phys. **119**(12), 6373–6377 (2003)
90. Weiss, G.: Aspects and Applications of the Random Walk. Random Materials and Processes. North-Holland, Amsterdam (1994)
91. Papadopoulou, V., Kosmidis, K., Vlachou, M., Macheras, P.: On the use of the Weibull function for the discernment of drug release mechanisms. Int. J. Pharm. **309**(1–2), 44–50 (2006)
92. Rinaki, E., Dokoumetzidis, A., Macheras, P.: The mean dissolution time depends on the dose/solubility ratio. Pharm. Res. **20**(3), 406–408 (2003)
93. Siepmann, J., Faisant, N., Benoit, J.: A new mathematical model quantifying drug release from bioerodible microparticles using Monte Carlo simulations. Pharm. Res. **19**(12), 1885–1893 (2002)
94. Siepmann, J., Faisant, N., Akiki, J., Richard, J., Benoit, J.: Effect of the size of biodegradable microparticles on drug release: experiment and theory. J. Control. Release **96**(1), 123–134 (2004)

95. Mazer, N.: Pharmacokinetic and pharmacodynamic aspects of polypeptide delivery. J. Control. Release **11**(1–3), 343–356 (1990)
96. Heller, J.: Modulated release from drug delivery devices. Crit. Rev. Ther. Drug Carrier Syst. **10**(3), 253–305 (1993)
97. Siegel, R., Ziaie, B.: Biosensing and drug delivery at the microscale. Adv. Drug Deliv. Rev. **56**(2), 121–123 (2004)
98. Misra, G., Siegel, R.: Ionizable drugs and pH oscillators: buffering effects. J. Pharm. Sci. **91**(9), 2003–2015 (2002)
99. Leroux, J., Siegel, R.: Autonomous gel/enzyme oscillator fueled by glucose: preliminary evidence for oscillations. Chaos **9**(2), 267–275 (1999)
100. Giannos, S., Dinh, S., Berner, B.: Temporally controlled drug delivery systems: coupling of pH oscillators with membrane diffusion. J. Pharm. Sci. **84**(5), 539–543 (1995)
101. Misra, G., Siegel, R.: Multipulse drug permeation across a membrane driven by a chemical pH-oscillator. J. Control. Release **79**(1–3), 293–297 (2002)
102. Misra, G., Siegel, R.: New mode of drug delivery: long term autonomous rhythmic hormone release across a hydrogel membrane. J. Control. Release **81**(1–2), 1–6 (2002)
103. Li, B., Siegel, R.: Global analysis of a model pulsing drug delivery oscillator based on chemomechanical feedback with hysteresis. Chaos **10**(3), 682–690 (2000)
104. Levich, V.: Physicochemical Hydrodynamics. Prentice Hall, Englewood Cliffs (1962)
105. Noyes, A., Whitney, W.: The rate of solution of solid substances in their own solutions. J. Am. Chem. Soc. **19**, 930–934 (1897)
106. Nernst, W., Brunner, E.: Theorie der Reaktionsgeschwindigkeit in heterogen Systemen. Z. Phys. Chem. **47**, 52–110 (1904)
107. Hixson, A., Crowell, J.: Dependence of reaction velocity upon surface and agitation. I. Theoretical considerations. Ind. Eng. Chem. Res. **51**, 923–931 (1931)
108. Avnir, D., Farin, D., Pfeifer, P.: Surface geometric irregularity of particulate materials. J. Colloid Interface Sci., **103**, 112–123 (1985)
109. Li, T., Park, K.: Fractal analysis of pharmaceutical particles by atomic force microscopy. Pharm. Res. **15**(8), 1222–1232 (1998)
110. Schroder, M., Kleinebudde, P.: Structure of disintegrating pellets with regard to fractal geometry. Pharm. Res. **12**(11), 1694–1700 (1995)
111. Tromelin, A., Hautbout, G., Pourcelot, Y.: Application of fractal geometry to dissolution kinetic study of a sweetener excipient. Int. J. Pharm. **224**(1–2), 131–140 (2001)
112. Cavallari, C., Abertini, B., Gonzalez-Rodriguez, M., Rodriguez, L.: Improved dissolution behavior of steam-granulated piroxicam. Eur. J. Pharm. Biopharm. **54**(1), 65–73 (2002)
113. Fini, A., Holgado, M., Rodriguez, L., Cavallari, C.: Ultrasound-compacted indomethacin/polyvinylpyrrolidone systems: effect of compaction process on particle morphology and dissolution behavior. J. Pharm. Sci. **91**(8), 1880–1890 (2002)
114. Pettit, F., Bowie, J.: Protein surface roughness and small molecular binding sites. J. Mol. Biol. **285**(4), 1377–1382 (1999)
115. Farin, D., Avnir, D.: Use of fractal geometry to determine effects of surface morphology on drug dissolution. J. Pharm. Sci. **81**(1), 54–57 (1992)
116. Farin, D., Avnir, D.: The reaction dimension in catalysis on dispersed metals. J. Am. Chem. Soc. **110**, 2039–2045 (1988)
117. Valsami, G., Macheras, P.: Determination of fractal reaction dimension in dissolution studies. Eur. J. Pharm. Sci. **3**(3), 163–169 (1995)
118. Weibull, W.: A statistical distribution function of wide applicability. J. Appl. Mech. **18**, 293–297 (1951)
119. Langenbucher, F.: Linearization of dissolution rate curves by the Weibull distribution. J. Pharm. Pharmacol. **24**(12), 979–981 (1972)
120. Djordjevic, A., Mendas, I.: Method for modelling in vitro dissolution profiles of drugs using gamma distribution. Eur. J. Pharm. Biopharm. **44**(2), 215–217 (1997)
121. Smoluchowski, M.: Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Loesungen. Z. Phys. Chem. **92**, 129–168 (1917)

122. Elkoshi, Z.: On the variability of dissolution data. Pharm. Res. **14**(10), 1355–1362 (1997)
123. Dokoumetzidis, A., Macheras, P.: On the heterogeneity of drug dissolution and release. Pharm. Res. **17**(2), 108–112 (2000)
124. Lansky, P., Weiss, M.: Does the dose-solubility ratio affect the mean dissolution time of drugs? Pharm. Res. **16**(9), 1470–1476 (1999)
125. Dokoumetzidis, A., Macheras, P.: A population growth model of dissolution. Pharm. Res. **14**(9), 1122–1126 (1997)
126. Dokoumetzidis, A., Papadopoulou, V., Valsami, G., Macheras, P.: Development of a reaction-limited model of dissolution: application to official dissolution tests experiments. Int. J. Pharm. **355**(1–2), 114–125 (2008)
127. May, R.: Simple mathematical models with very complicated dynamics. Nature **261**(5560), 459–467 (1976)
128. Wiggins, S.: Introduction to Applied Nonlinear Dynamical Systems and Chaos, 3rd edn. Springer, New York (1990)
129. Shah, V., Noory, A., Noory, C., McCullough, B., Clarke, S., Everett, R., Naviasky, H., Srinivasan, B., Fortman, D., Skelly, J.: In vitro dissolution of sparingly water-soluble drug dosage forms. Int. J. Pharm. **125**, 99–106 (1995)
130. Yamamoto, K., Nakano, M., Arita, T., Takayama, Y., Nakai, Y.: Dissolution behavior and bioavailability of phenytoin from a ground mixture with microcrystalline cellulose. J. Pharm. Sci. **65**(10), 1484–1488 (1976)
131. Suzuki, H., Sunada, H.: Influence of water-soluble polymers on the dissolution of nifedipine solid dispersions with combined carriers. Chem. Pharm. Bull. **46**(3), 482–487 (1998)
132. Valsami, G., Dokoumetzidis, A., Macheras, P.: Modeling of supersaturated dissolution data. Int. J. Pharm. **181**(2), 153–157 (1999)
133. Charkoftaki, G., Dokoumetzidis, A., Valsami, G., Macheras, P.: Supersaturated dissolution data and their interpretation: the TPGS-carbamazepine model case. J. Pharm. Pharmacol. **63**(3), 352–361 (2011)
134. Lansky, P., Weiss, M.: Modeling heterogeneity of particles and random effects in drug dissolution. Pharm. Res. **18**(7), 1061–1067 (2001)
135. Lansky, P., Weiss, M.: Classification of dissolution profiles in terms of fractional dissolution rate and a novel measure of heterogeneity. J. Pharm. Sci. **92**(8), 1632–1647 (2003)
136. Lansky, P., Lanska, V., Weiss, M.: A stochastic differential equation model for drug dissolution and its parameters. J. Control. Release **100**(2), 267–274 (2004)
137. Lansky, P., Weiss, M.: Role of heterogeneity in deterministic models of drug dissolution and their statistical characteristics. Biosystems **71**(1–2), 123–131 (2003)
138. Tsong, Y., Hammerstrom, T., Sathe, P., Shah, V.: Statistical assessment of mean differences between two dissolution data sets. Drug Inf. J. **30**, 1105–1112 (1996)
139. Chow, S., Ki, F.: Statistical comparison between dissolution profiles of drug products. J. Biopharm. Stat. **7**(2), 241–258 (1997)
140. Costa, P., Lobo, J.: Modeling and comparison of dissolution profiles. Eur. J. Pharm. Sci. **13**, 123–133 (2001)
141. CDER: Guidance for Industry. Dissolution testing of immediate release solid oral dosage forms. Technical Report, Center for Drug Evaluation and Research, Food and Drug Administration (1997)
142. CDER: Guidance for Industry. Waiver of in vivo bioavailability and bioequivalence studies for immediate release solid oral dosage forms based on a biopharmaceutics classification system. Technical Report, Center for Drug Evaluation and Research, Food and Drug Administration (2000)
143. CPMP: Note for guidance on the investigation of bioavailability and bioequivalence. Technical Report, Committee for Proprietary Medicinal Products, European Medicines Agency (2001)
144. Liu, J., Ma, M., Chow, S.: Statistical evaluation of similarity factor f2 as a criterion for assessment of similarity between dissolution profiles. Drug Inf. J. **31**, 1255–1271 (1997)

145. Shah, V., Tsong, Y., Sathe, P., Liu, J.: In vitro dissolution profile comparison-statistics and analysis of the similarity factor, f2. Pharm. Res. **15**(6), 889–896 (1998)
146. Vertzoni, M., Symillides, M., Iliadis, A., Nicolaides, E., Reppas, C.: Comparison of simulated cumulative drug versus time data sets with indices. Eur. J. Pharm. Biopharm. **56**(3), 421–428 (2003)
147. Jacobs, M.: Some aspects of cell permeability to weak electrolytes. Cold Spring Harb. Symp. Quant. Biol. **8**, 30–39 (1940)
148. Shore, P., Brodie, B., Hogben, C.: The gastric secretion of drugs: a pH partition hypothesis. J. Pharmacol. Exp. Ther. **119**(3), 361–369 (1957)
149. Schanker, L., Shore, P., Brodie, B., Hogben, C.: Absorption of drugs from the stomach. I. The rat. J. Pharmacol. Exp. Ther. **120**(4), 528–539 (1957)
150. Hogben, C., Schanker, L., Tocco, D., Brodie, B.: Absorption of drugs from the stomach. II. The human. J. Pharmacol. Exp. Ther. **120**(4), 540–545 (1957)
151. Schanker, L., Tocco, D., Brodie, B., Hogben, C.: Absorption of drugs from the rat small intestine. J. Pharmacol. Exp. Ther. **123**(1), 81–88 (1958)
152. Hogben, C., Tocco, D., Brodie, B., Schanker, L.: On the mechanism of intestinal absorption of drugs. J. Pharmacol. Exp. Ther. **125**(4), 275–282 (1959)
153. Schanker, L.: On the mechanism of absorption of drugs from the gastrointestinal tract. J. Med. Pharm. Chem. **2**, 343–359 (1960)
154. Suzuki, A., Higuchi, W., Ho, N.: Theoretical model studies of drug absorption and transport in the gastrointestinal tract. I. J. Pharm. Sci. **59**(5), 644–651 (1970)
155. Suzuki, A., Higuchi, W., Ho, N.: Theoretical model studies of drug absorption and transport in the gastrointestinal tract. II. J. Pharm. Sci. **59**(5), 651–659 (1970)
156. Winne, D.: Shift of pH-absorption curves. J. Pharmacokinet. Biopharm. **5**(1), 53–94 (1977)
157. Amidon, G., Lennernas, H., Shah, V., Crison, J.: A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. Pharm. Res. **12**(3), 413–420 (1995)
158. Testa, B., Carrupt, P., Gaillard, P., Billois, F., Weber, P.: Lipophilicity in molecular modeling. Pharm. Res. **13**(3), 335–343 (1996)
159. VanDeWaterbeemd, H., Testa, B.: The parametrization of lipophilicity and other structural properties in drug design. In: Testa, B. (ed.) Advances in Drug Research, vol. 16, pp. 87–227. Academic, London (1988)
160. Dressman, J., Amidon, G., Fleisher, D.: Absorption potential: estimating the fraction absorbed for orally administered compounds. J. Pharm. Sci. **74**(5), 588–589 (1985)
161. Boxenbaum, H.: Absorption potential and its variants. Pharm. Res. **16**(12), 1893 (1999)
162. Macheras, P., Symillides, M.: Toward a quantitative approach for the prediction of the fraction of dose absorbed using the absorption potential concept. Biopharm. Drug Dispos. **10**(1), 43–53 (1989)
163. Balon, K., Riebesehl, B., Muller, B.: Drug liposome partitioning as a tool for the prediction of human passive intestinal absorption. Pharm. Res. **16**(6), 882–888 (1999)
164. Sanghvi, T., Ni, N., Yalkowsky, S.: A simple modified absorption potential. Pharm. Res. **18**(12), 1794–1796 (2001)
165. Sanghvi, T., Ni, N., Mayersohn, M., Yalkowsky, S.: Predicting passive intestinal absorption using a single parameter. Quant. Struct. Act. Relat. Comb. Sci. **22**(2), 247–257 (2003)
166. Johnson, K., Swindell, A.: Guidance in the setting of drug particle size specifications to minimize variability in absorption. Pharm. Res. **13**(12), 1795–1798 (1996)
167. Komiya, I., Park, J., Kamani, A., Ho, N., Higuchi, W.: Quantitative mechanistic studies in simultaneous fluid flow and intestinal absorption using steroids as model solutes. Int. J. Pharm. **4**(3), 249–262 (1980)
168. Ho, N., Raub, T., Burton, P., Barsuhn, C., Adson, A., Audus, K., Borchardt, R.: Quantitative approaches to delineate transport mechanisms in cell culture monolayers. In: Amidon, G., Lee, P. (eds.) Transport Processes in Pharmaceutical Systems, pp. 219–316. Marcel Dekker, New York (2000)

169. Hilgendorf, C., Spahn-Langguth, H., Regardh, C., Lipka, E., Amidon, G., Langguth, P.: Caco-2 versus Caco-2/HT29-MTX co-cultured cell lines: permeabilities via diffusion, inside- and outside-directed carrier-mediated transport. J. Pharm. Sci. **89**(1), 63–75 (2000)
170. Irvine, J., Takahashi, L., Lockhart, K., Cheong, J., Tolan, J., Selick, H., Grove, J.: MDCK (Madin-Darby canine kidney) cells: a tool for membrane permeability screening. J. Pharm. Sci. **88**(1), 28–33 (1999)
171. Zhu, C., Jiang, L., Chen, T., Hwang, K.: A comparative study of artificial membrane permeability assay for high throughput profiling of drug absorption potential. Eur. J. Med. Chem. **37**(5), 399–407 (2002)
172. Bergstrom, C., Strafford, M., Lazorova, L., Avdeef, A., Luthman, K., Artursson, P.: Absorption classification of oral drugs based on molecular surface properties. J. Med. Chem. **46**(4), 558–570 (2003)
173. Artursson, P., Palm, K., Luthman, K.: Caco-2 monolayers in experimental and theoretical predictions of drug transport. Adv. Drug Deliv. Rev. **46**(1–3), 27–43 (2001)
174. Sun, D., Lennernas, H., Welage, L., Barnett, J., Landowski, C., Foster, D., Fleisher, D., Lee, K., Amidon, G.: Comparison of human duodenum and Caco-2 gene expression profiles for 12,000 gene sequences tags and correlation with permeability of 26 drugs. Pharm. Res. **19**(10), 1400–1416 (2002)
175. Dressman, J., Fleisher, D., Amidon, G.: Physicochemical model for dose-dependent drug absorption. J. Pharm. Sci. **73**(9), 1274–1279 (1984)
176. Macheras, P., Reppas, C., Dressman, J.: Estimate of volume/flow ratio of gastrointestinal fluids in humans using pharmacokinetic data. Pharm. Res. **7**(5), 518–522 (1990)
177. Yu, L., Crison, J., Amidon, G.: Compartmental transit and dispersion model analysis of small intestinal transit flow in humans. Int. J. Pharm. **140**, 111–118 (1996)
178. Yu, L., Amidon, G.: Characterization of small intestinal transit time distribution in humans. Int. J. Pharm. **171**(2), 157–163 (1998)
179. Yu, L., Amidon, G.: Saturable small intestinal drug absorption in humans: modeling and interpretation of cefatrizine data. Eur. J. Pharm. Biopharm. **45**(2), 199–203 (1998)
180. Agoram, B., Woltosz, W., Bolger, M.: Predicting the impact of physiological and biochemical processes on oral drug bioavailability. Adv. Drug Deliv. Rev. **50**(S1), S41–67 (2001)
181. Ni, P., Ho, N., Fox, J., Leuenberger, H., Higuchi, W.: Theoretical model studies of intestinal drug absorption V. Non-steady-state fluid flow and absorption. Int. J. Pharm. **5**(1), 33–47 (1980)
182. Ho, N., Park, J., Ni, P., Higuchi, W.: Advancing quantitative and mechanistic approaches in interfacing gastrointestinal drug absorption studies in animals and humans. In: Crouthamel, W., Sarapu, A. (eds.) Animal Models for Oral Drug Delivery. In Situ and In Vivo Approaches, pp. 27–106. American Pharmaceutical Association, Washington (1983)
183. Unice, K., Logan, B.: The insignificant role of hydrodynamic dispersion on bacterial transport. J. Environ. Eng. **126**(6), 491–500 (2000)
184. Willmann, S., Schmitt, W., Keldenich, J., Dressman, J.: A physiologic model for simulating gastrointestinal flow and drug absorption in rats. Pharm. Res. **20**(11), 1766–1771 (2003)
185. Willmann, S., Schmitt, W., Keldenich, J., Lippert, J., Dressman, J.: A physiological model for the estimation of the fraction dose absorbed in humans. J. Med. Chem. **47**(16), 4022–4031 (2004)
186. Dokoumetzidis, A., Macheras, P.: A dispersion-convection model for the study of the gastrointestinal drug absorption. AAPS PharmSci **5**(4), R6086 (2003)
187. Agutter, P., Malone, P., Wheatley, D.: Intracellular transport mechanisms: a critique of diffusion theory. J. Theor. Biol. **176**(2), 261–272 (1995)
188. Wheatley, D.: Difusion theory in biology: its validity and relevance. J. Biol. Educ. **27**, 181–188 (1993)
189. Macheras, P., Argyrakis, P.: Gastrointestinal drug absorption: is it time to consider heterogeneity as well as homogeneity? Pharm. Res. **14**, 842–847 (1997)
190. Davenport, H.: Physiology of the Digestive Tract, 5th edn. Year Book Medical, Chicago (1982)

191. Davis, S., Hardy, J., Fara, J.: Transit of pharmaceutical dosage forms through the small intestine. Gut **27**(8), 886–892 (1986)
192. Digenis, G., Sandefer, E.: Gamma scintigraphy and neutron activation techniques in the in vivo assessment of orally administered dosage forms. Crit. Rev. Ther. Drug Carrier Syst. **7**(4), 309–345 (1991)
193. Davis, S., Stockwell, A., Taylor, M., Hardy, J., Whalley, D., Wilson, C., Bechgaard, H., Christensen, F.: The effect of density on the gastric emptying of single- and multiple-unit dosage forms. Pharm. Res. **3**(4), 208–213 (1986)
194. Khosla, R., Feely, L., Davis, S.: Gastrointestinal transit of non-disintegrating tablets in fed subjects. Int. J. Pharm. **53**(2), 107–117 (1989)
195. Wilding, I., Hardy, J., Maccari, M., Ravelli, V., Davis, S.: Scintigraphic and pharmacokinetic assessment of a multiparticulate sustained release formulation of diltiazem. Int. J. Pharm. **76**(1–2), 133–143 (1991)
196. Spiller, R., Brown, M., Phillips, S.: Emptying of the terminal ileum in intact humans. Influence of meal residue and ileal motility. Gastroenterology **92**(3), 724–729 (1987)
197. Krevsky, B., Malmud, L., D'Ercole, F., Maurer, A., Fisher, R.: Colonic transit scintigraphy. A physiologic approach to the quantitative measurement of colonic transit in humans. Gastroenterology **91**(5), 1102–1112 (1986)
198. Adkin, D., Davis, S., Sparrow, R., Wilding, I.: Colonic transit of different sized tablets in healthy subjects. J. Control. Release **23**(2), 147–156 (1993)
199. Scher, H., Shlesinger, M., Bendler, J.: Time-scale invariance in transport and relaxation. Phys. Today **44**(1), 26–34 (1991)
200. Katori, N., Aoyagi, N., Terao, T.: Estimation of agitation intensity in the GI tract in humans and dogs based on in vitro/in vivo correlation. Pharm. Res. **12**(2), 237–243 (1995)
201. Shameem, M., Katori, N., Aoyagi, N., Kojima, S.: Oral solid controlled release dosage forms: role of GI-mechanical destructive forces and colonic release in drug absorption under fasted and fed conditions in humans. Pharm. Res. **12**(7), 1049–1054 (1995)
202. Boxenbaum, H., Jodhka, G., Ferguson, A., Riegelman, S., MacGregor, T.: The influence of bacterial gut hydrolysis on the fate of orally administered isonicotinuric acid in man. J. Pharmacokinet. Biopharm. **2**(3), 211–237 (1974)
203. Sawamoto, T., Haruta, S., Kurosaki, Y., Higaki, K., Kimura, T.: Prediction of the plasma concentration profiles of orally administered drugs in rats on the basis of gastrointestinal transit kinetics and absorbability. J. Pharm. Pharmacol. **49**(4), 450–457 (1997)
204. Nittmann, J., Daccord, G., Stanley, H.: Fractal growth of viscous fingers: a quantitative characterization of a fluid instability phenomenon. Nature **314**, 141–144 (1985)
205. VanDamme, H.: Flow and interfacial instabilities in Newtonian and colloidal fluids (or the birth, life and death of a fractal). In: Avnir, D. (ed.) The Fractal Approach to Heterogeneous Chemistry, pp. 199–226. Wiley, Chishester (1989)
206. Bhaskar, K., Garik, P., Turner, B., Bradley, J., Bansil, R., Stanley, H., LaMont, J.: Viscous fingering of HCl through gastric mucin. Nature **360**(6403), 458–461 (1992)
207. Holm, L., Flemstrom, G.: Microscopy of acid transport at the gastric surface in vivo. J. Intern. Med. Suppl. **732**, 91–95 (1990)
208. Daccord, G., Lenormand, R.: Fractal patterns from chemical dissolution. Nature **325**, 41–43 (1987)
209. Daccord, G., Chemical dissolution of a porous medium by a reactive fluid. Phys. Rev. Lett. **58**(5), 479–482 (1987)
210. Veng-Pedersen, P., Modi, N.: Optimal extravascular dosing intervals. J. Pharmacokinet. Biopharm. **19**(4), 405–412 (1991)
211. Charter, M., Gull, S.: Maximum entropy and its application to the calculation of drug absorption rates. J. Pharmacokinet. Biopharm. **15**(6), 645–655 (1987)
212. Verotta, D.: Two constrained deconvolution methods using spline functions. J. Pharmacokinet. Biopharm. **21**(5), 609–636 (1993)
213. Cutler, D.: Numerical deconvolution by least squares: Use of prescribed input functions. J. Pharmacokinet. Biopharm. **6**(3), 227–241 (1978)

214. Cutler, D.: Numerical deconvolution by least squares: use of polynomials to represent the input function. J. Pharmacokinet. Biopharm. **6**(3), 243–263 (1978)
215. Veng-Pedersen, P.: An algorithm and computer program for deconvolution in linear pharmacokinetics. J. Pharmacokinet. Biopharm. **8**(5), 463–481 (1980)
216. Higaki, K., Yamashita, S., Amidon, G.: Time-dependent oral absorption models. J. Pharmacokinet. Pharmacodyn. **28**(2), 109–128 (2001)
217. Macheras, P., Argyrakis, P., Polymilis, C.: Fractal geometry, fractal kinetics and chaos en route to biopharmaceutical sciences. Eur. J. Drug Metab. Pharmacokinet. **21**(2), 77–86 (1996)
218. Tsang, Y., Pop, R., Gordon, P., Hems, J., Spino, M.: High variability in drug pharmacokinetics complicates determination of bioequivalence: experience with verapamil. Pharm. Res. **13**(6), 846–850 (1996)
219. Tozer, T., Bois, F., Hauck, W., Chen, M., Williams, R.: Absorption rate vs. exposure: which is more useful for bioequivalence testing? Pharm. Res. **13**(3), 453–456 (1996)
220. Kalampokis, A., Argyrakis, P., Macheras, P.: Heterogeneous tube model for the study of small intestinal transit flow. Pharm. Res. **16**(1), 87–91 (1999)
221. Kalampokis, A., Argyrakis, P., Macheras, P.: A heterogeneous tube model of intestinal drug absorption based on probabilistic concepts. Pharm. Res. **16**(11), 1764–1769 (1999)
222. Lipinski, C., Lombardo, F., Dominy, B., Feeney, P.: Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. **46**(1–3), 3–26 (2001)
223. Stenberg, P., Bergstrom, C., Luthman, K., Artursson, P.: Theoretical predictions of drug absorption in drug discovery and development. Clin. Pharmacokinet. **41**(11), 877–899 (2002)
224. Zhao, Y., Le, J., Abraham, M., Hersey, A., Eddershaw, P., Luscombe, C., Butina, D., Beck, G., Sherborne, B., Cooper, I., Platts, J.: Evaluation of human intestinal absorption data and subsequent derivation of a quantitative structure-activity relationship (QSAR) with the Abraham descriptors. J. Pharm. Sci. **90**(6), 749–784 (2001)
225. Zhao, Y., Abraham, M., Le, J., Hersey, A., Luscombe, C., Beck, G., Sherborne, B., Cooper, I.: Rate-limited steps of human oral absorption and QSAR studies. Pharm. Res. **19**(10), 1446–1457 (2002)
226. Turner, J., Maddalena, D., Agatonovic-Kustrin, S.: Bioavailability prediction based on molecular structure for a diverse series of drugs. Pharm. Res. **21**(1), 68–82 (2004)
227. CDER: Guidance for Industry. Waiver of in vivo bioavailability studies for immediate release solid oral dosageforms based on a biopharmaceutics classification system. Technical Report, Center for Drug Evaluation and Research, Food and Drug Administration (2000)
228. Rinaki, E., Valsami, G., Macheras, P.: Quantitative biopharmaceutics classification system: the central role of dose/solubility ratio. Pharm. Res. **20**(12), 1917–1925 (2003)
229. Yu, L., Amidon, G., Polli, J., Zhao, H., Mehta, M., Conner, D., Shah, V., Lesko, L., Chen, M., Lee, V., Hussain, A.: Biopharmaceutics classification system: the scientific basis for biowaiver extensions. Pharm. Res. **19**(7), 921–925 (2002)
230. Polli, J., Yu, L., Cook, J., Amidon, G., Borchardt, R., Burnside, B., Burton, P., Chen, M., Conner, D., Faustino, P., Hawi, A., Hussain, A., Joshi, H., Kwei, G., Lee, V., Lesko, L., Lipper, R., Loper, A., Nerurkar, S., Polli, J., Sanvordeker, D., Taneja, R., Uppoor, R., Vattikonda, C., Wilding, I., Zhang, G.: Summary workshop report: biopharmaceutics classification system - Implementation challenges and extension opportunities. J. Pharm. Sci. **93**(6), 1375–1381 (2004)
231. Yazdanian, M., Briggs, K., Jankovsky, C., Hawi, A.: The 'high solubility' definition of the current FDA Guidance on Biopharmaceutical Classification System may be too strict for acidic drugs. Pharm. Res. **21**(2), 293–299 (2004)
232. Yu, L., Carlin, A., Amidon, G., Hussain, A.: Feasibility studies of utilizing disk intrinsic dissolution rate to classify drugs. Int. J. Pharm. **270**(1–2), 221–227 (2004)
233. Ginski, M., Taneja, R., Polli, J.: Prediction of dissolution-absorption relationships from a continuous dissolution/Caco-2 system. AAPS PharmSci **1**(2), E3 (1999)

234. Rinaki, E., Dokoumetzidis, A., Valsami, G., Macheras, P.: Identification of biowaivers among Class II drugs: theoretical justification and practical examples. Pharm. Res. **21**(9), 1567–1572 (2004)
235. Davies, N., Anderson, K.: Clinical pharmacokinetics of diclofenac. Therapeutic insights and pitfalls. Clin. Pharmacokinet. **33**(3), 184–213 (1997)
236. Blume, H., Schug, B.: The biopharmaceutics classification system (BCS): class III drugs - better candidates for BA/BE waiver? Eur. J. Pharm. Sci. **9**(2), 117–121 (1999)
237. CHMP: Guideline on the investigation of bioequivalence. Technical Report, Committee for Medicinal Products for Human Use, European Medicines Agency (2010)
238. Reppas, C., Vertzoni, M.: Biorelevant in-vitro performance testing of orally administered dosage forms. J. Pharm. Pharmacol. **64**(7), 919–930 (2012)
239. Daousani, C., Macheras, P.: Scientific considerations concerning the EMA change in the definition of "dose" of the BCS-based biowaiver guideline and implications for bioequivalence. Int. J. Pharm. **478**(2), 606–609 (2015)
240. Macheras, P., Karalis, V., Valsami, G.: Keeping a critical eye on the science and the regulation of oral drug absorption: a review. J. Pharm. Sci. **102**(9), 3018–3036 (2013)
241. Papadopoulou, V., Valsami, G., Dokoumetzidis, A., Macheras, P.: Biopharmaceutics classification systems for new molecular entities (BCS-NMEs) and marketed drugs (BCS-MD): theoretical basis and practical examples. Int. J. Pharm. **361**(1–2), 70–77 (2008)
242. Charkoftaki, G., Dokoumetzidis, A., Valsami, G., Macheras, P.: Elucidating the role of dose in the biopharmaceutics classification of drugs: the concepts of critical dose, effective in vivo solubility, and dose-dependent BCS. Pharm. Res. **29**(11), 3188–3198 (2012)
243. WHO: Proposal to waive in vivo bioequivalence requirements for WHO model list of essential medicines immediate-release, solid oral dosage forms. Technical Report, World Health Organization (2006)
244. Sediq, A., Kubbinga, M., Langguth, P., Dressman, J.: The impact of the EMA change in definition of "dose" on the BCS dose-solubility ratio: a review of the biowaiver monographs. J. Pharm. Sci. **103**(1), 65–70 (2014)
245. Bergstrom, C., Andersson, S., Fagerberg, J., Ragnarsson, G., Lindahl, A.: Is the full potential of the biopharmaceutics classification system reached? Eur. J. Pharm. Sci. **57**, 224–231 (2014)
246. Macheras, P., Karalis, V.: A non-binary biopharmaceutical classification of drugs: the ABGamma system. Int. J. Pharm. **464**(1–2), 85–90 (2014)
247. Wu, C., Benet, L.: Predicting drug disposition via application of BCS: transport/ absorption/elimination interplay and development of a biopharmaceutics drug disposition classification system. Pharm. Res. **22**(1), 11–23 (2005)
248. Dokoumetzidis, A., Macheras, P.: The changing face of the rate concept in biopharmaceutical sciences: from classical to fractal and finally to fractional. Pharm. Res. **28**(5), 1229–1232 (2011)
249. Benet, L., Larregieu, C.: The FDA should eliminate the ambiguities in the current BCS biowaiver guidance and make public the drugs for which BCS biowaivers have been granted. Clin. Pharmacol. Ther. **88**(3), 405–407 (2010)
250. Custodio, J., Wu, C., Benet, L.: Predicting drug disposition, absorption/elimination/ transporter interplay and the role of food on drug absorption. Adv. Drug Deliv. Rev. **60**(6), 717–733 (2008)
251. Benet, L.: Predicting drug disposition via application of a biopharmaceutics drug disposition classification system. Basic Clin. Pharmacol. Toxicol. **106**(3), 162–167 (2010)
252. Khandelwal, A., Bahadduri, P., Chang, C., Polli, J., Swaan, P., Ekins, S.: Computational models to assign biopharmaceutics drug disposition classification from molecular structure. Pharm. Res. **24**(12), 2249–2262 (2007)
253. Benet, L., Amidon, G., Barends, D., Lennernas, H., Polli, J., Shah, V., Stavchansky, S., Yu, L.: The use of BDDCS in classifying the permeability of marketed drugs. Pharm. Res. **25**(3), 483–488 (2008)
254. Benet, L., Broccatelli, F., Oprea, T.: BDDCS applied to over 900 drugs. AAPS J. **13**(4), 519–547 (2011)

255. Broccatelli, F., Cruciani, G., Benet, L., Oprea, T.: BDDCS class prediction for new molecular entities. Mol. Pharm. **9**(3), 570–580 (2012)
256. Benet, L.: The role of BCS (biopharmaceutics classification system) and BDDCS (biopharmaceutics drug disposition classification system) in drug development. J. Pharm. Sci. **102**(1), 34–42 (2013)
257. Larregieu, C., Benet, L.: Distinguishing between the permeability relationships with absorption and metabolism to improve BCS and BDDCS predictions in early drug discovery. Mol. Pharm. **11**(4), 1335–1344 (2014)
258. Zheng, Y., Benet, L., Okochi, H., Chen, X.: pH dependent but not P-gp dependent bidirectional transport study of S-propranolol: the importance of passive diffusion. Pharm. Res. **32**(8), 2516–2526 (2015)
259. VanRossum, J., DeBie, J.: Chaos and illusion. Trends Pharmacol. Sci. **12**(10), 379–383 (1991)
260. Dokoumetzidis, A., Macheras, P.: Investigation of absorption kinetics by the phase plane method. Pharm. Res. **15**(8), 1262–1269 (1998)
261. Dokoumetzidis, A., Iliadis, A., Macheras, P.: An alternative method for the estimation of the terminal slope when a few data points are available. J. Pharm. Sci. **88**(5), 557–560 (1999)
262. Tothfalusi, L., Endrenyi, L., Midha, K.: Scaling or wider bioequivalence limits for highly variable drugs and for the special case of C(max). Int. J. Clin. Pharmacol. Ther. **41**(5), 217–225 (2003)
263. Tothfalusi, L., Endrenyi, L.: Limits for the scaled average bioequivalence of highly variable drugs and drug products. Pharm. Res. **20**(3), 382–389 (2003)
264. Karalis, V., Symillides, M., Macheras, P.: Novel scaled average bioequivalence limits based on GMR and variability considerations. Pharm. Res. **21**(10), 1933–1942 (2004)
265. Wang, Z., Zhenya, H., Chen, J.: Chaotic behavior of gastric migrating myoelectrical complex. IEEE Trans. Biomed. Eng. **51**(8), 1401–1406 (2004)
266. Marathe, P., Sandefer, E., Kollia, G., Greene, D., Barbhaiya, R., Lipper, R., Page, R., Doll, W., Ryo, U., Digenis, G.: In vivo evaluation of the absorption and gastrointestinal transit of avitriptan in fed and fasted subjects using gamma scintigraphy. J. Pharmacokinet. Biopharm. **26**(1), 1–29 (1998)
267. Richalet, J., Rault, A., Pouliquen, R.: Identification des Processus par la Methode du Modele. Gordon and Breach, Paris (1971)
268. Balant, L., Gex-Fabry, M.: Modelling during drug development. Eur. J. Pharm. Biopharm. **50**(1), 13–26 (2000)
269. Wise, M.: Negative power functions of time in pharmacokinetics and their implications. J. Pharmacokinet. Biopharm. **13**(3), 309–346 (1985)
270. Weiss, M.: Use of gamma distributed residence times in pharmacokinetics. Eur. J. Clin. Pharmacol. **25**, 695–702 (1983)
271. Weiss, M.: A note on the role of generalized inverse Gaussian distributions of circulatory transit times in pharmacokinetics. J. Math. Biol. **20**, 95–102 (1984)
272. Beard, D., Bassingthwaighte, J.: Power-law kinetics of tracer washout from physiological systems. Ann. Biomed. Eng. **26**, 775–779 (1998)
273. Norwich, K.: Noncompartment models of whole-body clearance of tracers: a review. Ann. Biomed. Eng. **25**, 421–439 (1997)
274. Macheras, P.: Carrier-mediated transport can obey fractal kinetics. Pharm. Res. **12**(4), 541–548 (1995)
275. Havlin, S., Ben-Avraham, D.: Diffusion in disordered media. Adv. Phys. **51**(1), 187–292 (2002)
276. Mandelbrot, B.: The Fractal Geometry of Nature. W.H. Freeman, New York (1983)
277. VanBeek, J., Roger, S., Bassingthwaighte, J.: Regional myocardial flow heterogeneity explained with fractal networks. Am. J. Physiol. **257**(5(2)), H1670–1680 (1989)
278. VanDerGraff, K., Fox, S.: Concepts of Human Anatomy and Physiology, 5th edn. McGraw Hill, New York (2000)
279. Suwa, N., Niwa, T., Fukasawa, H., Sasaki, Y.: Estimation of intravascular blood pressure gradient by mathematical analysis of arterial casts. Tohoku J. Exp. Med. **79**, 168–198 (1963)

280. Suwa, N., Takahashi, T.: Morphological and Morphometrical Analysis of Circulation in Hypertension and Ischemic Kidney. Urban und Schwarzenberg, München (1971)
281. Macheras, P.: A fractal approach to heterogeneous drug distribution: calcium pharmacokinetics. Pharm. Res. **13**(5), 663–670 (1996)
282. Gerlowski, L., Jain, R.: Physiologically based pharmacokinetic modeling: principles and applications. J. Pharm. Sci. **72**(10), 1103–1127 (1983)
283. Wiltshire, H., Sutton, B., Heeps, G., Betty, A., Angus, D., Harris, S., Worth, E., Welker, H.: Metabolism of the calcium antagonist, mibefradil (POSICOR, Ro 40-5967). Part III. Comparative pharmacokinetics of mibefradil and its major metabolites in rat, marmoset, cynomolgus monkey and man. Xenobiotica **27**(6), 557–571 (1997)
284. Campra, J., Reynolds, T.: The hepatic circulation. In: Arias, I., Popper, H., Schachter, D., Shafritz, D. (eds.) In The Liver: Biology and Pathobiology, pp. 627–645. Raven, New York (1982)
285. Glenny, R., Robertson, H.: Fractal properties of pulmonary blood flow: characterization of spatial heterogeneity. J. Appl. Physiol. **69**(2), 532–545 (1990)
286. Glenny, R., Robertson, H.: Fractal modeling of pulmonary blood flow heterogeneity. J. Appl. Physiol. **70**(3), 1024–1030 (1991)
287. Lin, J., Lu, A.: Role of pharmacokinetics and metabolism in drug discovery and development. Pharmacol. Rev. **49**(4), 403–449 (1997)
288. Houston, J.B., Carlile, D.: Prediction of hepatic clearance from microsoms, hepatocytes, and liver slices. Drug Metab. Rev. **29**(4), 891–922 (1997)
289. Obach, R.: Prediction of human clearance of twenty-nine drugs from hepatic microsomal intrinsic clearance data: an examination of in vitro half-life approach and nonspecific binding to microsomes. Drug Metab. Dispos. **27**(11), 1350–1359 (1999)
290. Ito, K., Houston, J.: Comparison of the use of liver models for predicting drug clearance using in vitro kinetic data from hepatic microsomes and isolated hepatocytes. Pharm. Res. **21**(5), 785–792 (2004)
291. Rowland, M., Benet, L., Graham, G.: Clearance concepts in pharmacokinetics. J. Pharmacokinet. Biopharm. **1**(2), 123–136 (1973)
292. Winkler, K., Keiding, S., Tygstrup, N.: Clearance as a quantitative measure of liver function. In: Paumgartner, P., Preisig, R. (eds.) The Liver: Quantitative Aspects of Structure and Function, pp. 144–155. Karger, Basle (1973)
293. Roberts, M., Donaldson, J., Rowland, M.: Models of hepatic elimination: comparison of stochastic models to describe residence time distributions and to predict the influence of drug distribution, enzyme heterogeneity, and systemic recycling on hepatic elimination. J. Pharmacokinet. Biopharm. **16**(1), 41–83 (1988)
294. Roberts, M., Rowland, M.: A dispersion model of hepatic elimination: 1. Formulation of the model and bolus concentrations. J. Pharmacokinet. Biopharm. **14**(3), 227–260 (1986)
295. Anissimov, Y., Bracken, A., Roberts, M.: Interconnected-tubes model of hepatic elimination. J. Theor. Biol. **188**(1), 89–101 (1997)
296. Anissimov, Y., Bracken, A., Roberts, M.: Interconnected-tubes model of hepatic elimination: steady-state considerations. J. Theor. Biol. **199**(4), 435–447 (1999)
297. Tirona, R., Schwab, A., Geng, W., Pang, K.: Hepatic clearance models: comparison of the dispersion and Goresky models in outflow profiles from multiple indicator dilution rat liver studies. Drug Metab. Dispos. **26**(5), 465–475 (1998)
298. Niro, R., Byers, J., Fournier, R., Bachmann, K.:" Application of a convective-dispersion model to predict in vivo hepatic clearance from in vitro measurements utilizing cryopreserved human hepatocytes. Curr. Drug Metab. **4**(5), 357–369 (2003)
299. Ridgway, D., Tuszynski, J., Tam, Y.: Reassessing models of hepatic extraction. J. Biol. Phys. **29**(1), 1–21 (2003)
300. Javanaud, C.: The application of a fractal model to the scattering of ultrasound in biological media. J. Acoust. Soc. Am. **86**(2), 493–496 (1989)
301. Liang, E., Derendorf, H.: Pitfalls in pharmacokinetic multicompartment analysis. J. Pharmacokinet. Biopharm. **26**(2), 247–260 (1998)

302. Berry, H.: Monte Carlo simulations of enzyme reactions in two dimensions: fractal kinetics and spatial segregation. Biophys. J. **83**(4), 1891–1901 (2002)
303. Fuite, J., Marsh, R., Tuszynski, J.: Fractal pharmacokinetics of the drug mibefradil in the liver. Phys. Rev. E, Stat. Nonlinear, Soft Matter Phys. **66**(2), 021904 (2002)
304. Skerjanec, A., Tawfik, S., Tam, Y.: Nonlinear pharmacokinetics of mibefradil in the dog. J. Pharm. Sci. **85**(2), 189–192 (1996)
305. Welker, H.: Single- and multiple-dose mibefradil pharmacokinetics in normal and hypertensive subjects. J. Pharm. Pharmacol. **50**(9), 983–987 (1998)
306. Kosmidis, K., Karalis, V., Argyrakis, P., Macheras, P.: Michaelis-Menten kinetics under spatially constrained conditions: application to mibefradil pharmacokinetics. Biophys. J. **87**(3), 1498–1506 (2004)
307. Sun, H., Charef, A.: Fractal system - A time domain approach. Ann. Biomed. Eng. **18**, 597–621 (1990)
308. Suki, B., Barabasi, A., Hantos, Z., Petak, F., Stanley, H.: Avalanches and power-law behaviour in lung inflation. Nature **368**, 615–618 (1994)
309. Maddox, J.: Long-range correlations within DNA. Nature **358**, 103 (1992)
310. Peng, C., Buldyrev, S., Goldberger, A., Havlin, S., Sciortino, F., Simons, M., Stanley, H.: Long-range correlations in nucleotide sequences. Nature **356**, 168–170 (1992)
311. Mandelbrot, B., Wallis, J.: Some long-run properties of geophysical records. Water Resour. Res. **5**, 321–340 (1969)
312. Tang, C., Bak, P.: Critical exponents and scaling relations for self organized critical phenomena. Phys. Rev. Lett. **60**, 2347–2350 (1988)
313. Rhodes, C., Anderson, R.: Power laws governing epidemics in isolated populations. Nature **381**, 600–602 (1996)
314. Mantegna, R., Stanley, H.: Scaling behaviour in the dynamics of an economic index. Nature **376**, 46–49 (1995)
315. West, B., Shlesinger, M.: On the ubiquity of 1/f noise. Int. J. Mod. Phys. B **3**(6), 795–819 (1989)
316. Stanley, M., Amaral, L., Buldyrev, S., Havlin, S., Leschhorn, H., Maass, P., Salinger, M., Stanley, H.: Can statistical physics contribute to the science of economics? Fractals **4**(3), 415–425 (1996)
317. Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality: an explanation of 1/f noise. Phys. Rev. Lett. **59**, 381–384 (1987)
318. Bak, P., Tang, C., Wiesenfeld, K.: Self-organized criticality. Phys. Rev. A **38**, 364–374 (1988)
319. Havlin, S.: Molecular diffusion and reactions. In: Avnir, D. (ed.) The Fractal Approach to Heterogeneous Chemistry, pp. 251–270. Wiley, Chishester (1989)
320. Giona, M.: First-order reaction-diffusion kinetics in complex fractal media. Chem. Eng. Sci. **47**(6), 1503–1515 (1992)
321. Koo, Y., Kopelman, R.: Space- and time-resolved diffusion-limited binary reaction kinetics in capillaries. Experimental observation of segregation, anomalous exponents and depletion zone. J. Stat. Phys. **65**, 893–918 (1991)
322. Kopelman, R., Lin, A., Argyrakis, P.: Non-classical kinetics and reactant segregation in d-dimensional tubular spaces. Phys. Lett. A **232**(1–2), 34–40 (1997)
323. Wise, M., Osborn, S., Anderson, J., Tomlinson, R.: A stochastic model for turnover of radiocalcium based on the observed power laws. Math. Biosci. **2**, 199–224 (1968)
324. Wise, M.: Interpreting both short- and long-term power laws in physiological clearance curves. Math. Biosci. **20**, 327–337 (1974)
325. Marcus, A.: Power laws in compartmental analysis. Part I: a unified stochastic model. Math. Biosci. **23**, 337–350 (1975)
326. Matis, J., Wehrly, T., Generalized stochastic compartmental models with Erlang transit times. J. Pharmacokinet. Biopharm. **18**(6), 589–607 (1990)
327. Eckerman, K., Leggett, R., Williams, L.: An elementary method for solving compartmental models with time-dependent coefficients. Radiat. Prot. Dosim. **41**, 257–263 (1992)

328. Farris, F., Dedrick, R., Allen, P., Smith, J.: Physiological model for the pharmacokinetics of methyl mercury in the growing rat. Toxicol. Appl. Pharmacol. **119**(1), 74–90 (1993)
329. Matis, J., Wehrly, T.: Stochastic models of compartmental systems. Biometrics **35**, 199–220 (1979)
330. Matis, J., An introduction to stochastic compartmental models in pharmacokinetics. In: Pecile, A., Rescigno, A. (eds.) Pharmacokinetics-Mathematical and Statistical Approaches in Metabolism and Distribution of Chemicals and Drugs, pp. 113–128. Plenum, New York (1988)
331. Matis, J., Wehrly, T.: A general approach to non-Markovian compartment models. J. Pharmacokinet. Biopharm. **26**(4), 437–456 (1998)
332. Wagner, J.: Fundamentals of Clinical Pharmacokinetics, 2nd edn. Drug Intelligence, Hamilton (1979)
333. Ljung, L., Gunnarsson, S.: Adaptation and tracking in system identification: a survey. Automatica **26**(1), 7–21 (1990)
334. Ljung, L.: System Identification: Theory for the User. PTR Prentice Hall Information and System Sciences Series, 2nd ed. Prentice Hall, Englewood Cliffs (1999)
335. Savageau, M., Voit, E.: Recasting nonlinear differential equations as S-systems: a canonical nonlinear form. Math. Biosci. **87**(1), 83–115 (1987)
336. Irvine, D., Savageau, M.: Efficient solution of nonlinear ordinary differential equations expressed in S-system canonical form. SIAM J. Numer. Anal. **27**, 704–735 (1990)
337. Sung, S., Lee, J., Lee, I.: Process Identification and PID Control. Wiley, Singapore (2009)
338. Mohler, R.: Bilinear Control Processes with Applications to Engineering, Ecology, and Medicine. Mathematics in Science and Engineering. Academic, New York (1973)
339. Riggs, D.: The Mathematical Approach to Physiological Problems. MIT, Cambridge (1963)
340. Press, W., Teukolsky, S., Vetterling, W., Flannery, B.: Numerical Recipies in C. The Art of Scientific Computing, 2nd edn. Cambridge University Press, Victoria (1992)
341. Gibaldi, M., Perrier, D.: Pharmacokinetics. Drugs and the Pharmaceutical Science, vol. 15, 2nd edn. Marcel Dekker, New York (1982)
342. Matis, J., Wehrly, T., Metzler, C.: On some stochastic formulations and related statistical moments of pharmacokinetic models. J. Pharmacokinet. Biopharm. **11**(1), 77–92 (1983)
343. Soong, T.: Pharmacokinetics with uncertainties in rate constants. Math. Biosci. **12**, 235–243 (1971)
344. Soong, T., Dowdee, J.: Pharmacokinetics with uncertainties in rate constants, III: the inverse problem. Math. Biosci. **19**, 343–353 (1974)
345. Tsokos, J., Tsokos, C.: Statistical modeling of pharmacokinetic systems. J. Dyn. Syst. Meas. Control. **98**, 37–43 (1976)
346. Michaelis, L., Menten, M.: Die Kinetik der Invertinwirkung. Biochem. Z. **49**, 333–369 (1913)
347. Segel, L.: Simplification and scaling. SIAM Rev. **14**(4), 547–571 (1972)
348. Fung, Y.: Biomechanics: Circulation. Springer, New York (1997)
349. Brown, J., West, G.: Scaling in Biology. Oxford University Press, New York (2000)
350. Lefevre, J.: Teleonomical optimization of a fractal model of the pulmonary arterial bed. J. Theor. Biol. **102**(2), 225–248 (1983)
351. Horsfield, K.: Diameters, generations, and orders of branches in the bronchial tree. J. Appl. Physiol. **68**(2), 457–461 (1990)
352. Zamir, M.: On fractal properties of arterial trees. J. Theor. Biol. **197**(4), 517–526 (1999)
353. West, G., Brown, J., Enquist, B.: A general model for the origin of allometric scaling laws in biology. Science **276**, 122–126 (1997)
354. Wagner, J.: Pharmacokinetics for the Pharmaceutical Scientist. Technomic, Lancaster (1993)
355. Weiss, M., Foster, W.: Pharmacokinetic model based on circulatory transport. Eur. J. Clin. Pharmacol. **16**, 287–293 (1979)
356. Vicini, P., Bonadonna, R., Lehtovirta, M., Groop, L., Cobelli, C.: Estimation of blood flow heterogeneity in human skeletal muscle using intravascular tracer data: importance for modeling transcapillary exchange. Ann. Biomed. Eng. **26**, 764–774 (1998)

357. Beard, D., Bassingthwaighte, J.: Advection and diffusion of substances in biological tissues with complex vascular networks. Ann. Biomed. Eng. **28**(3), 253–268 (2000)
358. Oliver, R., Jones, A., Rowland, M.: A whole-body physiologically based pharmacokinetic model incorporating dispersion concepts: short and long time characteristics. J. Pharmacokinet. Biopharm. **28**(1), 27–55 (2001)
359. Dokoumetzidis, A., Macheras, P.: A model for transport and dispersion in the circulatory system based on the vascular fractal tree. Ann. Biomed. Eng. **31**(3), 284–293 (2003)
360. Edwards, D.: A general theory of the macrotransport of nondepositing particles in the lung by convective dispersion. J. Aerosol Sci. **25**(3), 543–565 (1994)
361. Huang, W., Yen, R., McLaurine, M., Bledsoe, G.: Morphometry of the human pulmonary vasculature. J. Appl. Physiol. **81**(5), 2123–2133 (1996)
362. Niemann, C., Henthorn, T., Krejcie, T., Shanks, C., Enders-Klein, C., Avram, M.: Indocyanine green kinetics characterize blood volume and flow distribution and their alteration by propranolol. Clin. Pharmacol. Ther. **67**(4), 342–350 (2000)
363. Karalis, V., Dokoumetzidis, A., Macheras, P.: A physiologically based approach for the estimation of recirculatory parameters. J. Pharmacol. Exp. Ther. **308**(1), 198–205 (2004)
364. Hashimoto, M., Watanabe, G.: Simultaneous measurement of effective hepatic blood flow and systemic circulation. Hepatogastroenterology **47**(36), 1669–1674 (2000)
365. Rescigno, A., Segre, G.: Drug and Tracer Kinetics. Blaisdell, Waltham (1966)
366. Jacquez, J.: Compartmental Analysis in Biology and Medicine, 3rd edn. BioMedware, Ann Arbor (1996)
367. Wise, M., Borsboom, G.: Two exceptional sets of physiological clearance curves and their mathematical form: test cases? Bull. Math. Biol. **51**, 579–596 (1989)
368. Rescigno, A.: The rise and fall of compartmental analysis. Pharm. Res. **44**(4), 335–340 (2001)
369. Weiss, M.: The anomalous pharmacokinetics of amiodarone explained by nonexponential tissue trapping. J. Pharmacokinet. Biopharm. **27**(4), 383–396 (1999)
370. Dokoumetzidis, A.: Contribution on the study of heterogeneity of biopharmaceutical, pharmacokinetic and pharmacodynamic processes. Ph.D. Thesis, National University of Athens (2002)
371. West, B., Deering, W.: Fractal physiology for physicists: Levy statistics. Phys. Rep. **246**(1), 1–100 (1994)
372. Tucker, G., Jackson, P., Storey, G., Holt, D.: Amiodarone disposition: polyexponential, power and gamma functions. Eur. J. Clin. Pharmacol. **26**(5), 655–656 (1984)
373. Phan, G., LeGall, B., Deverre, J., Fattal, E., Benech, H.: Predicting plutonium decorporation efficacy after intravenous administration of DTPA formulations: study of pharmacokinetic pharmacodynamic relationships in rats. Pharm. Res. **23**(9), 2030–2035 (2006)
374. Podlubny, I.: Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications. Mathematics in Science and Engineering, vol. 198. Academic, San Diego (1999)
375. Magin, R.: Fractional calculus in bioengineering, part 1. Crit. Rev. Biomed. Eng. **32**(1), 1–104 (2004)
376. Magin, R.: Fractional calculus in bioengineering, part 2. Crit. Rev. Biomed. Eng. **32**(2), 105–193 (2004)
377. Magin, R.: Fractional calculus in bioengineering, part 3. Crit. Rev. Biomed. Eng. **32**(3–4), 195–377 (2004)
378. Hennion, M., Hanert, E.: How to avoid unbounded drug accumulation with fractional pharmacokinetics. J. Pharmacokinet. Pharmacodyn. **40**(6), 691–700 (2013)
379. DeHoog, F., Knight, J., Stokes, A.: An improved method for numerical inversion of Laplace transforms. SIAM J. Sci. Stat. Comput. **3**(3), 357–366 (1982)
380. Dokoumetzidis, A., Macheras, P.: Fractional kinetics in drug absorption and disposition processes. J. Pharmacokinet. Pharmacodyn. **36**(2), 165–178 (2009)
381. Dokoumetzidis, A., Magin, R., Macheras, P.: A commentary on fractionalization of multicompartmental models. J. Pharmacokinet. Pharmacodyn. **37**(2), 203–207 (2010)

382. Dokoumetzidis, A., Magin, R., Macheras, P.: Fractional kinetics in multi-compartmental systems. J. Pharmacokinet. Pharmacodyn. **37**(5), 507–524 (2010)
383. Kilbas, A., Srivastava, H., Trujillo, J.: Theory and Applications of Fractional Differential Equations, vol. 204. Elsevier, Amsterdam (2006)
384. Moloni, S.: Applications of fractional calculus to pharmacokinetics. Ph.D. Thesis, University of Patras (2015)
385. Holt, D., Tucker, G., Jackson, P., Storey, G.: Amiodarone pharmacokinetics. Am. Heart J. **106**(4), 840–847 (1983)
386. Popovic, J., Atanackovic, M., Pilipovic, A., Rapaic, M., Pilipovic, S., Atanackovic, T.: A new approach to the compartmental analysis in pharmacokinetics: fractional time evolution of diclofenac. J. Pharmacokinet. Pharmacodyn. **37**(2), 119–134 (2010)
387. Popovic, J., Dolicanin, D., Rapaic, M., Popovic, S., Pilipovic, S., Atanackovic, T.: A nonlinear two compartmental fractional derivative model. Eur. J. Drug Metab. Pharmacokinet. **36**(4), 189–196 (2011)
388. Popovic, J., Posa, M., Popovic, K., Popovic, D., Milosevic, N., Tepavcevic, V.: Individualization of a pharmacokinetic model by fractional and nonlinear fit improvement. Eur. J. Drug Metab. Pharmacokinet. **38**(1), 69–76 (2013)
389. Popovic, J., Spasic, D., Tosic, J., Kolarovic, J., Malti, R., Mitic, I., Pilipovic, S., Atanackovic, T.: Fractional model for pharmacokinetics of high dose methotrexate in children with acute lymphoblastic leukaemia. Commun. Nonlinear Sci. Numer. Simul. **22**(1), 451–471 (2015)
390. Copot, D., Chevalier, A., Ionescu, C., DeKeyser, R.: A two-compartment fractional derivative model for propofol diffusion in anesthesia. In: 2013 IEEE International Conference on Control Applications (CCA), pp. 264–269. IEEE (2013)
391. Verotta, D.: Fractional dynamics pharmacokinetics pharmacodynamic models. J. Pharmacokinet. Pharmacodyn. **37**(3), 257–276 (2010)
392. CDER: Bioavailability and bioequivalence studies for orally administered drug products: general considerations. Technical Report, Center for Drug Evaluation and Research, Food and Drug Administration (2003)
393. Niazi, S.: Handbook of Bioequivalence Testing. Informa Healthcare, New York (2007)
394. Bois, F., Tozer, T., Hauck, W., Chen, M., Patnaik, R., Williams, R.: Bioequivalence: performance of several measures of extent of absorption. Pharm. Res. **11**(5), 715–722 (1994)
395. Bois, F., Tozer, T., Hauck, W., Chen, M., Patnaik, R., Williams, R.: Bioequivalence: performance of several measures of rate of absorption. Pharm. Res. **11**(7), 966–974 (1994)
396. El-Tahtawy, A., Jackson, A., Ludden, T.: Evaluation of bioequivalence of highly variable drugs using Monte Carlo simulations. I. Estimation of rate of absorption for single and multiple dose trials using Cmax. Pharm. Res. **12**(11), 1634–1641 (1995)
397. Endrenyi, L., Csizmadia, F., Tothfalusi, L., Balch, A., Chen, M.: The duration of measuring partial AUCs for the assessment of bioequivalence. Pharm. Res. **15**(3), 399–404 (1998)
398. Endrenyi, L., Csizmadia, F., Tothfalusi, L., Chen, M.: Metrics comparing simulated early concentration profiles for the determination of bioequivalence. Pharm. Res. **15**(8), 1292–1299 (1998)
399. Karalis, V., Macheras, P.: Pharmacodynamic considerations in bioequivalence assessment: comparison of novel and existing metrics. Eur. J. Pharm. Sci. **19**(1), 45–56 (2003)
400. Chen, M., Lesko, L., Williams, R.: Measures of exposure versus measures of rate and extent of absorption. Clin. Pharmacokinet. **40**(8), 565–572 (2001)
401. Jackson, A.: Determination of in vivo bioequivalence. Pharm. Res. **19**(3), 227–228 (2002)
402. Endrenyi, L., Tothfalusi, L.: Metrics for the evaluation of bioequivalence of modified-release formulations. AAPS J. **14**(4), 813–819 (2012)
403. Stier, E., Davit, B., Chandaroy, P., Chen, M., Fourie-Zirkelbach, J., Jackson, A., Kim, S., Lionberger, R., Mehta, M., Uppoor, R., Wang, Y., Yu, L., Conner, D.: Use of partial area under the curve metrics to assess bioequivalence of methylphenidate multiphasic modified release formulations. AAPS J. **14**(4), 925–926 (2012)
404. Basson, R., Cerimele, B., DeSante, K., Howey, D.: Tmax: an unconfounded metric for rate of absorption in single dose bioequivalence studies. Pharm. Res. **13**(2), 324–328 (1996)

405. Rostami-Hodjegan, A., Jackson, P., Tucker, G.: Sensitivity of indirect metrics for assessing "rate" in bioequivalence studies: moving the "goalposts" or changing the "game". J. Pharm. Sci. **83**(11), 1554–1557 (1994)
406. Chen, M.: An alternative approach for assessment of rate of absorption in bioequivalence studies. Pharm. Res. **9**(11), 1380–1385 (1992)
407. Endrenyi, L., Al-Shaikh, P.: Sensitive and specific determination of the equivalence of absorption rates. Pharm. Res. **12**(12), 1856–1864 (1995)
408. Lacey, L., Keene, O., Duquesnoy, C., Bye, A.: Evaluation of different indirect measures of rate of drug absorption in comparative pharmacokinetic studies. J. Pharm. Sci. **83**(2), 212–215 (1994)
409. Macheras, P., Symillides, M., Reppas, C.: The cutoff time point of the partial area method for assessment of rate of absorption in bioequivalence studies. Pharm. Res. **11**(6), 831–834 (1994)
410. Macheras, P., Symillides, M., Reppas, C.: An improved intercept method for the assessment of absorption rate in bioequivalence studies. Pharm. Res. **13**(11), 1755–1758 (1996)
411. Schall, R., Luus, H.: Comparison of absorption rates in bioequivalence studies of immediate release drug formulations. Int. J. Clin. Pharmacol. Ther. Toxicol. **30**(5), 153–159 (1992)
412. Tothfalusi, L., Endrenyi, L.: Without extrapolation, Cmax/AUC is an effective metric in investigations of bioequivalence. Pharm. Res. **12**(6), 937–942 (1995)
413. Chinchilli, V., Elswick, R.: The multivariate assessment of bioequivalence. J. Biopharm. Stat. **7**(1), 113–123 (1997)
414. Marston, S., Polli, J.: Evaluation of direct curve comparison metrics applied to pharmacokinetic profiles and relative bioavailability and bioequivalence. Pharm. Res. **14**(10), 1363–1369 (1997)
415. Polli, J., McLean, A.: Novel direct curve comparison metrics for bioequivalence. Pharm. Res. **18**(6), 734–741 (2001)
416. Rescigno, A.: Bioequivalence. Pharm. Res. **9**(7), 925–928 (1992)
417. Hauck, W., Anderson, S.: A new statistical procedure for testing equivalence in two-group comparative bioavailability trials. J. Pharmacokinet. Biopharm. **12**(1), 83–91 (1984)
418. Schuirmann, D.: A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. J. Pharmacokinet. Biopharm. **15**(6), 657–680 (1987)
419. Hauschke, D., Steinijans, V., Diletti, E.: A distribution-free procedure for the statistical analysis of bioequivalence studies. Int. J. Clin. Pharmacol. Ther. Toxicol. **28**(2), 72–78 (1990)
420. Hauschke, D., Steinijans, V., Diletti, E.: A distribution-free procedure for the statistical analysis of bioequivalence studies. Int. J. Clin. Pharmacol. Ther. Toxicol. **30**(Suppl. 1), S37–43 (1992)
421. Tucker, G., Rostami-Hodjekan, A., Jackson, P.: Bioequivalence - a measure of therapeutic equivalence? In: Blume, H., Midha, K. (eds.) Bio-International 2, pp. 35–43. Scientific publishers, Stuttgart (1995)
422. CDER: Statistical approaches to establishing bioequivalence. Technical Report, Center for Drug Evaluation and Research, Food and Drug Administration (2001)
423. Westlake, W.: The design and analysis of comparative blood-level trials. In: Swarbrick, J. (ed.) Current Concepts in the Pharmaceutical Sciences, Dosage Form Design and Bioavailability, pp. 149–179. Lea and Febiger, Philadelphia (1973)
424. Westlake, W.: Bioavailability and bioequivalence of pharmaceutical formulations. In: Peace, K. (ed.) Biopharmaceutical Statistics for Drug Development, pp. 329–352. Marcel Dekker, New York (1988)
425. Midha, K., Rawson, M., Hubbard, J.: bioequivalence: switchability and scaling. Eur. J. Pharm. Sci. **6**(2), 87–91 (1998)
426. Hauck, W., Anderson, S.: Types of bioequivalence and related statistical considerations. Int. J. Clin. Pharmacol. Ther. Toxicol. **30**(5), 181–187 (1992)
427. Schall, R., Luus, H.: On population and individual bioequivalence. Stat. Med. **12**(12), 1109–1124 (1993)

428. Chinchilli, V.: The assessment of individual and population bioequivalence. J. Biopharm. Stat. **6**(1), 1–14 (1996)
429. Anderson, S., Hauck, W.: Consideration of individual bioequivalence. J. Pharmacokinet. Biopharm. **18**(3,) 259–273 (1990)
430. Hauck, W., Anderson, S.: Individual bioequivalence: what matters to the patient. Stat. Med. **10**(6), 959–960 (1991)
431. Schall, R., Williams, R.: Towards a practical strategy for assessing individual bioequivalence. Food and Drug Administration Individual Bioequivalence Working Group. J. Pharmacokinet. Biopharm. **24**(1), 133–149 (1996)
432. Patnaik, R., Lesko, L., Chen, M., Williams, R.: Individual bioequivalence. New concepts in the statistical assessment of bioequivalence metrics. FDA Individual Bioequivalence Working Group. Clin. Pharmacokinet. **33**(1), 1–6 (1997)
433. Midha, K., Rawson, M., Hubbard, J.: Individual and average bioequivalence of highly variable drugs and drug products. J. Pharm. Sci. **86**(11), 1193–1197 (1997)
434. Endrenyi, L., Amidon, G., Midha, K., Skelly, J.: Individual bioequivalence: attractive in principle, difficult in practice. Pharm. Res. **15**(9), 1321–1325 (1998)
435. Chen, M., Patnaik, R., Hauck, W., Schuirmann, D., Hyslop, T., Williams, R.: An individual bioequivalence criterion: regulatory considerations. Stat. Med. **19**(20), 2821–2842 (2000)
436. Hyslop, T., Hsuan, F., Holder, D.: A small sample confidence interval approach to assess individual bioequivalence. Stat. Med. **19**(20), 2885–2897 (2000)
437. McGilveray, I., Midha, K., Skelly, J., Dighe, S., Doluisio, J., French, I., Karim, A., Burford, R.: Consensus report from "Bio International 89": Issues in the evaluation of bioavailability data. J. Pharm. Sci. **79**(10), 945–946 (1990)
438. Blume, H., Midha, K., Bio-International 2: Bioavailability, bioequivalence, and pharmacokinetic studies. J. Pharm. Sci. **82**(11), 1186–1189 (1993)
439. Blume, H., Elze, M., Potthast, H., Schug, B.: Practical strategies and design advantages in highly variable drug studies: multiple dose and replicate administration design. In: Blume, H., Midha, K. (eds.) Bio-International 2: Bioavailability, Bioequivalence, and Pharmacokinetic Studies, pp. 117–122. Medpharm, Stuttgart (1995)
440. Boddy, A., Snikeris, F., Kringle, R., Wei, G., Oppermann, J., Midha, K.: An approach for widening the bioequivalence acceptance limits in the case of highly variable drugs. Pharm. Res. **12**(12), 1865–1868 (1995)
441. Karalis, V., Macheras, P., Symillides, M.: Geometric mean ratio-dependent scaled bioequivalence limits with leveling-off properties. Eur. J. Pharm. Sci. **26**(1), 54–61 (2005)
442. Kytariolos, J., Karalis, V., Macheras, P., Symillides, M.: Novel scaled bioequivalence limits with leveling-off properties. Pharm. Res. **23**(11), 2657–2664 (2006)
443. Haidar, S., Makhlouf, F., Schuirmann, D., Hyslop, T., Davit, B., Conner, D., Yu, L.: Evaluation of a scaling approach for the bioequivalence of highly variable drugs. AAPS J. **10**(3), 450–454 (2008)
444. CHMP: Questions and answers: Positions on specific questions addressed to the pharmacokinetics working party. Technical Report, Committee for Medicinal Products for Human Use, European Medicines Agency (2014)
445. Haidar, S., Davit, B., Chen, M., Conner, D., Lee, L., Li, Q., Lionberger, R., Makhlouf, F., Patel, D., Schuirmann, D., Yu, L.: Bioequivalence approaches for highly variable drugs and drug products. Pharm. Res. **25**(1), 237–241 (2008)
446. Karalis, V., Symillides, M., Macheras, P.: Bioequivalence of highly variable drugs: a comparison of the newly proposed regulatory approaches by FDA and EMA. Pharm. Res. **29**(4), 1066–1077 (2012)
447. Tothfalusi, L., Endrenyi, L.: Sample sizes for designing bioequivalence studies for highly variable drugs. J. Pharm. Pharm. Sci. **15**(1), 73–84 (2012)
448. Shah, V., Yacobi, A., Barr, W., Benet, L., Breimer, D., Dobrinska, M., Endrenyi, L., Fairweather, W., Gillespie, W., Gonzalez, M., Hooper, J., Jackson, A., Lesko, L., Midha, K., Noonan, P., Patnaik, R., Williams, R.: Evaluation of orally administered highly variable drugs and drug formulations. Pharm. Res. **13**(11), 1590–1594 (1996)

449. Lachin, J.: Statistical properties of randomization in clinical trials. Control. Clin. Trials **9**(4), 289–311 (1988)
450. Demets, D.: Group sequential procedures: calendar versus information time. Stat. Med. **8**(10), 1191–1198 (1989)
451. Shih, W.,: Group sequential, sample size re-estimation and two-stage adaptive designs in clinical trials: a comparison. Stat. Med. **25**(6), 933–941 (2006)
452. Pong, A., Chow, S.: Handbook of Adaptive Designs in Pharmaceutical and Clinical Development. CRC/Taylor and Francis, Boca Raton (2011)
453. Pocock, S.: Size of cancer clinical trials and stopping rules. Br. J. Cancer **38**(6), 757–766 (1978)
454. Montague, T., Potvin, D., Diliberti, C., Hauck, W., Parr, A., Schuirmann, D.: Additional results for "Sequential design approaches for bioequivalence studies with crossover designs". Pharm. Stat. **11**(1), 8–13 (2012)
455. Potvin, D., DiLiberti, C., Hauck, W., Parr, A., Schuirmann, D., Smith, R.: Sequential design approaches for bioequivalence studies with crossover designs. Pharm. Stat. **7**(4), 245–262 (2008)
456. Fuglsang, A.: Sequential bioequivalence trial designs with increased power and controlled type I error rates. AAPS J. **15**(3), 659–661 (2013)
457. Fuglsang, A.: Futility rules in bioequivalence trials with sequential designs. AAPS J. **16**(1), 79–82 (2014)
458. Fuglsang, A.: Sequential bioequivalence approaches for parallel designs. AAPS J. **16**(3), 373–378 (2014)
459. Fuglsang, A.: A sequential bioequivalence design with a potential ethical advantage. AAPS J. **16**(4), 843–846 (2014)
460. Karalis, V.: The role of the upper sample size limit in two-stage bioequivalence designs. Int. J. Pharm. **456**(1), 87–94 (2013)
461. Karalis, V., Macheras, P.: An insight into the properties of a two-stage design in bioequivalence studies. Pharm. Res. **30**(7), 1824–1835 (2013)
462. Karalis, V., Macheras, P.: On the statistical model of the two-stage designs in bioequivalence assessment. J. Pharm. Pharmacol. **66**(1), 48–52 (2014)
463. Diletti, E., Hauschke, D., Steinijans, V.: Sample size determination for bioequivalence assessment by means of confidence intervals. Int. J. Clin. Pharmacol. Ther. Toxicol. **29**(1), 1–8 (1991)
464. Diletti, E., Hauschke, D., Steinijans, V.: Sample size determination for bioequivalence assessment by means of confidence intervals. Int. J. Clin. Pharmacol. Ther. Toxicol. **30** Suppl 1, S51–58 (1992)
465. Hauschke, D., Steinijans, V., Diletti, E., Burke, M.: Sample size determination for bioequivalence assessment using a multiplicative model. J. Pharmacokinet. Biopharm. **20**(5), 557–561 (1992)
466. Phillips, K.: Power of the two one-sided tests procedure in bioequivalence. J. Pharmacokinet. Biopharm. **18**(2), 137–144 (1990)
467. Liu, J., Chow, S.: Sample size determination for the two one-sided tests procedure in bioequivalence. J. Pharmacokinet. Biopharm. **20**(1), 101–104 (1992)
468. Liu, J., Weng, C.: Estimation of direct formulation effect under log-normal distribution in bioavailability/bioequivalence studies. Stat. Med. **11**(7), 881–896 (1992)
469. Julious, S.: Sample sizes for clinical trials with normal data. Stat. Med. **23**(12), 1921–1986 (2004)
470. Midha, K., Rawson, M., Hubbard, J.: The role of metabolites in bioequivalence. Pharm. Res. **21**(8), 1331–1344 (2004)
471. Midha, K., Shah, V., Singh, G., Patnaik, R.: Conference report: bio-International 2005. J. Pharm. Sci. **96**(4), 747–754 (2007)
472. Houston, J.: Importance of metabolites in bioequivalence. In: McGilveray, J., Dighe, S. (eds.) Proceedings Bio-International 89: Issues in the Evaluation of Bioavailability Data, pp. 99–100. Toronto (1989)

473. Jackson, A., Robbie, G., Marroum, P.: Metabolites and bioequivalence: past and present. Clin. Pharmacokinet. **43**(10), 655–672 (2004)
474. Midha, K., Hubbard, J., McKay, G., Hawes, E., Hsia, D.: The role of metabolites in a bioequivalence study 1: loxapine, 7-hydroxyloxapine and 8-hydroxyloxapine. Int. J. Clin. Pharmacol. Ther. Toxicol. **31**(4), 177–183 (1993)
475. Midha, K., Hubbard, J., McKay, G., Rawson, M., Hsia, D.: The role of metabolites in a bioequivalence study II: amoxapine, 7-hydroxyamoxapine, and 8-hydroxyamoxapine. Int. J. Clin. Pharmacol. Ther. Toxicol. **37**(9), 428–438 (1999)
476. Chen, M., Jackson, A.: The role of metabolites in bioequivalency assessment. I. Linear pharmacokinetics without first-pass effect. Pharm. Res. **8**(1), 25–32 (1991)
477. Chen, M., Jackson, A.: The role of metabolites in bioequivalency assessment. II. Drugs with linear pharmacokinetics and first-pass effect. Pharm. Res. **12**(5), 700–708 (1995)
478. Rosenbaum, S., Lam, J.: Bioequivalence parameters of parent drug and its first-pass metabolite: comparative sensitivity to sources of pharmacokinetic variability. Drug Dev. Ind. Pharm. **23**, 337–344 (1997)
479. Jackson, A.: The role of metabolites in bioequivalency assessment. III. Highly variable drugs with linear kinetics and first-pass effect. Pharm. Res. **17**(11), 1432–1436 (2000)
480. Braddy, A., Jackson, A.: Role of metabolites for drugs that undergo nonlinear first-pass effect: impact on bioequivalency assessment using single-dose simulations. J. Pharm. Sci. **99**(1), 515–523 (2010)
481. Fernandez-Teruel, C., Gonzalez-Alvarez, I., Navarro-Fontestad, C., Garcia-Arieta, A., Bermejo, M., Casabo, V.: Computer simulations of bioequivalence trials: selection of design and analyte in BCS drugs with first-pass hepatic metabolism: Part II. Non-linear kinetics. Eur. J. Pharm. Sci. **36**(1), 147–156 (2009)
482. Fernandez-Teruel, C., Nalda-Molina, R., Gonzalez-Alvarez, I., Navarro-Fontestad, C., Garcia-Arieta, A., Casabo, V., Bermejo, M.: Computer simulations of bioequivalence trials: selection of design and analyte in BCS drugs with first-pass hepatic metabolism: linear kinetics (I). Eur. J. Pharm. Sci. **36**(1), 137–146 (2009)
483. Karalis, V., Macheras, P.: Examining the role of metabolites in bioequivalence assessment. J. Pharm. Pharm. Sci. **13**(2), 198–217 (2010)
484. Gwaza, L., Gordon, J., Welink, J., Potthast, H., Hansson, H., Stahl, M., Garcia-Arieta, A.: Statistical approaches to indirectly compare bioequivalence between generics: a comparison of methodologies employing artemether/lumefantrine 20/120 mg tablets as prequalified by WHO. Eur. J. Clin. Pharmacol. **68**(12), 1611–1618 (2012)
485. Herranz, M., Morales-Alcelay, S., Corredera-Hernandez, M., DeLaTorre-Alvarado, J., Blazquez-Perez, A., Suarez-Gea, M., Alvarez, C., Garcia-Arieta, A.: Bioequivalence between generic tacrolimus products marketed in Spain by adjusted indirect comparison. Eur. J. Clin. Pharmacol. **69**(5), 1157–1162 (2013)
486. Chow, S., Liu, J.: Meta-analysis for bioequivalence review. J. Biopharm. Stat. **7**(1), 97–111 (1997)
487. Chow, S., Shao, J.: Bioequivalence review for drug interchangeability. J. Biopharm. Stat. **9**(3), 485–497 (1999)
488. Karalis, V., Bialer, M., Macheras, P.: Quantitative assessment of the switchability of generic products. Eur. J. Pharm. Sci. **50**(3–4), 476–483 (2013)
489. Karalis, V., Macheras, P., Bialer, M.: Generic products of antiepileptic drugs: a perspective on bioequivalence, bioavailability, and formulation switches using Monte Carlo simulations. CNS Drugs **28**(1), 69–77 (2014)
490. Panhard, X., Mentre, F.: Evaluation by simulation of tests based on non-linear mixed-effects models in pharmacokinetic interaction and bioequivalence cross-over trials. Stat. Med. **24**(10), 1509–1524 (2005)
491. Dubois, A., Gsteiger, S., Pigeolet, E., Mentre, F.: Bioequivalence tests based on individual estimates using non-compartmental or model-based analyses: evaluation of estimates of sample means and type I error for different designs. Pharm. Res. **27**(1), 92–104 (2010)

492. Maier, G., Lockwood, G., Oppermann, J., Wei, G., Bauer, P., Fedler-Kelly, J., Grasela, T.: Characterization of the highly variable bioavailability of tiludronate in normal volunteers using population pharmacokinetic methodologies. Eur. J. Drug Metab. Pharmacokinet. **24**(3), 249–254 (1999)
493. Velasque, L., Estrela, R., Suarez-Kurtz, G., Struchiner, C.: A new model for the population pharmacokinetics of didanosine in healthy subjects. Braz. J. Med. Biol. Res. **40**(1), 97–104 (2007)
494. Stanhope, R., Sorgel, F., Gravel, P., Pannatier-Schuetz, Y., Zabransky, M., Muenzberg, M.: Bioequivalence studies of omnitrope, the first biosimilar/rhGH follow-on protein: two comparative phase 1 randomized studies and population pharmacokinetic analysis. J. Clin. Pharmacol. **50**(11), 1339–1348 (2010)
495. Seng-Yue, C., Gallicano, K., Labbe, L., Ducharme, M.: Novel population pharmacokinetic method compared to the standard noncompartmental approach to assess bioequivalence of iron gluconate formulations. J. Pharm. Pharm. Sci. **16**(3), 424–440 (2013)
496. Soulele, K., Macheras, P., Silvestro, L., Rizea-Savu, S., Karalis, V.: Population pharmacokinetics of fluticasone propionate/salmeterol using two different dry powder inhalers. Eur. J. Pharma. Sci. **80**, 33–42 (2015)
497. Zelenetz, A., Ahmed, I., Braud, E., Cross, J., Davenport-Ennis, N., Dickinson, B., Goldberg, S., Gottlieb, S., Johnson, P., Lyman, G., Markus, R., Matulonis, U., Reinke, D., Li, E., DeMartino, J., Larsen, J., Hoffman, J.: NCCN biosimilars white paper: regulatory, scientific, and patient safety perspectives. J. Natl. Compr. Cancer Netw. **9**(Suppl. 4), S1–22 (2011)
498. Schellekens, H.: Bioequivalence and the immunogenicity of biopharmaceuticals. Nat. Rev. Drug Discov. **1**(6), 457–462 (2002)
499. Schellekens, H.: Biosimilar therapeutic agents: issues with bioequivalence and immunogenicity. Eur. J. Clin. Invest. **34**(12), 797–799 (2004)
500. Chow, S., Wang, J., Endrenyi, L., Lachenbruch, P.: Scientific considerations for assessing biosimilar products. Stat. Med. **32**(3), 370–381 (2013)
501. Zhang, N., Yang, J., Chow, S., Endrenyi, L., Chi, E.: Impact of variability on the choice of biosimilarity limits in assessing follow-on biologics. Stat. Med. **32**(3), 424–433 (2013)
502. Endrenyi, L., Chang, C., Chow, S., Tothfalusi, L.: On the interchangeability of biologic drug products. Stat. Med. **32**(3), 434–441 (2013)
503. Chow, S., Endrenyi, L., Lachenbruch, P.: Comments on the FDA draft guidance on biosimilar products. Stat. Med. **32**(3), 364–369 (2013)
504. Tothfalusi, L., Endrenyi, L., Chow, S.: Statistical and regulatory considerations in assessments of interchangeability of biological drug products. Eur. J. Health Econ. **15**(Suppl. 1), S5–11 (2014)
505. Chen, X., Hickling, T., Kraynov, E., Kuang, B., Parng, C., Vicini, P.: A mathematical model of the effect of immunogenicity on therapeutic protein pharmacokinetics. AAPS J. **15**(4), 1141–1154 (2013)
506. Perez-Ruixo, J., Ma, P., Chow, A.: The utility of modeling and simulation approaches to evaluate immunogenicity effect on the therapeutic protein pharmacokinetics. AAPS J. **15**(1), 172–182 (2013)
507. Driscoll, R., Zhou, L., Moxness, M., Mytych, D., Chirmule, N., Jawa, V.: Statistical and bioanalytical considerations for establishing a depletion criterion for specificity testing during immunogenicity assessment of a biotherapeutic. AAPS J. **15**(4), 1160–1167 (2013)
508. Matis, J., Hartley, H.: Stochastic compartmental analysis: model and least squares estimation from time series data. Biometrics **27**, 77–102 (1971)
509. Matis, J.: Gamma time-dependency in Blaxter's compartmental model. Biometrics **28**, 597–602 (1972)
510. Takur, A., Rescigno, A., Schafer, D.: On the stochastic theory of compartments. II. Multicompartment systems. Bull. Math. Biophys. **35**, 263–271 (1973)
511. Purdue, P.: Stochastic theory of compartments. Bull. Math. Biol. **36**, 305–309 (1974)
512. Purdue, P.: Stochastic theory of compartments: one and two compartment systems. Bull. Math. Biol. **36**, 577–587 (1974)

513. Epperson, J., Matis, J.: On the distribution of the general irreversible n-compartmental model having time-dependent transition probabilities. Bull. Math. Biol. **41**, 737–749 (1979)
514. Chiang, C.: Introduction to Stochastic Processes and Their Applications. Krieger, New York (1980)
515. Mehata, K., Selvam, D.: A stochastic model for the n-compartment irreversible system. Bull. Math. Biol. **43**(5), 549–561 (1981)
516. Weiss, M.: Theorems on log-convex disposition curves in drug and tracer kinetics. J. Theor. Biol. **116**, 355–368 (1985)
517. Weiss, M.: Generalizations in linear pharmacokinetics using properties of certain classes of residence time distributions. I. Log-convex drug disposition curves. J. Pharmacokinet. Biopharm. **14**(6), 635–657 (1986)
518. Weiss, M.: Generalizations in linear pharmacokinetics using properties of certain classes of residence time distributions. II. Log-concave concentration-time curves following oral administration. J. Pharmacokinet. Biopharm. **15**(1), 57–74 (1987)
519. Agrafiotis, G.: On the stochastic theory of compartments: a semi-Markov approach for the analysis of the k-compartmental systems. Bull. Math. Biol. **44**(6), 809–817 (1982)
520. Schalla, M., Weiss, M.: Pharmacokinetic curve fitting using numerical inverse Laplace transformation. Eur. J. Pharm. Sci. **7**, 305–309 (1999)
521. Cutler, D.: A linear recirculation model for drug disposition. J. Pharmacokinet. Biopharm. **7**(1), 101–116 (1979)
522. Yamaoka, K., Kanba, M., Toyoda, Y., Yano, Y., Nakagawa, T.: Analysis of enterohepatic circulation of ceficime in rat by fast inverse Laplace transform (FILT). J. Pharmacokinet. Biopharm. **18**(6), 545–559 (1990)
523. Gallo, J., Lam, F., Perrier, D.: Area method for the estimation of partition coefficients for physiological pharmacokinetic models. J. Pharmacokinet. Biopharm. **15**(3), 271–280 (1987)
524. Veng-Pedersen, P.: Linear and nonlinear system approaches in pharmacokinetics: how much do they offer. I. General considerations. J. Pharmacokinet. Biopharm. **16**(4), 413–472 (1988)
525. Weiss, M.: Moments of physiological transit time distributions and the time course of drug disposition in the body. J. Math. Biol. **15**, 305–318 (1982)
526. Pipes, L., Harvill, L.: Applied Mathematics for Engineers and Scientists. Pure and Applied Mathematics, 3rd edn. McGraw Hill, Kogakusha (1970)
527. Duffy, D.: On the numerical inversion of Laplace transforms: comparison of three new methods on characteristic problems from applications. ACM Trans. Math. Softw. **19**(3), 333–359 (1993)
528. Durisova, M., Dedik, L., Balan, M.: Building a structured model of a complex pharmacokinetic system with time delays. Bull. Math. Biol. **57**(6), 787–808 (1995)
529. Purdue, P.: Stochastic theory of compartments: an open, two-compartment, reversible system with independent Poisson arrivals. Bull. Math. Biol. **37**, 269–275 (1975)
530. Neuts, M.: Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach. Johns Hopkins University Press, Baltimore (1981)
531. Faddy, M.: Compartmental models with phase-type residence-time distributions. Appl. Stoch. Model Data Anal. **6**, 121–127 (1990)
532. Faddy, M.: A structured compartmental model for drug kinetics. Biometrics **49**(1), 243–248 (1993)
533. Johnson, M., Taaffe, M.: Matching moments to phase distributions: mixtures of Erlang distributions of common order. Commun. Stat. Stoch. Model. **5**, 711–743 (1989)
534. Ramaswami, V., Latouche, G.: An experimental evaluation of the matrix-geometric method for the G1/PH/1 queue. Commun. Stat. Stoch. Model. **5**, 629–667 (1989)
535. Karalis, V., Claret, L., Iliadis, A., Macheras, P.: Fractal volume of drug distribution: it scales proportionally to body mass. Pharm. Res. **18**(7), 1056–1060 (2001)
536. Matis, J., Wehrly, T.: On the use of residence time moments in the statistical analysis of age-department stochastic compartmental systems. In: Capasso, V., Grosso, E., Paveri-Fontana, S. (eds.) Mathematics in Biology and Medicine, pp. 386–398. Springer, New York (1985)

537. McInnis, B., El-Asfouri, S., Kapadia, S.: On stochastic compartmental modeling. Bull. Math. Biol. **41**, 611–613 (1979)
538. Rescigno, A., Matis, J.: On the relevance of stochastic compartmental models to pharmacokinetic systems. Bull. Math. Biol. **43**(2), 245–247 (1981)
539. Seber, G., Wild, C.: Nonlinear Regression Analysis. Wiley Series in Probability and Mathematical Statistics. Wiley, New York (1989)
540. Kodell, R., Matis, J.: Estimating the rate constants in a two-compartment stochastic model. Biometrics **32**(2), 377–390 (1976)
541. Piotrovskii, V.: Pharmacokinetic stochastic model with Weibull-distributed residence times of drug molecules in the body. Eur. J. Clin. Pharmacol. **32**, 515–523 (1987)
542. Matis, J., Tolley, H.: Compartmental models with multiple sources of stochastic variability: the one compartment, time invariant hazard rate case. Bull. Math. Biol. **41**, 491–515 (1979)
543. Matis, J., Tolley, H.: On the stochastic modeling of tracer kinetics. Fed. Proc. **39**, 104–109 (1980)
544. Rao, C.: Linear Statistical Inference and Its Applications. General Probability and Mathematical Statistics, 2nd edn. Wiley, New York (2002)
545. Matis, J., Wehrly, T., Ellis, W.: Some generalized stochastic compartment models for digesta flow. Biometrics **45**, 703–720 (1989)
546. Wing, G.: A Primer on Integral Equations of the First Kind: The Problem of Deconvolution and Unfolding. Society for Industrial and Applied Mathematics, Philadelphia (1991)
547. Charter, M., Gull, S.: Maximum entropy and drug absorption. J. Pharmacokinet. Biopharm. **19**(5), 497–520 (1991)
548. Gillespie, D.: Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. **81**(25), 2340–2361 (1977)
549. BenAvraham, D., Havlin, S.: Diffusion and Reactions in Fractals and Disordered Systems. Cambridge University Press, Cambridge (2000)
550. McQuarrie, D.: Stochastic approach to chemical kinetics. J. Appl. Probab. **4**, 413–478 (1967)
551. Gardenas, M., Matis, J.: On the time-dependent reversible stochastic compartmental model. II. A class of n-compartment systems. Bull. Math. Biol. **37**, 555–564 (1975)
552. Matis, J., Kiffe, T.: Stochastic Population Models. A Compartmental Perspective. Lecture Notes in Statistics, vol. 145. Springer, New York (2000)
553. Renshaw, E.: Saddlepoint approximations for stochastic processes with truncated cumulant generating functions. J. Math. Appl. Med. Biol. **15**, 1–12 (1998)
554. Bailey, N.: Elements of Stochastic Processes with Applications to the Natural Sciences. General Applied Probability and Statistics. Wiley, New York (1964)
555. Gillespie, D.: A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys. **22**, 403–434 (1976)
556. Gibson, M., Bruck, J.: Efficient exact stochastic simulation of chemical systems with many species and many channels. J. Phys. Chem. **104**(9), 1876–1889 (2000)
557. Gillespie, D.: Approximate accelerated stochastic simulation of chemically reacting systems. J. Chem. Phys. **115**(4), 1716–1733 (2001)
558. Stundzia, A., Lumsden, C.: Stochastic simulation of coupled reaction-diffusion processes. J. Comput. Phys. **127**(1), 196–207 (1996)
559. Oppenheim, I., Shuler, K., Weiss, G.: Stochastic Processes in Chemical Physics: The Master Equation. MIT, Cambridge (1977)
560. Matheson, I., Walls, D., Gardiner, C.: Stochastic models of first-order nonequilibrium phase transitions in chemical reactions. J. Stat. Phys. **12**, 21–34 (1975)
561. Liebovitch, L., Fischbarg, J., Koniarek, J.: Ion channel kinetics: a model based on fractal scaling rather than multistate Markov processes. Math. Biosci. **84**(1), 37–68 (1987)
562. Shooman, M.: Probabilistic Reliability: An Engineering Approach. McGraw-Hill Electrical and Electronic Engineering Series. McGraw-Hill, New York (1968)
563. Mann, N., Schafer, R., Singpurwalla, N.: Methods for Statistical Analysis of Reliability and Life Data. Wiley, New York (1974)

564. Billinton, R., Allan, R.: Reliability Evaluation of Engineering Systems. Concepts and Techniques, 2nd edn. Plenum, New York (1992)
565. Kalbfleisch, J., Prentice, R.: The Statistical Analysis of Failure Time Data. Wiley Series in Probability and Statistics, 2nd edn. Wiley, Hoboken (2002)
566. Singer, K.: Application of the theory of stochastic processes to the study of irreproducible chemical reactions and nucleation processes. J. R. Stat. Soc. **15**, 92–106 (1953)
567. McAdams, H., Arkin, A.: Stochastic mechanisms in gene expression. Proc. Natl. Acad. Sci. U.S.A. **94**, 814–819 (1997)
568. Lee, J., Hulse, J., Colburn, W.: Surrogate biochemical markers: precise measurement for strategic drug and biologics development. J. Clin. Pharmacol. **35**(5), 464–470 (1995)
569. McCullagh, P.: Regression models for ordinal data. J. R. Stat. Soc. **42**(2), 109–142 (1980)
570. Black, J.: Receptors as pharmaceutical targets. In: Foreman, J., Johansen, T. (eds.) Textbook of Receptor Pharmacology, 2nd edn., pp. 271–278. CRC, New York (2002)
571. Gieschke, R., Reigner, B., Steimer, J.: Exploring clinical study design by computer simulation based on pharmacokinetic/pharmacodynamic modelling. Int. J. Clin. Pharmacol. Ther. **55**(10), 469–474 (1997)
572. Derendorf, H., Meibohm, B.: Modeling of pharmacokinetic/pharmacodynamic (PK/PD) relationships: concepts and perspectives. Pharm. Res. **16**(2), 176–185 (1999)
573. Hill, A.: The mode of action of nicotine and curari, determined by the form of the contraction curve and the method of temperature coefficients. J. Physiol. **39**, 361–373 (1909)
574. Clark, A.: Applied Pharmacology, 5th edn. Churchill, London (1933)
575. Clark, A.: General Pharmacology. Handbook of Experimental Pharmacology Series, vol. 4. Springer, Berlin (1937)
576. Black, J., Leff, P.: Operational models of pharmacological agonism. Proc. R. Soc. Lond. B. Biol. Sci. **220**(1219), 141–162 (1983)
577. Kenakin, T.: Principles: receptor theory in pharmacology. Trends Pharmacol. Sci. **25**(4), 186–192 (2004)
578. Visser, S., Wolters, F., Gubbens-Stibbe, J., Tukker, E., VanDerGraaf, P., Peletier, L., Danhof, M.: Mechanism-based pharmacokinetic/pharmacodynamic modeling of the electroencephalogram effects of GABAA receptor modulators: in vitro-in vivo correlations. J. Pharmacol. Exp. Ther. **304**(1), 88–101 (2003)
579. Levy, G.: Relationship between elimination rate of drugs and rate of decline of their pharmacological effects. J. Pharm. Sci. **53**, 342–343 (1964)
580. Levy, G.: Kinetics of pharmacologic effects. Clin. Pharmacol. Ther. **7**, 362–372 (1966)
581. Wagner, J., Aghajanian, G., Bing, O.: Correlation of performance test scores with 'tissue concentration' of lysergic acid diethylamide in human subjects. Clin. Pharmacol. Ther. **9**(5), 635–638 (1968)
582. Segre, G.: Kinetics of interaction between drugs and biological systems. Il Farmaco **23**, 907–918 (1968)
583. Sheiner, L., Stanski, D., Vozeh, S., Miller, R., Ham, J.: Simultaneous modeling of pharmacokinetics and pharmacodynamics: application to d-tubocurarine. Clin. Pharmacol. Ther. **25**(3), 358–371 (1979)
584. Holford, N., Sheiner, L.: Kinetics of pharmacologic response. Pharmacol. Ther. **16**, 143–166 (1982)
585. Racine-Poon, A., Botta, L., Chang, T., Davis, F., Gygax, D., Liou, R., Rohane, P., Staehelin, T., VanSteijn, A., Frank, W.: Efficacy, pharmacodynamics, and pharmacokinetics of CGP 51901, an anti-immunoglobulin E chimeric monoclonal antibody, in patients with seasonal allergic rhinitis. Clin. Pharmacol. Ther. **62**(6), 675–690 (1997)
586. D'Hollander, A., Delcroix, C.: An analytical pharmacodynamic model for nondepolarizing neuromuscular blocking agents. J. Pharmacokinet. Biopharm. **9**(1), 27–40 (1981)
587. Perry, S., Inturrisi, C.: Analgesia and morphine disposition in burn patients. J. Burn Care Rehabil. **4**(4), 276–279 (1983)
588. Inturrisi, C., Colburn, W., Kaiko, R., Houde, R., Foley, K.: Pharmacokinetics and pharmacodynamics of methadone in patients with chronic pain. Clin. Pharmacol. Ther. **41**(4), 392–401 (1987)

589. Scott, J., Cooke, J., Stanski, D.: Electroencephalographic quantitation of opioid effect: comparative pharmacodynamics of fentanyl and sufentanil. Anesthesiology **74**(1), 34–42 (1991)
590. Stanski, D., Hudson, R., Homer, T., Saidman, L., Meathe, E.: Pharmacodynamic modeling of thiopental anesthesia. J. Pharmacokinet. Biopharm. **12**(2), 223–240 (1984)
591. Homer, T., Stanski, D.: The effect of increasing age on thiopental disposition and anesthetic requirement. Anesthesiology **62**(6), 714–724 (1985)
592. Dingemanse, J., Haussler, J., Hering, W., Ihmsen, H., Albrecht, S., Zell, M., Schwilden, H., Schuttler, J.: Pharmacokinetic-pharmacodynamic modelling of the EEG effects of Ro 48-6791, a new short-acting benzodiazepine, in young and elderly subjects. Br. J. Anaesth. **79**(5), 567–574 (1997)
593. Salazar, D., Much, D., Nichola, P., Seibold, J., Shindler, D., Slugg, P.: A pharmacokinetic-pharmacodynamic model of d-sotalol Q-Tc prolongation during intravenous administration to healthy subjects. J. Clin. Pharmacol. **37**(9), 799–809 (1997)
594. Derendorf, H., Hochhaus, G.: Handbook of Pharmacokinetic/Pharmacodynamic Correlation. CRC, Boca Raton (1995)
595. Iliadis, A., Cano, J.: A new pharmacodynamic modelling approach based on receptors theory. In: Aiache, J., Hirtz, J. (eds.) 3th European Congress of Biopharmaceutics and Pharmacokinetics, vol. 3, pp. 65–74. Imprimerie de l'Universite, Clermont-Ferrand/Freiburg (1987)
596. Shimada, S., Nakajima, Y., Yamamoto, K., Sawada, Y., Iga, T.: Comparative pharmacodynamics of eight calcium channel blocking agents in Japanese essential hypertensive patients. Biol. Pharm. Bull. **19**(3), 430–437 (1996)
597. Karlsson, M., Molnar, V., Bergh, J., Freijs, A., Larsson, R.: A general model for time-dissociated pharmacokinetic-pharmacodynamic relationships exemplified by paclitaxel myelosuppression. Clin. Pharmacol. Ther. **63**, 11–25 (1998)
598. Dayneka, N., Garg, V., Jusko, W.: Comparison of four basic models of indirect pharmacodynamic response. J. Pharmacokinet. Biopharm. **21**(4), 457–478 (1993)
599. Ariens, E.: Affinity and intrinsic activity in the theory of competitive inhibition. I. Problems and theory. Archives Internationales de Pharmacodynamie et de Therapie **99**(1), 32–49 (1954)
600. Nagashima, R., O'Reilly, R., Levy, G.: Kinetics of pharmacologic effects in man: the anticoagulant action of warfarin. Clin. Pharmacol. Ther. **10**, 22–35 (1969)
601. Jusko, W., Ko, H.: Physiologic indirect response models characterize diverse types of pharmacodynamic effects. Clin. Pharmacol. Ther. **56**(4), 406–419 (1994)
602. Mager, D., Wyska, E., Jusko, W.: Diversity of mechanism-based pharmacodynamic models. Drug Metab. Dispos. **31**(5), 510–518 (2003)
603. Sharma, A., Ebling, W., Jusko, W.: Precursor-dependent indirect pharmacodynamic response model for tolerance and rebound phenomena. J. Pharm. Sci. **87**(12), 1577–1584 (1998)
604. Lew, K., Ludwig, E., Milad, M., Donovan, K., Middleton, E., Ferry, J., Jusko, W.: Gender-based effects on methylprednisolone pharmacokinetics and pharmacodynamics. Clin. Pharmacol. Ther. **54**(4), 402–414 (1993)
605. Chakraborty, A., Krzyzanski, W., Jusko, W.: Mathematical modeling of circadian cortisol concentrations using indirect response models: comparison of several methods. J. Pharmacokinet. Biopharm. **27**(1), 23–43 (1999)
606. Li, H., Meno-Tetang, G., Chiba, K., Arima, N., Heining, P., Jusko, W.: Pharmacokinetics and cell trafficking dynamics of 2-Amino-2-[2-(4-octylphenyl)ethyl]propane-1,3-diol Hydrochloride (FTY720) in cynomolgus monkeys after single oral and intravenous doses. J. Pharmacol. Exp. Ther. **301**(2), 519–526 (2002)
607. Wakelkamp, M., Alvan, G., Paintaud, G.: The time of maximum effect for model selection in pharmacokinetic-pharmacodynamic analysis applied to frusemide. Br. J. Clin. Pharmacol. **45**(1), 63–70 (1998)
608. Trzeciakowski, J.: Analysis of stimulus-response chains using nonlinear dynamics. J. Pharmacol. Toxicol. Methods **36**(2), 103–121 (1996)

609. Yamamoto, K., Abe, M., Katashima, M., Yamada, Y., Sawada, Y., Iga, T.: Pharmacodynamic analysis of antiplatelet effect of aspirin in the literature. Modeling based on inhibition of cyclooxygenase in the platelet and the vessel wall endothelium. Jpn. J. Hosp. Pharm. **22**, 133–141 (1996)
610. Katashima, M., Yamamoto, K., Tokuma, Y., Hata, T., Sawada, Y., Iga, T.: Pharmacokinetic and pharmacodynamic study of a new nonsteroidal 5 alpha-reductase inhibitor, 4-[3-[3-[Bis(4-isobutylphenyl)methylamino]benzoyl]-1H-indol-1-yl]-butyric acid, in rats. J. Pharmacol. Exp. Ther. **284**(3), 914–920 (1998)
611. Abelo, A., Eriksson, U., Karlsson, M., Larsson, H., Gabrielsson, J.: A turnover model of irreversible inhibition of gastric acid secretion by omeprazole in the dog. J. Pharmacol. Exp. Ther. **295**(2), 662–669 (2000)
612. Mouton, J., Vinks, A., Punt, N.: Pharmacokinetic-pharmacodynamic modeling of activity of ceftazidime during continuous and intermittent infusion. Antimicrob. Agents Chemother. **41**(4), 733–738 (1997)
613. Jusko, W.: Pharmacodynamics of chemotherapeutic effects: dose-time-response relationships for phase-nonspecific agents. J. Pharm. Sci. **60**(6), 892–895 (1971)
614. Zhi, J., Nightingale, C., Quintiliani, R.: Microbial pharmacodynamics of piperacillin in neutropenic mice of systematic infection due to Pseudomonas aeruginosa. J. Pharmacokinet. Biopharm. **16**(4), 355–375 (1988)
615. Meibohm, B., Hochhaus, G., Derendorf, H.: Time dependency of the pharmacologic response to glucocorticoids. Clin. Pharmacol. Ther. **61**(2), 155 (PI–71) 1997
616. Bauer, J., Fung, H.: Pharmacodynamic models of nitroglycerin-induced hemodynamic tolerance in experimental heart failure. Pharm. Res. **11**(6), 816–823 (1994)
617. Kenakin, T.: Pharmacological Analysis of Drug-Receptor Interaction. Lippincott-Raven, New York (1997)
618. VanDenMeiracker, A., ManIn'tVeld, A., Boomsma, F., Fischberg, D., Molinoff, P., and Schalekamp, M.: Hemodynamic and beta-adrenergic receptor adaptations during long-term beta-adrenoceptor blockade. Studies with acebutolol, atenolol, pindolol, and propranolol in hypertensive patients. Circulation **80**(4), 903–914 (1989)
619. Sheiner, L., Verotta, D.: Further notes on physiologic indirect response models. Clin. Pharmacol. Ther. **58**(2), 238–240 (1995)
620. Tallarida, R.: Further characterization of a control model for ligand-receptor interaction: phase plane geometry, stability, and oscillation. Ann. Biomed. Eng. **18**, 671–684 (1990)
621. Paalzow, L.: Integrated pharmacokinetic-dynamic modeling of drugs acting on the CNS. Drug Metab. Rev. **15**, 383–400 (1984)
622. Perez-Urizar, J., Granados-Soto, V., Flores-Murrieta, F., Castaneda-Hernandez, G.: Pharmacokinetic-pharmacodynamic modeling: why? Arch. Med. Res. **31**, 539–545 (2000)
623. Gabrielsson, J., Weiner, D.: Methodology for pharmacokinetic/pharmacodynamic data analysis. Pharm. Sci. Technol. Today **2**(6), 244–252 (1999)
624. Meibohm, B., Derendorf, H.: Pharmacokinetic/pharmacodynamic studies in drug product development. J. Pharm. Sci. **91**(1), 18–31 (2002)
625. Meille, C., Iliadis, A., Barbolosi, D., Frances, N., Freyer, G.: An interface model for dosage adjustment connects hematotoxicity to pharmacokinetics. J. Pharmacokinet. Pharmacodyn. **35**(6), 619–633 (2008)
626. Abdi, M.: Best practice in toxicological pathology. Spontaneous and induced histopathological lesions in preclinical studies. In: Faqi, A. (ed.) A Comprehensive Guide to Toxicology in Preclinical Drug Development, pp. 231–236. Academic, London (2013)
627. Milton, J., VanDerHeiden, U., Longtin, A., Mackey, M.: Complex dynamics and noise in simple neural networks with delayed mixed feedback. Biomed. Biochim. Acta **49**(8–9), 697–707 (1990)
628. Tallarida, R.: Control and oscillation in ligand receptor interactions according to the law of mass action. Life Sci. **46**(22), 1559–1568 (1990)
629. Tallarida, R.: On stability and control of ligand-receptor interactions according to the mass action law: a theoretical model of chaos. Drug Dev. Res. **19**, 257–274 (1990)

630. Tallarida, R., Freeman, K.: Control and stability of ligand receptors interaction in the presence of competitive compound. Life Sci. **48**(6), PL19–PL24 (1991)
631. Tallarida, R., Freeman, K.: Chaos and control in mass-action binding of endogeneous compounds. Ann. Biomed. Eng. **22**, 153–161 (1994)
632. Freeman, K., Tallarida, R.: A quantitative study of dopamine control in the rat striatum. J. Pharmacol. Exp. Ther. **268**(2), 629–638 (1994)
633. Sun, Y., Jusko, W.: Role of baseline parameters in determining indirect pharmacodynamic responses. J. Pharm. Sci. **88**(10), 987–990 (1999)
634. Mackey, M., AnDerHeiden, U.: The dynamics of recurrent inhibition. J. Math. Biol. **19**(2), 211–225 (1984)
635. Klausen, T.: The feed-back regulation of erythropoietin production in healthy humans. Dan. Med. Bull. **45**(4), 345–353 (1998)
636. Boggs, D., Athens, J., Cartwright, G., Wintrobe, M.: Leukokinetic studies. IX. Experimental evaluation of a model of granulopoiesis. J. Clin. Invest. **44**, 643–656 (1965)
637. Rubinow, S., Lebowitz, J.: A mathematical model of neutrophil production and control in normal man. J. Math. Biol. **1**, 187–225 (1975)
638. Lange, R.: Cyclic hematopoiesis: human cyclic neutropenia. Exp. Hematol. **11**(6), 435–451 (1983)
639. Cohen, T., Cooney, D.: Cyclic thrombocytopenia. Case report and review of literature. Scand. J. Haematol. **12**(1), 9–17 (1974)
640. Mackey, M.: Unified hypothesis for the origin of aplastic anemia and periodic hematopoiesis. Blood **51**(5), 941–956 (1978)
641. Mackey, M.: Periodic auto-immune hemolytic anemia: an induced dynamical disease. Bull. Math. Biol. **41**(6), 829–834 (1979)
642. Dunn, C.: Cyclic hematopoiesis: the biomathematics. Exp. Hematol. **11**(9), 779–791 (1983)
643. Lieschke, G., Grail, D., Hodgson, G., Metcalf, D., Stanley, E., Cheers, C., Fowler, K., Basu, S., Zhan, Y., Dunn, A.: Mice lacking granulocyte colony-stimulating factor have chronic neutropenia, granulocyte and macrophage progenitor cell deficiency, and impaired neutrophil mobilization. Blood **84**(6), 1737–1746 (1994)
644. Price, T., Chatta, G., Dale, D.: Effect of recombinant granulocyte colony-stimulating factor on neutrophil kinetics in normal young and elderly humans. Blood **88**(1), 335–340 (1996)
645. Haurie, C., Dale, D., Mackey, M.: Cyclical neutropenia and other periodic hematological disorders: a review of mechanisms and mathematical models. Blood **92**(8), 2629–2640 (1998)
646. Degn, H., Holden, A., Olsen, L.: Chaos in Biological Systems. Plenum, New York (1987)
647. Bai-Lin, H.: Chaos II. World Scientific, Singapore (1990)
648. Quesenberry, P., Levitt, L.: Hematopoietic stem cells (second of three parts). N. Engl. J. Med. **301**(15), 819–823 (1979)
649. Terashi, K., Oka, M., Ohdo, S., Furukubo, T., Ikeda, C., Fukuda, M., Soda, H., Higuchi, S., Kohno, S.: Close association between clearance of recombinant human granulocyte colony-stimulating factor (G-CSF) and G-CSF receptor on neutrophils in cancer patients. Antimicrob. Agents Chemother. **43**(1), 21–24 (1999)
650. Wheldon, T.: Mathematical models of oscillatory blood cell production. Math. Biosci. **24**(3–4), 289–305 (1975)
651. Kazarinoff, N., VanDenDriessche, P.: Control of oscillations in hematopoiesis. Science **203**(4387), 1348–1349 (1979)
652. Hearn, T., Haurie, C., Mackey, M.: Cyclical neutropenia and the peripheral control of white blood cell production. J. Theor. Biol. **192**(2), 167–181 (1998)
653. Haurie, C., Dale, D., Rudnicki, R., Mackey, M.: Modeling complex neutrophil dynamics in the grey collie. J. Theor. Biol. **204**(4), 505–519 (2000)
654. Bernard, S., Belair, J., Mackey, M.: Oscillations in cyclical neutropenia: new evidence based on mathematical modeling. J. Theor. Biol. **223**(3), 283–298 (2003)
655. Gu, K., Kharitonov, V., Chen, J.: Stability of Time-Delay Systems. Control Engineering, Birkhauser, Boston (2003)

656. Karlsson, M., Port, R., Ratain, M., Sheiner, L.: A population model for the leukopenic effect of etoposide. Clin. Pharmacol. Ther. **57**, 325–334 (1995)
657. Minami, H., Sasaki, Y., Saijo, N., Ohtsu, T., Fujii, H., Igarashi, T., Itoh, K.: Indirect-response model for the time course of leukopenia with anticancer drugs. Clin. Pharmacol. Ther. **64**(5), 511–521 (1998)
658. Fokas, A., Keller, J., Clarkson, B.: Mathematical model of granulocytopoiesis and chronic myelogenous leukemia. Cancer Res. **51**(8), 2084–2091 (1991)
659. Iliadis, A., Barbolosi, D.: Optimizing drug regimens in cancer chemotherapy by an efficacy-toxicity mathematical model. Comput. Biomed. Res. **33**(3), 211–226 (2000)
660. Viens, P., Roche, H., Kerbrat, P., Fumoleau, P., Guastalla, J., Delozier, T.: Epirubicin-docetaxel combination in first-line chemotherapy for patients with metastatic breast cancer: final results of a dose-finding and efficacy study. Am. J. Clin. Oncol. **24**(4), 328–335 (2001)
661. Seymour, A., DeCampos, E., Thatcher, N., DeGreve, J., Cunningham, D., Howell, A., Tueni, E., Bron, D., Steward, W., Berdel, W., Knuth, A., Lorenz, J., Timothy, A., Yver, A., Richards, M.: A single-blind, randomised, vehicle-controlled dose-finding study of recombinant human granulocyte colony-stimulating factor (lenograstim) in patients undergoing chemotherapy for solid cancers and lymphoma. Eur. J. Cancer **31A**(13–14), 2157–2163 (1995)
662. Mackey, M., AnDerHeiden, U.: Dynamic diseases and bifurcations in physiological control systems. Funkt. Biol. Med. **1**, 156–164 (1982)
663. Gatti, R., Robinson, W., Deinard, A., Nesbit, M., McCullough, J., Ballow, M., Good, R.: Cyclic leukocytosis in chronic myelogenous leukemia: new perspectives on pathogenesis and therapy. Blood **41**(6), 771–782 (1973)
664. Chikkappa, G., Borner, G., Burlington, H., Chanana, A., Cronkite, E., Ohl, S., Pavelec, M., Robertson, J.: Periodic oscillation of blood leukocytes, platelets, and reticulocytes in a patient with chronic myelocytic leukemia. Blood **47**(6), 1023–1030 (1976)
665. Goldman, J., Melo, J.: Chronic myeloid leukemia. Advances in biology and new approaches to treatment. N. Engl. J. Med. **349**(15), 1451–1464 (2003)
666. Beresford, C.: Time: a biological dimension. J. R. Coll. Physicians Lond. **22**(2), 94–96 (1988)
667. Reimann, H.: Periodic diseases in the aged. Geriatrics **24**(5), 146–149 (1969)
668. Hellman, L., Nakada, F., Curti, J., Weitzman, E., Kream, J., Roffwarg, H., Ellman, S., Fukushima, D., Gallagher, T.: Cortisol is secreted episodically by normal man. J. Clin. Endocrinol. Metab. **30**(4), 411–422 (1970)
669. Ilias, I., Vgontzas, A., Provata, A., Mastorakos, G.: Complexity and non-linear description of diurnal cortisol and growth hormone secretory patterns before and after sleep deprivation. Endocr. Regul. **36**(2), 63–72 (2002)
670. West, B.: Fractal Physiology and Chaos in Medicine. World Scientific, Singapore (1990)
671. Lenbury, Y., Pacheenburawana, P.: Modelling fluctuation phenomena in the plasma cortisol secretion system in normal man. Biosystems **26**(2), 117–125 (1991)
672. Dokoumetzidis, A., Iliadis, A., Macheras, P.: Nonlinear dynamics in clinical pharmacology: the paradigm of cortisol secretion and suppression. Br. J. Clin. Pharmacol. **54**(1), 21–29 (2002)
673. Kraan, P., Dullaart, R., Pratt, J., Wolthers, B., Drayer, N., DeBruin, R.: The daily cortisol production reinvestigated in healthy men. The serum and urinary cortisol production rates are not significantly different. J. Clin. Endocrinol. Metab. **83**(4), 1247–1252 (1998)
674. Murray, J.: Mathematical Biology. Biomathematics Texts, vol. 19, 2nd edn. Springer, Berlin (1993)
675. Goodwin, B.: Oscillatory behavior in enzymatic control processes. Adv. Enzym. Regul. **3**, 425–438 (1965)
676. Leloup, J., Goldbeter, A.: A model for circadian rhythms in drosophila incorporating the formation of a complex between the PER and TIM proteins. J. Biol. Rhythm. **13**(1), 70–87 (1998)
677. Rohatagi, S., Bye, A., Mackie, A., Derendorf, H.: Mathematical modeling of cortisol circadian rhythm and cortisol suppression. Eur. J. Pharm. Biopharm. **4**(6), 341–350 (1996)

678. Hairer, E., Norsett, S., Wanner, G.: Solving Ordinary Differential Equations: Nonstiff Problems, 2nd edn. Springer, Berlin (1993)
679. Farmer, J.: Chaotic attractors of an infinite-dimensional dynamical system. Phys. D **4**(3), 366–393 (1982)
680. Gupta, S., Ritchie, J., Ellinwood, E., Wiedemann, K., Holsboer, F.: Modeling the pharmacokinetics and pharmacodynamics of dexamethasone in depressed patients. Eur. J. Clin. Pharmacol. **43**(1), 51–55 (1992)
681. Meibohm, B., Hochhaus, G., Mollmann, H., Barth, J., Wagner, M., Krieg, M., Stockmann, R., Derendorf, H.: A pharmacokinetic/pharmacodynamic approach to predict the cumulative cortisol suppression of inhaled corticosteroids. J. Pharmacokinet. Pharmacodyn. **27**(2), 127–147 (1999)
682. Smith, W.: Hypothalamic regulation of pituitary secretion of luteinizing hormone. II. Feedback control of gonadotropin secretion. Bull. Math. Biol. **42**(1), 57–78 (1980)
683. Cartwright, M., Husain, M.: A model for the control of testosterone secretion. J. Theor. Biol. **123**(2), 239–250 (1986)
684. Liu, B., Deng, G.: An improved mathematical model of hormone secretion in the hypothalamo-pituitary-gonadal axis in man. J. Theor. Biol. **150**(1), 51–58 (1991)
685. Das, P., Roy, A., Das, A.: Stability and oscillations of a negative feedback delay model for the control of testosterone secretion. Biosystems **32**(1), 61–69 (1994)
686. Topp, B., Promislow, K., Devries, G., Miura, R., Finegood, D.: A model of [beta] -cell mass, insulin, and glucose kinetics: pathways to diabetes. J. Theor. Biol. **206**(4), 605–619 (2000)
687. Londergan, C., Peacock-Lopez, E.: Dynamic model of hormonal systems coupled by negative feedback. Biophys. Chem. **73**(1–2), 85–107 (1998)
688. Fattinger, K., Verotta, D., Porchet, H., Munafo, A., LeCotonnec, J., Sheiner, L.: Modeling a bivariate control system: LH and testosterone response to the GnRH antagonist antide. Am. J. Physiol. **271**(4/1), E775–E787 (1996)
689. Francheteau, P., Steimer, J., Dubray, C., Lavene, D.: Mathematical model for in vivo pharmacodynamics integrating fluctuation of the response: application to the prolactin suppressant effect of the dupaminomimetic drug DCN 230-922. J. Pharmacokinet. Biopharm. **19**(3), 287–309 (1991)
690. Lalonde, R., Gaudreault, J., Karhu, D., Marriott, T.: Mixed-effects modeling of the pharmacodynamic response to the calcimimetic agent R-568. Clin. Pharmacol. Ther. **64**(1), 45–49 (1999)
691. Gobburu, J., Agerso, H., Jusko, W., Ynddal, L.: Pharmacokinetic-pharmacodynamic modeling of ipamorelin, a growth hormone releasing peptide, in human volunteers. Pharm. Res. **16**(9), 1412–1416 (1999)
692. Tolic, I., Mosekilde, E., Sturis, J.: Modeling the insulin-glucose feedback system: the significance of pulsatile insulin secretion. J. Theor. Biol. **207**(3), 361–375 (2000)
693. Prank, K., Harms, H., Dammig, M., Brabant, G., Mitschke, F., Hesch, R.: Is there low-dimensional chaos in pulsatile secretion of parathyroid hormone in normal human subjects? Am. J. Physiol. Endocrinol. Matabolism **266**(4), E653–E658 (1994)
694. Papavasiliou, S., Brue, T., Jaquet, P., Castanas, E.: Pattern of prolactin diurnal secretion in normal humans: evidence for nonlinear dynamics. Neuroendocrinology **62**(5), 444–453 (1995)
695. Pincus, S.: Approximate entropy as a measure of system complexity. Proc. Natl. Acad. Sci. U.S.A. **88**(6), 2297–2301 (1991)
696. Veldhuis, J., Pincus, S.: Orderliness of hormone release patterns: a complementary measure to conventional pulsatile and circadian analyses. Eur. J. Endocrinol. **138**(2), 164–169 (1998)
697. Kroll, M.: Biological variation of glucose and insulin includes a deterministic chaotic component. Biosystems **50**(3), 189–201 (1999)
698. Zeisberger, E.: Central modulators of thermoregulation. J. Basic Clin. Physiol. Pharmacol. **1**(1–4), 277–289 (1990)
699. Zeisberger, E.: Biogenic amines and thermoregulatory changes. Prog. Brain Res. **115**, 159–176 (1998)

700. Schwartz, P., Rosenthal, N., Wehr, T.: Serotonin 1A receptors, melatonin, and the proportional control thermostat in patients with winter depression. Arch. Gen. Psychiatry **55**(10), 897–903 (1998)
701. Yu, H., Lewander, T.: Pharmacokinetic and pharmacodynamic studies of (R)-8-hydroxy-2-(di-n-propylamino)tetralin in the rat. Eur. Neuropsychopharmacol. **7**(3), 165–172 (1997)
702. Abdel-Fattah, A., Matsumoto, K., El-Hady, K., Watanabe, H.: 5-HT1A and 5-HT2 receptors mediate hypo- and hyperthermic effects of tryptophan in pargyline-pretreated rats. Pharmacol. Biochem. Behav. **52**(2), 379–384 (1995)
703. Salmi, P., Ahlenius, S.: Evidence for functional interactions between 5-HT1A and 5-HT2A receptors in rat thermoregulatory mechanisms. Pharmacol. Toxicol. **82**(3), 122–127 (1998)
704. Hjorth, S.: Hypothermia in the rat induced by the potent serotoninergic agent 8-OH-DPAT. J. Neural Transm. **61**(1–2), 131–135 (1985)
705. Goodwin, G., DeSouza, R., Green, A., Heal, D.: The pharmacology of the behavioural and hypothermic responses of rats to 8-hydroxy-2-(di-n-propylamino)tetralin (8-OH-DPAT). Psychopharmacol. (Berl) **91**(4), 506–511 (1987)
706. Wakelkamp, M., Alvan, G., Gabrielsson, J., Paintaud, G.: Pharmacodynamic modeling of furosemide tolerance after multiple intravenous administration. Clin. Pharmacol. Ther. **60**(1), 75–88 (1996)
707. Bauer, J., Balthasar, J., Fung, H.: Application of pharmacodynamic modeling for designing time-variant dosing regimens to overcome nitroglycerin tolerance in experimental heart failure. Pharm. Res. **14**(9), 1140–1145 (1997)
708. Werner, J.: Modeling homeostatic responses to heat and cold. In: Blatteis, C., Fregly, M. (eds.) Handbook of Physiology, vol. sect. 4: Adaptation to the Environment, pp. 613–626. Oxford University Press, Oxford (1996)
709. Zuideveld, K., Maas, H., Treijtel, N., Hulshof, J., VanDerGraaf, P., Peletier, L., Danhof, M.: A set-point model with oscillatory behavior predicts the time course of 8-OH-DPAT-induced hypothermia. Am. J. Physiol. Regul. Integr. Comp. Physiol. **281**(6), R2059–R2071 (2001)
710. Gabrielsson, J., Jusko, W., Alari, L.: Modeling of dose-response-time data: four examples of estimating the turnover parameters and generating kinetic functions from response profiles. Biopharm. Drug Dispos. **21**(2), 41–52 (2000)
711. Kitney, R.: A nonlinear model for studying oscillations in the blood pressure control system. J. Biomed. Eng. **1**(2), 89–99 (1979)
712. Gordon, C., Heath, J.: Reassessment of the neural control of body temperature: importance of oscillating neural and motor components. Comp. Biochem. Physiol. A **74**(3), 479–489 (1983)
713. Zuideveld, K., Treijtel, N., Maas, H., Gubbens-Stibbe, J., Peletier, L., VanDerGraaf, P., Danhof, M.: A competitive interaction model predicts the effect of WAY-100,635 on the time course of R-(+)-8-hydroxy-2-(di-n-propylamino)tetralin-induced hypothermia. J. Pharmacol. Exp. Ther. **300**(1), 330–338 (2002)
714. Zuideveld, K., Rusic-Pavletic, J., Maas, H., Peletier, L., VanDerGraaf, P., Danhof, M.: Pharmacokinetic-pharmacodynamic modeling of buspirone and its metabolite 1-(2-pyrimidinyl)-piperazine in rats. J. Pharmacol. Exp. Ther. **303**(3), 1130–1137 (2002)
715. Zuideveld, K., VanDerGraaf, P., Newgreen, D., Thurlow, R., Petty, N., Jordan, P., Peletier, L., Danhof, M.: Mechanism-based pharmacokinetic-pharmacodynamic modeling of 5-HT1A receptor agonists: estimation of in vivo affinity and intrinsic efficacy on body temperature in rats. J. Pharmacol. Exp. Ther. **308**(3), 1012–1020 (2004)
716. Accardo, A., Affinito, M., Carrozzi, M., Bouquet, F.: Use of the fractal dimension for the analysis of electroencephalographic time series. Biol. Cybern. **77**(5), 339–350 (1997)
717. Silva, C., Pimentel, I., Andrade, A., Foreid, J., Ducla-Soares, E.: Correlation dimension maps of EEG from epileptic absences. Brain Topogr. **11**(3), 201–209 (1999)
718. Stam, K., Tavy, D., Jelles, B., Achtereekte, H., Slaets, J., Keunen, R.: Non-linear dynamical analysis of multichannel EEG: clinical applications in dementia and Parkinson's disease. Brain Topogr. **7**(2), 141–150 (1994)

719. Jeong, J., Kim, D., Chae, J., Kim, S., Ko, H., Paik, I.: Nonlinear analysis of the EEG of schizophrenics with optimal embedding dimension. Med. Eng. Phys. **20**(9), 669–676 (1998)
720. Ehlers, C., Havstad, J., Prichard, D., Theiler, J.: Low doses of ethanol reduce evidence for nonlinear structure in brain activity. J. Neurosci. **18**(18), 7474–7486 (1998)
721. Mandema, J., Danhof, M.: Electroencephalogram effect measures and relationships between pharmacokinetics and pharmacodynamics of centrally acting drugs. Clin. Pharmacokinet. **23**(3), 191–215 (1992)
722. Lehnertz, K.: Non-linear time series analysis of intracranial EEG recordings in patients with epilepsy - An overview. Int. J. Psychophysiol. **34**(1), 45–52 (1999)
723. Elger, C., Widman, G., Andrzejak, R., Arnhold, J., David, P., Lehnertz, K.: Nonlinear EEG analysis and its potential role in epileptology. Epilepsia **41**(S3), S34–38 (2000)
724. Roschke, J., Aldenhoff, J.: A nonlinear approach to brain function: deterministic chaos and sleep EEG. Sleep **15**(2), 95–101 (1992)
725. Isaacson, R., Varner, J., Baars, J., DeWied, D.: The effects of pregnenolone sulfate and ethylestrenol on retention of a passive avoidance task. Brain Res. **689**(1), 79–84 (1995)
726. DiMascio, M., DiGiovanni, G., DiMatteo, V., Esposito, E.: Decreased chaos of midbrain dopaminergic neurons after serotonin denervation. Neuroscience **92**(1), 237–243 (1999)
727. Ishizuka, S., Hayashi, H.: Spontaneous epileptiform bursts and long-term potentiation in rat CA3 hippocampal slices induced by chaotic stimulation of mossy fibers. Brain Res. **790**(1–2), 108–114 (1998)
728. Karlsson, M., Schoemaker, R., Kemp, B., Cohen, A., VanGerven, J., Tuk, B., Peck, C., Danhof, M.: A pharmacodynamic Markov mixed-effects model for the effect of temazepam on sleep. Clin. Pharmacol. Ther. **68**(2), 175–188 (2000)
729. Larter, R., Speelman, B., Worth, R.: A coupled ordinary differential equation lattice model for the simulation of epileptic seizures. Chaos **9**(3), 795–804 (1999)
730. Cleton, A., Mazee, D., Voskuyl, R., Danhof, M.: Rate of change of blood concentrations is a major determinant of the pharmacodynamics of midazolam in rats. Br. J. Pharmacol. **127**(1), 227–235 (1999)
731. Widman, G., Lehnertz, K., Jansen, P., Meyer, W., Burr, W., Elger, C.: A fast general purpose algorithm for the computation of auto- and cross-correlation integrals from single channel data. Phys. D **121**(1–2), 65–74 (1998)
732. Grassberger, P., Procaccia, I.: Measuring the strangeness of strange attractors. Phys. D **9**(1–2), 189–208 (1983)
733. Widman, G., Schreiber, T., Rehberg, B., Hoeft, A., Elger, C.: Quantification of depth of anesthesia by nonlinear time series analysis of brain electrical activity. Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics **62**(4 Pt A), 4898–4903 (2000)
734. Rampil, I.: A primer for EEG signal processing in anesthesia. Anesthesiology **89**(4), 980–1002 (1998)
735. Bruhn, J., Lehmann, L., Ropcke, H., Bouillon, T., Hoeft, A.: Shannon entropy applied to the measurement of the electroencephalographic effects of desflurane. Anesthesiology **95**(1), 30–35 (2001)
736. Wagner, C., Persson, P.: Chaos in the cardiovascular system: an update. Cardiovasc. Res. **40**(2), 257–264 (1998)
737. Sugihara, G., Allan, W., Sobel, D., Allan, K.: Nonlinear control of heart rate variability in human infants. Proc. Natl. Acad. Sci. U.S.A. **93**(6), 2608–2613 (1996)
738. Karagueuzian, H., Kogan, B., Khan, S., Denton, T., Karplus, W., Mandel, W., Diamond, G.: Induction of cellular chaos during quinidine toxicity. Predictive power of nonlinear dynamic analysis for drug-induced proarrhythmia- A hypothesis. J. Electrocardiol. **24** Suppl, 91–96 (1992)
739. Garfinkel, A., Spano, M., Ditto, W., and Weiss, J.: Controlling cardiac chaos. Science **257**, 1230–1235 (1992)
740. Scheinin, H., Helminen, A., Huhtala, S., Gronroos , P., Bosch, J., Kuusela, T., Kanto, J., Kaila, T.: Spectral analysis of heart rate variability as a quantitative measure of parasympatholytic effect-integrated pharmacokinetics and pharmacodynamics of three anticholinergic drugs. Ther. Drug Monit. **21**(2), 141–151 (1999)

741. DeBrouwer, S., Edwards, D., Griffith, T.: Simplification of the quasiperiodic route to chaos in agonist-induced vasomotion by iterative circle maps. Am. J. Physiol. **274**(4(2)), H1315–1326 (1998)
742. Weiss, J., Garfinkel, A., Karagueuzian, H., Qu, Z., Chen, P.: Chaos and the transition to ventricular fibrillation: a new approach to antiarrhythmic drug evaluation. Circulation **99**(21), 2819–2826 (1999)
743. Exner, D., Reiffel, J., Epstein, A., Ledingham, R., Reiter, M., Yao, Q., Duff, H., Follmann, D., Schron, E., Greene, H., Carlson, M., Brodsky, M., Akiyama, T., Baessler, C., Anderson, J.: Beta-blocker use and survival in patients with ventricular fibrillation or symptomatic ventricular tachycardia: the Antiarrhythmics Versus Implantable Defibrillators (AVID) trial. J. Am. Coll. Cardiol. **34**(2), 325–333 (1999)
744. Echt, D., Liebson, P., Mitchell, L., Peters, R., Obias-Manno, D., Barker, A., Arensberg, D., Baker, A., Friedman, L., Greene, H., Juther, M., Richardson, D.: Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. N. Engl. J. Med. **324**(12), 781–788 (1991)
745. Starmer, C., Romashko, D., Reddy, R., Zilberter, Y., Starobin, J., Grant, A., Krinsky, V.: Proarrhythmic response to potassium channel blockade. Numerical studies of polymorphic tachyarrhythmias. Circulation **92**(3), 595–605 (1995)
746. Qu, Z., Weiss, J., Garfinkel, A.: Cardiac electrical restitution properties and stability of reentrant spiral waves: a simulation study. Am. J. Physiol. **276**, No. 1(2), H269–283 (1999)
747. Garfinkel, A., Kim, Y., Voroshilovsky, O., Qu, Z., Kil, J., Lee, M., Karagueuzian, H., Weiss, J., Chen, P.: Preventing ventricular fibrillation by flattening cardiac restitution. Proc. Natl. Acad. Sci. U.S.A. **97**(11), 6061–6066 (2000)
748. Kowey, P.: Pharmacological effects of antiarrhythmic drugs. Review and update. Arch. Intern. Med. **158**(4), 325–332 (1998)
749. Reiffel, J., Kowey, P.: Generic antiarrhythmics are not therapeutically equivalent for the treatment of tachyarrhythmias. Am. J. Cardiol. **85**(9), 1151–1153, A10 (2000)
750. Steimer, J., Vozeh, S., Racine-Poon, A., Holford, N., O'Neil, R.: The population approach: rationale, methods, applications in clinical pharmacology drug development. In: Welling, P., Balant, L. (eds.) Pharmacokinetics of Drugs, Handbook of Experimental Pharmacology, pp. 405–451. Springer, Berlin (1994)
751. Sheiner, L., Rosenberg, B., Melmon, K.: Modelling of individual pharmacokinetics for computer-aided drug dosage. Comput. Biomed. Res. **4**, 441–459 (1972)
752. Sheiner, L., Rosenberg, B., Marathe, V.: Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. J. Pharmacokinet. Biopharm. **5**(5), 445–479 (1977)
753. Aarons, L., Balant, L., Danhof, M., Gex-Fabry, M., Gundert-Remy, U., Karlsson, M., Mentre, F., Morselli, P., Rombout, F., Rowland, M., Steimer, J., Vozeh, S.: The Population Approach: Measuring and Managing Variability in Response, Concentration and Dose, COST B1 Medicine. Commission of European Communities, Luxembourg (1997)
754. Beal, S., Sheiner, L.: The NONMEM system. Am. Stat. **34**(2), 118–119 (1980)
755. Lindstrom, M., Bates, D.: Nonlinear mixed effects models for repeated measures data. Biometrics **46**(3), 673–687 (1990)
756. Kuhn, E., Lavielle, M.: Maximum likelihood estimation in mixed effects models. Comput. Stat. Data Anal. **49**(4), 1020–1038 (2005)
757. Lunn, D., Best, N., Thomas, A., Wakefield, J., Spiegelhalter, D.: Bayesian analysis of population PK/PD models: general concepts and software. J. Pharmacokinet. Pharmacodyn. **29**(3), 271–307 (2002)
758. Davidian, M., Gallant, R.: The nonlinear mixed effect model with a smooth random effects density. Biometrika **80**, 475–488 (1993)
759. Mallet, A.: A maximum likelihood estimation method for random coefficient regression models. Biometrika **73**(3), 645–656 (1986)
760. Schumitzky, A.: Nonparametric EM algorithms for estimating prior distributions. Appl. Math. Comput. **45**(2), 143–157 (1991)

761. Aarons, L.: Software for population pharmacokinetics and pharmacodynamics. Clin. Pharmacokinet. **36**(4), 255–264 (1999)
762. Gurrin, L., Moss, T., Sloboda, D., Hazelton, M., Challis, J., Newnham, J.: Using WinBUGS to fit nonlinear mixed models with an application to pharmacokinetic modelling of insulin response to glucose challenge in sheep exposed antenatally to glucocorticoids. J. BioPharm. Stat. **13**(1), 117–139 (2003)
763. Mu, S., Ludden, T.: Estimation of population pharmacokinetic parameters in the presence of non-compliance. J. Pharmacokinet. Pharmacodyn. **30**(1), 53–81 (2003)
764. Duffull, S., Kirkpatrick, C., Green, B., Holford, N.: Analysis of population pharmacokinetic data using NONMEM and WinBUGS. J. BioPharm. Stat. **15**(1), 53–73 (2005)
765. Candy, J.: Signal Processing. The Model-based Approach. McGraw-Hill Series in Electrical Engineering, 2nd edn. McGraw-Hill, Singapore (1987)
766. Dresher, M.: The Mathematics of Games of Strategy: Theory and Applications. Dover, New York (1981)

Index

P. Macheras, A. Iliadis, *Modeling in Biopharmaceutics, Pharmacokinetics and Pharmacodynamics*, Interdisciplinary Applied Mathematics 30,
DOI 10.1007/978-3-319-27598-7

N

O

S