ELSEVIER

European Journal of Pharmaceutics and Biopharmaceutics 55 (2003) 77–84

European Journal of Pharmaceutics and Biopharmaceutics

www.elsevier.com/locate/ejphabio

Research paper

Handling of computational in vitro/in vivo correlation problems by Microsoft Excel II. Distribution functions and moments

Frieder Langenbucher*

BioVista LLC, Riehen, Switzerland

Received 25 April 2002; accepted in revised form 16 August 2002

Abstract

MS Excel is a useful tool to handle in vitro/in vivo correlation (IVIVC) distribution functions, with emphasis on the Weibull and the biexponential distribution, which are most useful for the presentation of cumulative profiles, e.g. release in vitro or urinary excretion in vivo, and differential profiles such as the plasma response in vivo. The discussion includes moments (AUC and mean) as summarizing statistics, and data-fitting algorithms for parameter estimation.

Keywords: In vitro/in vivo correlation; Excel; Distribution functions; Probability density function; Cumulative distribution function; Rate; Extent; Moments; Polyexponential; Weibull; Curve fitting

1. Introduction

General applications of MS Excel were discussed in a previous communication [1]. The present paper discusses the treatment of in vitro/in vivo correlation (IVIVC) time profiles by means of mathematical functions and stresses three important features. (i) Time profiles are distribution functions, where time t is the 'distributed' quantity, and the functions are used as either probability density function (PDF) (e.g. plasma response in vivo) or cumulative distribution function (CDF) (e.g. in vitro drug release or urinary excretion in vivo). (ii) While the polyexponential is well established for in vivo profiles, as is the Weibull distribution for in vitro release profiles, other distributions mentioned in the literature, e.g. [2], are deemed as less suitable. (iii) Distribution functions provide a set of moments (area, mean, variance) as useful statistics to summarize the entire time profile.

Again, MS Excel is found adequate for the computation of relevant functions and moments, non-linear curve fitting, or database tabulation, without a need for specialized and expensive software. (A collection of Excel worksheets, relevant to the entire series of papers, has been prepared in a workbook IVIVC.XLS, together with a description in IVIVC.DOC; both files may be requested from the author.)

2. Theory

2.1. Distribution functions

Elementary statistical tabulations, e.g. Table 2.2 and Figs. 2.2 and 2.3 of Ref. [3], or those similar in Ref. [4], distinguish between absolute and relative frequencies, where relative frequencies are expressed as fractions or percentages. Both can be presented as either PDF or CDF, as illustrated below:

	Absolute	Relative	Absolute cumulative	Relative cumulative
	PDF	PDF	CDF	RCF
	$f°(t)$	$f(t)$	$F°(t)$	$F(t)$
Class 1	3	0.200	3	0.200
Class 2	7	0.467	10	0.667
Class 3	5	0.333	15	1.000
Sum	15	1.000		

An essential link between the four profiles is the parameter F_∞ (= 15 in the example), division which converts absolute to relative frequencies. For discrete data, it is the

* Bio Vista LLC, Ruechligweg 101, CH-4125, Riehen, Switzerland. Tel.: +41-61603-2735; fax: +41-61601-2738.
E-mail address: mysickaj@ivivc.com (F. Langenbucher).

0939-6411/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0939-6411(02)00115-7

sum of all$f^{\circ}(t)$, for continuous profiles it is the AUC of the absolute PDF. On the other hand, it represents the final plateau of the CDF, whether this is estimated as the last observed data point, extrapolated from this to infinity, or supplied by a 'theoretical' value.

Time profiles encountered in IVIVC are generally represented by distribution functions, where time t is 'distributed' between classes: 0, t_1, t_2, t_3, …, ∞. In vitro, a cumulative release profile $F_D(t)$ expresses the distribution of drug substance dissolved (released) at time t; the corresponding differential profile $f_D(t)$ characterizes the rate of release. In both cases, F_∞ is supplied from the dose D. In vivo, a plasma concentration profile $f_P(t)$ represents the distribution of the drug in the plasma, i.e. absorbed but not yet eliminated; its cumulative equivalent, $F_P(t)$, represents the drug already eliminated. In these cases, F_∞ must be estimated from the data themselves, either as the AUC of the plasma profile or the final value of the cumulative urinary excretion.

Any distribution function can be defined with a time lag t_0 as additional function parameter, if t is substituted by $(t - t_0)$. This shifts the original profile to the left, such that initial values $\leq t_0$ are assumed to be zero. It must be judged as accidental that the Weibull distribution is usually reported including t_0, while others are not.

2.1.1. Elementary distributions

Fig. 1 illustrates the general behavior of four elementary distribution functions, for a PDF (top) and a corresponding CDF (bottom).

(a) The *unit pulse* represents a borderline case where the entire input occurs at time T. Its PDF, Eq. (1a), is the Dirac delta function δ, i.e. a rectangle of length ε, height $1/\varepsilon$ and area 1, with the condition that ε tends to 0. The corresponding PDF is the *step* function of Eq. (1b).

$$f(t) = \left\{ \begin{matrix} 1 & (t = T) \\ 0 & (t \neq T) \end{matrix} \right\} \tag{1a}$$

$$F(t) = \left\{ \begin{matrix} 0 & (t \leq T) \\ 1 & (t > T) \end{matrix} \right\} \tag{1b}$$

(b) A process with constant rate over an interval $0 < t < T$, e.g. drug release from an idealized 'zero-order' device, is described by a *rectangular* PDF and a corresponding *ramp* CDF.

$$f(t) = \left\{ \begin{matrix} 1/T & (t \leq T) \\ 0 & (t > T) \end{matrix} \right\} \tag{2a}$$

$$F(t) = \left\{ \begin{matrix} t/T & (t \leq T) \\ 1 & (t > T) \end{matrix} \right\} \tag{2b}$$

Eq. (2) implies a linear profile in the time range $0 < t < T$, where T is the time of completed dissolution/release. If appropriate, this profile may be replaced by a *polynomial* of differing order. E.g. the square-root law

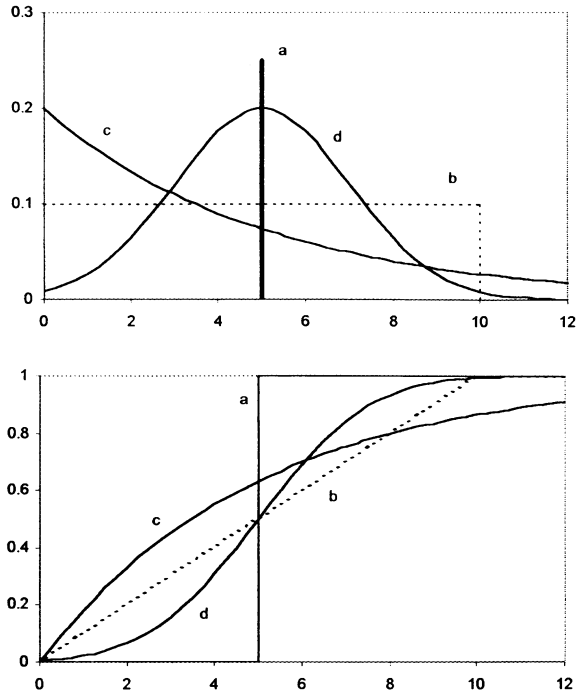

Fig. 1. Elementary relative distribution functions, with mean $\mu = 5$, displayed as PDF (top) or CDF (bottom). (a) Unit pulse ($\sigma = 0$); (b) Rectangular ($\sigma = 2.89$); (c) Exponential ($\sigma = 5$); (d) Normal ($\sigma = 2$).

$F(t) = k\sqrt{t}$ describes diffusion-controlled release, e.g. from a matrix product or an ointment; the cube-root or Hixson–Crowell law, $F(t) = 1 - (1 - t/T)^3$, describes dissolution of a bed of particles of regular shape.

(c) The *mono-exponential* function characterizes simple first-order processes by means of a single parameter, the rate constant k expressed in reciprocal time units (e.g. h^{-1}). Relative PDF and CDF are as follows[3].

$$f(t) = k\, e^{-kt} \tag{3a}$$

$$F(t) = 1 - e^{-kt} \tag{3b}$$

Note that $f(t)$ has the same unit as k, and its area AUC between 0 and ∞ equals 1. $F(t)$ is dimensionless and approaches a final value $F_\infty = 1$ for $t = \infty$.

(d) The *normal* distribution $N(\mu,\sigma)$ provides two parameters, mean μ and standard deviation σ.

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(t-\mu)^2/2\sigma^2} \tag{4}$$

The profile ranges from $-\infty$ to $+\infty$ and is symmetric about the mean, which is far from typical for data in vivo or in vitro. The same holds true for related functions: Probits shift normal deviates by 5σ, in order to avoid negative values. The logistic (Pearl–Reed) distribution, $F(t) = 1/(1 + \exp\{-(t - a))/b\}$, is similar to the normal one, but simpler to compute [3]. The log–normal distribution uses a logarithmic scale for the abscissa; this gives a skewed profile.

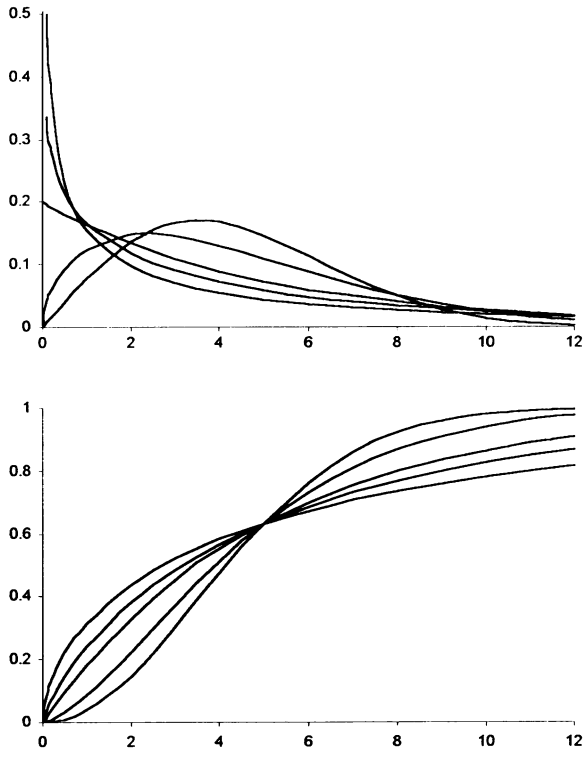

Fig. 2. Relative Weibull time profiles shown as PDF (top) and CDF (bottom). Parameters: $\beta = 5$ and $\alpha = 0.6, 0.8, 1, 1.5, 2$.

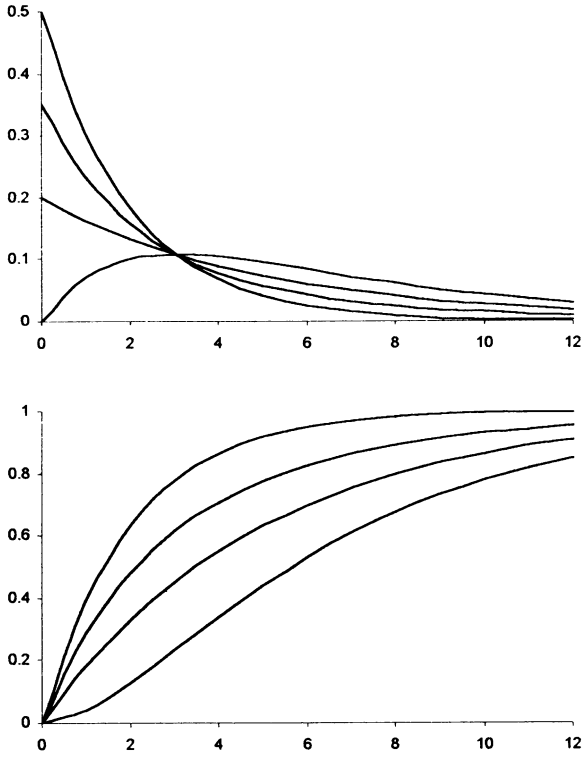

Fig. 3. Relative biexponential time profiles shown as PDF (top) and CDF (bottom). Parameters: $b_1 = 0.2$, $b_2 = 0.5$, $\alpha_1 = 0, 0.1, 0.2, 0.333$.

2.1.2. Weibull distribution

A useful extension of the exponential is the Weibull distribution which, when relative to $F_\infty = 1$, provides two parameters [5–7].

$$f(t) = (\alpha/\beta)(t/\beta)^{\alpha-1}\, e^{-(t/\beta)^\alpha} = (\alpha/\beta^\alpha)t^{\alpha-1}\, e^{-(t/\beta)^\alpha} \quad (5a)$$

$$F(t) = 1 - e^{-(t/\beta)^\alpha} \quad (5b)$$

The scale parameter β characterizes the overall rate corresponding with the time constant $1/b$ in the exponential. Flexibility is added by the dimensionless shape parameter α raising the time scale to a power: $\alpha = 1$ represents a monoexponential as special case, $\alpha > 1$ describes a 'sigmoid' profile retarded in the beginning, and $\alpha < 1$ represents a profile faster in the beginning but retarded in the tail. Fig. 2 illustrates this for $\beta = 5$ and five differing values of α. When substituting $t = \beta$ into Eq. (5a), the result is $F = 0.632$, independent of α. Accordingly, all cumulative profiles intersect at a point ($t = 5$, $F = 0.632$), which closely reflects the mean of the distribution.

The *Gompertz* distribution, similar to the Weibull, provides two parameters: an overall rate constant c, and a parameter k representing the initial slope of the CDF as well as the ordinate intercept of the PDF.

$$f(t) = k\, e^{ct} \exp\{-(k/c)(e^{ct} - 1)\} \quad (6a)$$

$$F(t) = 1 - \exp\{-(k/c)(e^{ct} - 1)\} \quad (6b)$$

The function may have certain advantages, but its documentation in the literature is poor and the interpretation of the parameters is clumsy.

2.1.3. Biexponential distribution

A most versatile expansion of the monoexponential is the *polyexponential* distribution, where each term adds two parameters, thus increasing flexibility.

$$f^\circ(t) = a_1\, e^{-b_1 t} + a_2\, e^{-b_2 t} + \cdots \quad (7a)$$

$$F^\circ(t) = (a_1/b_1)(1 - e^{-b_1 t}) + (a_2/b_2)(1 - e^{-b_2 t}) + \cdots \quad (7b)$$

where the area under the PDF and the final value of the CDF is $\mathrm{AUC} = F_\infty = \sum a/b$. For convenience, terms are arranged such that $0 < b_1 < b_2$, i.e. rate constants have positive sign and are ordered from slowest to fastest. Since $f^\circ(t)$ is confined to positive values, the coefficient of the smallest rate constant must have a positive sign ($0 < a_1$).

With two terms and four parameters (a_1, b_1, a_1, b_2), the polyexponential of Eqs. (7) reduces to the biexponential distribution, the AUC of which is $F_\infty = a_1/b_1 + a_2/b_2$. When relative to $F_\infty \equiv 1$, one of its four parameters can be expressed in terms of the other three, e.g. $a_2 = b_2 - a_1(b_2/b_1)$. Substitution of a_2 into Eqs. (7) gives the *relative* biexponential with three parameters (a, b_1, b_2).

$$f(t) = a\,e^{-b_1 t} + b_2(1 - a/b_1)\,e^{-b_2 t} \quad (8a)$$

$$F(t) = (a/b_1)(1 - e^{-b_1 t}) + (1 - a/b_1)(1 - e^{-b_2 t}) \quad (8b)$$

The relative ordinate of Eqs. (8) is easily verified: If F_∞ is computed for Eq. (8a), the result is $F_\infty = a/b_1 + 1 - a/b_1 = 1$; the same value is found if Eq. (8b) is evaluated for $t = \infty$.

Fig. 3 illustrates this for $b_1 = 0.2$ and $b_2 = 0.5$, and typical cases with increasing a_1

	(a)	(b)	(c)	(d)	(e)	(f)
a_1	0	0.10	0.2	0.250	0.333	0.50
a_2	0.5	0.25	0.0	− 0.125	− 0.333	− 0.75
$f(0)$	0.5	0.35	0.2	0.125	0	− 0.25

This covers monoexponential functions as borderline cases as well as the Bateman function. However, it is concluded from own experience that the function cannot compete with the Weibull for simplicity and flexibility.

In Fig. 3, (a) and (c) are monoexponential borderline cases where one of the two terms is zero: (a) is defined by b_2, and (c) by b_1. In the range between, e.g. (b), the faster term has a positive coefficient $a_2 > 0$. Hence, initial rates are faster than the exponential tail, which is typical for in vivo data of a two-compartment model after intravenous (i.v.) input. Beyond (c) the faster term has a negative sign, which is typical for profiles where the initial rate is slower than the exponential tail. For in vivo data, this covers cases of a one-compartment model with first-order input and output. (e) represents the *Bateman* function with $a_2 = -a_1$

$$f(t) = a(e^{-b_1 t} - e^{-b_2 t}) \quad (9)$$

where $a = 1/(1/b_1 - 1/b_2) = b_1 b_2/(b_2 - b_1)$. Cases beyond (e) have no physical significance.

For given values of b_1 and b_2 (e.g. 0.2 and 0.5), all PDF profiles intersect at a pivot point ($t^* = 3.054, f^* = 0.1086$). The value of t^* is found from the fact that the independence of a requires that the two terms in Eq. (8a), which contains a as a factor, must compensate to zero. Further rearrangement on both sides gives the following.

$$a\,e^{-b_1 t} - a(b_2/b_1)e^{-b_2 t} = 0$$

$$e^{-b_1 t} - (b_2/b_1)e^{-b_2 t} = 0$$

$$b_1\,e^{-b_1 t} = b_2\,e^{-b_2 t} \quad (10)$$

$$\ln b_1 - b_1 t = \ln b_2 - b_2 t$$

$$t = \ln(b_1/b_2)/(b_1 - b_2)$$

The ordinate f^* is obtained from Eq. (8a), if the terms containing a are dropped, t is substituted from Eq. (10), and the exponential is simplified by two elementary formulas: $e^{ab} = (e^a)^b = (e^b)^a$ and $e^{\ln(x)} = x$.

$$f(t) = b_2\,e^{-b_2 t} = b_2\,e^{-b_2 \ln(b_1/b_2)/(b_1 - b_2)}$$

$$f = b_2(b_1/b_2)^{b_2/(b_2 - b_1)} \quad (11)$$

2.2. Moments (semi-invariants)

Distribution function can be summarized by a set of *moments* of order $r = 0,\ 1,\ 2\ldots$, each of which is a single number obtained by integration over the entire time range [3], thus representing the profile as a whole:

$$S_r = \int_0^\infty t^r f(t)\,dt \quad (12)$$

Based on Eq. (12), a set of semi-invariants (cumulants) is defined, which characterize the time profile in a logical sequence from 'essential' to 'less important'. The first three of these have been reviewed [8–12], and are summarized in Table 1. Semi-invariants of higher order are not yet used in the context of IVIVC: Skewness κ_3 characterizes the degree of asymmetry around the mean, and Kurtosis κ_4 the proportion of the wings in relation to the center.

2.2.1. Area

κ_0 is the area under the PDF, usually denoted as AUC; and with regard to the CDF, it represents the final value F_∞. It clearly defines the scaling of the vertical response axis (ordinate), i.e. the total *extent* of input including the applied dose as

Table 1
Semi-invariants (area, mean, variance) of some typical distribution functions

Distribution	Equation	Figure	Area $\kappa_0 = AUC = F_\infty = S_0$	Mean $\kappa_1 = MDT = \mu = S_1/S_0$	Variance $\kappa_2 = \mathrm{VDT} = \sigma^2 = S_2/S_0 - (S_1/S_0)^2$
Unit pulse	(1)	1a	1	T	0
Rectangular	(2)	1b	1	$T/2$	$T^2/12$
Exponential	(3)	1c	1	$\tau = k^{-1}$	k^{-2}
Normal	(4)	1d	1	μ	σ^2
Weibull	(5)	2a–e	1	$\beta\Gamma(1 + 1/\alpha)$	$[\beta^2\Gamma(1 + 2/\alpha)] - \kappa_1^2$
Polyexponential	(7)	–	$\sum(a/b)$	$\sum(a/b^2)/\sum(a/b)$	$[\sum(2a/b^3)/\sum(a/b)] - \kappa_1^2$
Biexponential	(8)	3a–d	1	$a/b_1^2 + (1 - a/b_1)/b_2$	$[2a/b_1^3 + 2b_2(1 - a/b_1)/b_2^3] - \kappa_1^2$

well as effects such as overdose, incomplete release, or first-pass effect. For relative functions, it equals '1' by definition.

Numerical computation of area (and mean) by means of Excel was demonstrated earlier [1]. For a general polyexponential, the area is $F_\infty = \sum(a_i/b_i)$. The Weibull distribution, when expressed according to Eqs. (5), is relative to $F_\infty = 1$; to adjust to another area, the equations are multiplied by F_∞ as an additional parameter.

2.2.2. Mean

The mean κ_1 may alternatively be denoted as μ as in statistical functions, $\bar{t}$ as mean time, or a time constant $\tau = 1/k$ as reciprocal of a rate constant k. According to the nature of the distribution, acronyms such as MDT, MAT, or MRT are used for the mean time of dissolution, absorption, or residence.

The mean summarizes the overall *rate* of the process as a characteristic time, e.g. expressed in [h], and corresponding with a horizontal scaling along the time axis. The meaning is illustrated by the four profiles in Fig. 1 which are all constructed to the same mean, $\mu = 5$, by formulas given in Table 1. For the unit pulse, the mean is the time T of the pulse; for the rectangular distribution, it is half of the zero-order duration T. For the exponential, it is the time constant τ defined as reciprocal of the rate constant k; this is exactly the time $t_{0.632}$, where the CDF equals 63.2% of the final plateau. For the normal distribution, μ is the usual mean.

The mean of the polyexponential and the Weibull are shown in Table 1. Note that, in both cases, $t_{0.632}$, which is easily read or interpolated from the CDF, comes close to the true mean if the profile does not deviate too much from a monoexponential. The mean of the *Weibull* distribution uses the gamma function as extension of the factorial $n!$ to non-integer arguments [3]; it is larger than β for $\alpha < 1$, and smaller for $\alpha > 1$

α	0.6	0.8	1.0	1.5	2.0
$\Gamma(1 + 1/\alpha)$	1.505	1.133	1	0.903	0.886

Peak coordinates t_{max} and C_{max} are frequently used for summarizing time profiles, but this is a poor alternative, because both are usually obtained from one single data point. From a more general point of view, t_{max} may be considered as equivalent to the mean; C_{max} is a characteristic with its own merits, with respect to drug safety; in a systematic context, it is dubious in that it reflects both, extent and rate.

2.2.3. Variance

The variance characterizes the variability (dispersion) about the mean; a small value indicates close centering of the values about the mean. Its square root, the standard deviation σ, expresses the variability directly in time units, and is easier to judge from graphical presentations. The meaning is illustrated in Fig. 1, where the variance increases in the sequence (a) < (d) < (b) < (c). The coefficient of variation, CV = σ/μ, provides a relative statistic, frequently expressed as percentage.

Table 2
Excel sheet showing felodipine raw data and fitted parameters

TIME	OBS	FIT	RES
0	—	-35.093	—
0.33	5.110	5.110	3E-05
0.67	8.645	8.646	-0.001
1	7.788	7.782	0.006
1.5	6.082	6.114	-0.032
2	4.803	4.835	-0.032
3	3.416	3.185	0.231
4	2.132	2.245	-0.113
5.5	1.279	1.473	-0.194
7	0.961	1.049	-0.088
8.5	0.745	0.781	-0.036
10	0.588	0.594	-0.006
12	0.426	0.418	0.008

SSQ	T0
0.115	0.224

	B	A	A'
#1	0.170	3.200	3.080
#2	0.705	10.483	8.949
#3	6.240	-48.776	-12.029
CO:		-35.093	0

2.3. Curve-fitting algorithms

In general, fitting a distribution function to experimental data requires non-linear regression techniques. A well-known example is the 'method of residuals', which fits a polyexponential PDF, according to Eq. (7a), based on the fact that the tail of the profile is exclusively determined by the slowest term (e.g. see [13]). Excel's Solver tool [1] provides a more convenient technique applicable to any PDF or CDF, regardless of the specific function.

The technique is illustrated in Table 2 and Fig. 4, for polyexponential fitting of experimental data of felodipine oral solution [14]. Columns *TIME* and *OBS* list the observations, *FIT* the corresponding fitted values and *RES* the residuals = *OBS* − *FIT*. Note that the observation at $t = 0$ is not included for fitting, since this is a parameter to be estimated by the fitting procedure rather than a true obser-

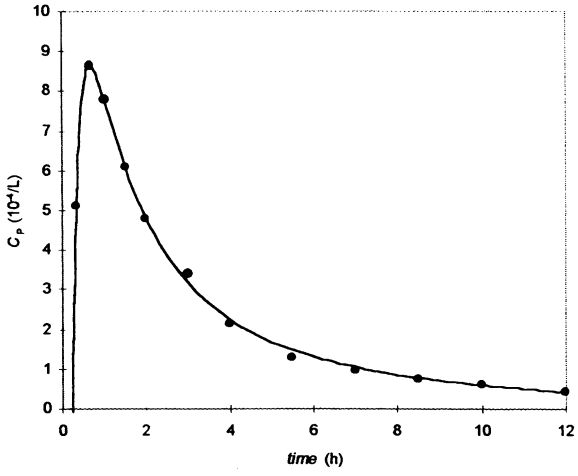

Fig. 4. Fitted triexponential function according to Table 2.

Table 3
Excel table storing polyexponential parameters of relevant weighting functions

DRUGNAME	FORM	*A*1 [1/L]	*B*1[1/h]	*A*2 [1/L]	*B*2 [1/h]	*A*3 [1/L]	*B*3 [1/h]	*T*0 [h]	REF
Baclofen	Syrup	0.0186	0.1900	− 0.0186	2.2000			0	P1104_B
Benazepril	Solution	0.0040	0.7020	0.0460	3.1950	− 0.0500	10.380	0	P1410_C
Cadralazine	Solution	0.0168	0.3040	0.0135	3.6500	− 0.0303	13.980	0	P1303_B
Carbamazepine	Suspension	0.0170	0.0182	− 0.0170	2.0000			0	P1474_C
Diclofenac	i.v. Bolus	0.0096	0.6600	0.0680	3.2000	0.5610	17.000	0	Willis et al. (1979)
Diclofenac	Solution	0.0058	0.3900	0.0580	3.2500	− 0.0638	9.200	0.04	Möller et al. (1984)
Diltiazem	Solution	0.2400	0.0710	1.8000	0.3240	− 2.0400	21.900	0.23	Ochs and Knüchel (1984)
Felodipine	Solution	3.0803	0.1700	8.9485	0.7054	− 12.0289	6.240	0.22	Edgar et al. (1987)
Formoterol	Syrup	1.1600	0.3470	4.5000	1.9800	− 5.6600	3.370	0	Sasaky and Kawai (1983)
Metoprolol	i.v. Bolus	0.0035	0.2050	0.0084	8.5400			0	P1001_B
Oxcarbazine	Solution	0.0284	0.0900	− 0.0284	0.5630			0	P1146_A
Oxprenolol	Solution	0.0077	0.3200	0.0053	0.9530	− 0.0130	7.720	0	P0997_A
Phenylbutazone	Solution	0.1050	0.0089	0.0633	0.2810	− 0.1680	1.040	0	P0757_A
Pirprofen	Solution	0.0884	0.1200	0.1110	0.6560	− 0.1990	8.340	0	P0697_B
Propyphenazone	Suspension	0.0180	0.5460	− 0.0180	6.4000			0	P0484_B

vation. The sum of squared residuals, *SSQ*, is minimized by the Solver to optimise the goodness of fit.

The data require at least two terms with parameters (a_1, b_1, a_2, b_2), but this fit is poor and considerably improved by including a third term (a_3, b_3). This best fit is shown by parameters in columns *A* and *B*, and *SSQ* = 0.115. The lag time is *T0* = 0, and the intercept has a large negative value *C0* = −35.093.

2.3.1. Time lag

The Solver is again invoked, in order to find the time t_0, where the function crosses the time axis. In the example, *T0* = 0.224 is found. With this value, each term of the polyexponential may be rewritten as follows.

$$a\,e^{-bt} = a\,e^{-bt_0}\,e^{-b(t-t_0)} = a'\,e^{-b(t-t_0)} \tag{13}$$

which makes use of the addition theorem of exponentials: $e^{p+q} = e^p\,e^q$. In this transformation, all rate constants b remain unchanged, and the coefficients a are replaced by

$$a' = a\,e^{-bt_0} \tag{14}$$

2.3.2. Parameter tabulation

For plasma-weighting functions (response to intravenous (i.v.) or post-operative (p.o.) bolus administration), final parameter estimates as obtained in Table 2, are stored in an Excel database table as reference for further (de)convolution analysis. Table 3 gives an example of such a compilation, where the field *DRUGNAME* identifies the drug substance, *FORM* the type of the bolus administration, and *REF* lists the data source as either literature reference or company-internal report.

Function parameters are stored for up to three exponential terms, in a sequence from slowest (*A1*, *B1*) to fastest (*A3*, *B3*). All time values are expressed in hours; hence, rate constants *B* are expressed in [1/h] and lag times *T0* in h. Drug amounts *DOSE* (not shown) are expressed in mg, and all amounts and concentrations are based on this unit. Volumes are expressed in [L], hence concentrations *m/V* are in [mg/L]. Coefficients *A* are primarily obtained in concentration units mg/l, but division by *DOSE* permits to tabulate them 'dose-independent' in the unit [1/L].

3. Relevant Excel features

3.1. Functions, charts, moments

3.1.1. Distribution functions

Distribution functions are specified in Excel by the abscissa value (e.g. time), followed by relevant function parameters. The last parameter specifies whether the function returns a value of the CDF (cum = true) or the PDF (cum = false). E.g. NORMDIST(t;mean;std;cum) evaluates the normal distribution for time t and mean and standard deviation. EXPONDIST(t;k;cum) evaluates the mono-exponential with rate constant k, but offers no advantage over the explicit formulas (1/k)*EXP(−k*t) or 1 − EXP(−k*t).

WEIBULL(t;*ALF*;*BET*;cum) computes the Weibull distribution, for scale parameter β and shape parameter α. For $t = 0$, the PDF according to Eq. (5a) has three distinct values $f(0)$: $1/\beta$ for $\alpha = 1$, ∞ for $\alpha < 1$, and 0 for $\alpha > 1$. The Excel function is erroneous in that it always returns a value of 0, which is true only for $\alpha > 1$. To overcome this problem, the value is best computed by this formula

$$(ALF < 1)^*1E + 99 + (ALF = 1)^*(1/BET) + (ALF > 1)^*0$$

where the logical expressions in parentheses evaluate to '1' if true, and to '0' if false.

3.1.2. Time plots

Time plots of computed functions are generated by first computing a series of equidistant time values in a column *TIME*. A typical example ranges from *START* = 0 to

FINTIM = 12, in equal steps of *DELT* = 0.2. This corresponds with *((FINTIM − START)/DELT)* + 1 = 61 rows; if these are entered in rows 5–65, they can be displayed on a single 'portrait' print page

		TIME		*FUN1*	*FUN2*
START→	R5	0	—		
DELT→	R6	0.2	=R[-1]C+*DELT*		
	R7	0.4	=R[-1]C+*DELT*		
		:			
FINTIM→	R65	12	=R[-1]C+*DELT*		

To create the series, enter *START* and *DELT* on top and use the Fill-handle to extend the selection up to *FINTIM*. A more flexible way is to enter *START* and *FINTIM* in the desired rows and enter = R[−1]C + DELT as recursive formula in all rows below *START*. Computed function values are entered in columns *FUN1*, *FUN2*….

3.1.3. Moments

Moments, in general, are computed by elementary Excel formulas according to Table 1. Mean and variance of the Weibull distribution are computed by using the function GAMMALN(x) which returns the natural logarithm of the gamma function Γ(x); e.g. the mean is computed as BET*EXP(GAMMALN(1 + 1/ALPH)).

3.2. Curve fitting by using the Solver

The following Excel Solver techniques directly apply to the polyexponential fitting illustrated in Table 2. Adaptation to other data types (e.g. CDF profiles) or distribution functions (e.g. Weibull) should be obvious.

3.2.1. Curve fitting

The columns *TIME* and *OBS* contain the observed data; two additional columns *FIT* and *RES* contain the computed values of the fitted values and the residuals (*OBS* − *FIT*). A block (*A1*;*B1*;*A2*;*B2*;*A3*;*B3*) contains the parameters of the specific function, with initial guesses such as *A1* = *B1* = 1 and zero otherwise. A cell *SSQ* computes the target cell which has to be minimized. The 'Solver Parameters' dialog box has these specifications:

SetTargetCell: SSQ	EqualTo: Min
ChangingCells:	A1;B1;A2;B2;A3;B3
Constraints:	B2 > B1;B3 > B2;A1 > 0

For convenience, the number of polyexponential terms is delimited to three, and ordered as $b_1 < b_2 < b_3$, i.e. from slowest to fast. Terms not used in the model have $a = 0$ and $b = 0$; since $e^0 = 1$, they contribute a value of '0'. Based on general pharmacokinetic principles, the following constraints must be observed: (i) The slowest term b_1 must always have a positive coefficient a_1. (ii) The sum of the coefficients determines the start of the function at $t = 0$. Delayed input requires $a_1 + a_2 + a_3 = 0$ to force the curve through the origin (0,0). *Bolus* input requires $a_1 + a_2 + a_3 > 0$ to ensure a positive starting point. (iii) For $n = 3$, the sign of a_2 depends on whether data prior to the final exponential tail are above ($a_2 > 0$) or below ($a_2 < 0$) the fitted curve.

3.2.2. Lag time

Lag time *T0* is estimated as the time value where the fitted function crosses the time axis, between meaningless negative values in the beginning and reasonable positive values thereafter. Initially, *T0* = 0 and *A′* contains the formula = *A*EXP(−*B***T0*). As a result, block *A′* displays the same values as the block *A*. Whenever another value is typed in *T0*, values in *A′* will change according to Eq. (11). The best estimate of *T0* is the value for which *C0′* = *A1′* + *A2′* + *A3′* equals 0. The Excel Solver finds this value, when invoked with the following specifications.

SetTargetCell: C0′ EqualTo: 0 ChangingCells: T0

Notation, Acronyms

AUC	Area Under the Curve
CDF	Cumulative Distribution Function
MAT	Mean Absorption Time
MDT	Mean Dissolution Time
MRT	Mean Residence Time
PDF	Probability Density Function
RCF	Relative Cumulative Function

Acknowledgements

The present manuscript was prepared with the assistance of J. Mysicka and Dr H. Rettig, whose contributions are gratefully acknowledged.

References

[1] F. Langenbucher, Handling of computational in vitro/in vivo correlation problems by Microsoft Excel, Eur. J. Pharm. Biopharm. 53 (2002) 1–7.
[2] P. Sathe, Y. Tsong, V.P. Shah, In vitro dissolution profile comparison and IVIVR, carbamazepine case, in: D. Young, J.G. Devane, J. Butler (Eds.), In vitro–in vivo correlations, Plenum Press, New York, NY, 1997, pp. 31–42.
[3] N.L. Johnson, F.C. Leone, Statistics and Experimental Design in Engineering and the Physical Sciences, Wiley, New York, NY, 1977.
[4] C.A. Bennett, N.L. Franklin, Statistical Analysis in Chemistry and the Chemical Industry, Wiley, New York, NY, 1963.
[5] F. Langenbucher, Linearization of dissolution rate curves by the Weibull distribution, J. Pharm. Pharmacol. 24 (1972) 979–981.
[6] F. Langenbucher, Parametric representation of dissolution rate curves by the RRSBW distribution, Pharm. Ind. 38 (1976) 472–477.

[7] S. Kotz, N.L. Johnson, Encyclopedia of Statistical Sciences, vol. 9, Wiley, New York, NY, 1988 pp. 549–556.
[8] K. Yamaoka, T. Nakagawa, T. Uno, Statistical moments in pharmacokinetics, J. Pharmacokin. Biopharm. 6 (1978) 547–558.
[9] D.J. Cutler, Theory of the mean absorption time, an adjunct to conventional bioavailability studies, J. Pharm. Pharmacol. 30 (1978) 476–478.
[10] D. Brockmeier, In vitro/in vivo correlation of dissolution using moments of disssolution and transit times, Acta Pharm. Technol. 32 (1986) 164–174.
[11] S.L. Beal, Some clarifications regarding moment of residence times with pharmacokinetic models, J. Pharmacokin. Biopharm. 15 (1987) 75–92.
[12] D. Brockmeier, Mean time concept and component analysis in pharmacokinetics, Int. J. Clin. Pharmacol. Ther. 37 (1999) 555–561.
[13] L. Shargel, A.B.C. Yu, Applied Biopharmaceutics and Pharmacokinetics, Appleton & Lange, Stamford, CT, 1999.
[14] B. Edgar, P. Lundborg, C.G. Regardh, Clinical Pharmacokinetics of Felodipine, Drugs 34 (Suppl. 3) (1987) 16–27.