ELSEVIER

International Journal of Pharmaceutics 240 (2002) 37–53

international journal of pharmaceutics

www.elsevier.com/locate/ijpharm

Non-linear mixed effects models for the evaluation of dissolution profiles

E. Adams [a,*], D. Coomans [b], J. Smeyers-Verbeke [a], D.L. Massart [a]

[a] *Pharmaceutical and Biomedical Analysis, Pharmaceutical Institute, Vrije Universiteit Brussel, Laarbeeklaan 103, B-1090 Brussels, Belgium*

[b] *School of Mathematical and Physical Sciences, James Cook University of North Queensland, Townsville Q4811, Australia*

Received 21 November 2001; accepted 12 March 2002

Abstract

The use of non-linear mixed effects models to describe dissolution data has been evaluated. A theoretical part is included to introduce this approach to scientists who are not familiar with this type of statistics. The standard settings of the statistical software package (S-plus) are used as much as possible. Several mathematical functions like the Weibull, logistic, first-order and Gompertz are employed as basis for the non-linear mixed effects models. Examples are given using dissolution data of immediate and extended release tablets. The results are compared with those obtained using linear mixed effects models.

Keywords: Dissolution; Non-linear mixed effects models; Pharmaceutical statistics; Longitudinal data

1. Introduction

The in vitro dissolution test is an important element in the development and quality control of solid oral dosage forms (tablets and capsules). In case of certain scale-up and post-approval changes (SUPAC), the Food and Drug Administration (FDA) even allows it to replace the costly in vivo experiments by in vitro dissolution tests (FDA Guidance for Industry, 1995, 1997). In general, a reference and a test batch have to be compared with each other. Dissolution data are obtained by measuring at certain time points the amount of the active substance released in the dissolution medium. To evaluate these dissolution data, several methods have been described which can be divided into ANOVA-based, model-independent and model-dependent methods (Polli et al., 1996).

The ANOVA-based procedures are sometimes also classified as model-independent (Polli et al., 1996). This can be somewhat confusing since they make use of an underlying model (Adams et al., 2001b). However, the results can not be used to fit a curve through the measured data points and so these methods are neither considered as model-dependent.

* Corresponding author. Tel.: +32-2-477-4723; fax: +32-2-477-4735.

E-mail address: eadams@fabi.vub.ac.be (E. Adams).

0378-5173/02/$ - see front matter
PII: S0378-5173(02)00127-8

The FDA proposes the calculation of the f_2 or similarity factor, originally introduced by Moore and Flanner (1996). This model-independent approach can be computed as:

$$f_2 = 50 \times \log\left\{\left[1 + \frac{1}{n}\sum_{t=1}^{n} w_t(R_t - T_t)^2\right]^{-0.5} \times 100\right\} \quad (1)$$

with R_t and T_t the average percentage dissolved at time t (for $t = 1, 2, \ldots, n$) for the reference and test set, respectively; w_t is an optional weight factor which is usually set equal to 1. When both sets are equal, $f_2 = 100$. The FDA allows an empirically determined 10% average difference at each sample time point, yielding a f_2 factor of 50. So, two sets are considered pharmaceutically equivalent when the f_2 factor lies between 50 and 100. Other model-independent techniques are based on the Mahalanobis distance (Tsong et al., 1996), the area under the curve (Anderson et al., 1998) or principal component analysis (Adams et al., 2001a).

The model-dependent methods all rely on a curve fitting procedure. Different mathematical functions have been used to model the observed data (Costa and Sousa Lobo, 2001). A distinction can be made between linear (zero-order, Higuchi, Hixson-Crowell, quadratic, polynomials) and non-linear models (first-order, Weibull, Korsmeyer-Peppas, logistic, Gompertz). Some of these models like Hixson-Crowell and Higuchi, are derived from the theoretical concepts of the dissolution process. Since the latter can be complicated, it is often difficult to describe it mathematically in a correct way. Hence, empirical equations like the Weibull have proven to be more adequate. The purpose of using mathematical models is that they facilitate the analysis and interpretation of the observed data because they describe the dissolution profiles as a function of only a few model parameters that can be statistically compared. A drawback is that these models are rather rigid and none of them is really suitable to fit all kinds of dissolution curves. In general, the Weibull is found to be the most successful (Sathe et al., 1996; Polli et al., 1997; Yuksel et al., 2000).

Another possibility to fit dissolution profiles is the use of mixed effects models. This approach is considered as superior to other modelling techniques since it takes into account the covariance structure of the data. Also here a distinction can be made between linear mixed effects (LME) and non-linear mixed effects (NLME) models. The evaluation of dissolution profiles by LME models, including an extensive theoretical part, has been described recently by Adams et al. (2001b). Using the standard settings of the statistical software (S-plus), convergence is not always reached and it is not easy to model information left in the residuals. On the other hand, LME models allow to analyse the dissolution data very accurately and are much more discriminative than the f_2 factor. An example of the application of NLME models to dissolution data can be found in Crowder (1996). Although the results are interesting, Crowder's procedure is difficult to implement in practice by people who are not familiar with NLME models since non-conventional statistical software is used.

In this paper the performance of non-linear mixed effects models is examined using the commercially available software package S-plus. The theoretical part, included to make the technique more accessible to non-statisticians, is followed by the analysis of real pharmaceutical data. Several mathematical functions were tried as starting point for the NLME models.

2. Theory

2.1. Linear versus non-linear models

Linear models are in essence often polynomial functions that are linear in their parameters. One has to keep in mind that for these models the conclusions are only valid within the observed data range (Pinheiro and Bates, 2000). Non-linear models on the other hand provide more reliable predictions for responses outside the observed range of the data. Another difference between both types is that non-linear models can be mechanistic, i.e. based on a (theoretical) model describing the underlying mechanism that produces the

data. As a consequence, the non-linear model parameters then have a more physical interpretation than the linear ones. A non-linear model usually contains also fewer parameters than a linear one. However, non-linear models are more computationally intensive and require starting estimates for the fixed effects coefficients. It is important to note that it is not always easy to choose reasonable values for these estimates and poor starting values may result in calculations that do not converge.

2.2. Mixed model analysis

Mixed effects models incorporate both fixed and random effects. Fixed effects are parameters associated with chosen, repeatable levels of experimental factors. Random effects are instead associated with experimental units drawn at random from a population. Applied to the dissolution data, a number of tablets are taken at random from each batch (= population). During the dissolution test, the percentage of drug dissolved is measured for each selected tablet at certain, previously determined, time points. So, 'batch' and 'time' are fixed factors while 'tablet' is a random one.

When two or more datasets have to be compared with each other, one can use one set as the reference, calculate confidence limits and check whether the test set meets these requirements. Another possibility is to compute the parameter differences between the data sets and examine if they differ significantly from zero or not. The latter approach is recommended when using mixed effects models.

2.3. NLME models

Although a lot of mathematical models have been described to fit dissolution curves, only the most 'promising' ones were selected to be translated to NLME models. As mentioned in the introduction, the dissolution process is difficult to describe pure theoretically and empirical functions give often better results.

2.3.1. Weibull

The Weibull function is often found to be the best choice to fit dissolution data:

$$X_{(t)} = X_{(\infty)} \times [1 - \exp(-\alpha t^{\beta})] \quad (2)$$

with $X_{(t)}$ the percentage drug dissolved at time t, $X_{(\infty)}$ the percentage drug dissolved at infinite time, α the scale factor and β the shape factor.

$X_{(\infty)}$ is usually set equal to 100%. The two non-linear model parameters to be determined and interpreted are α and β. The scale factor α corresponds to the time scale of the process. Parameter β describes the shape of the profile: exponential ($\beta = 1$), sigmoid ($\beta > 1$) or parabolic with a higher initial slope compared to the exponential ($\beta < 1$) (Langenbucher, 1972).

Model Eq. (2) is usually used to fit the mean profiles of each batch. Consequently, one has no idea about the behaviour of the separate tablets in a batch or in other words, the random effects are ignored. The application of a NLME model solves this problem. When model Eq. (2) is translated to a non-linear mixed effect model:

$$y_{ij} = X_{(\infty)} \times [1 - \exp(-(\alpha_0 + \alpha_1 B_i + a_i) t_{ij}^{(\beta_0 + \beta_1 B_i + b_i)})] + \varepsilon_{ij} \quad (3)$$

with the assumptions: a_i, $b_i \sim N(\boldsymbol{0}, \boldsymbol{D})$, $\varepsilon_{ij} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_i)$ and $a_1,\ldots, a_T$, $b_1,\ldots, b_T$, $\varepsilon_1,\ldots, \varepsilon_T$ independent (with T the number of tablets).

This equation must be interpreted as follows: y_{ij} the percentage dissolved at time point j for tablet i; α_0 and β_0 the fixed effects parameters for the reference set; α_1 and β_1 the differences between the fixed effects parameters for the reference and the test set; a_i and b_i the random effects associated with tablet i; B_i the indicator variables (0 if tablet i belongs to the reference and 1 if to the test set) and ε_{ij} the residual components or random errors.

So, in this mixed effects model, the fixed effects (α_0 and β_0 for the reference batch; $\alpha_0 + \alpha_1$ and $\beta_0 + \beta_1$ for the test batch) represent the mean values of the parameters for each batch. The deviations of the individual coefficients of tablet i from these mean values are represented as the random effects (a_i and b_i). The random effects and errors are assumed to be normally distributed with mean $\boldsymbol{0}$ and covariance matrix $\boldsymbol{D}$ and $\boldsymbol{\Sigma}_i$,

respectively. Random effects corresponding to different tablets are assumed to be independent of each other and of the random errors.

The problem with the Weibull function is the choice of the initial estimates for the fixed effects. For β, one can have a rough idea based on the shape of the profile: $\beta = 1$, $\beta > 1$ or $\beta < 1$ (see above). Estimators for α are more difficult to determine. In the literature, values for α vary between 0.03 and 2.1 and for β between 0.4 and 1.75 (Sathe et al., 1996; Sathe et al., 1997; Polli et al., 1997). Of course, these parameter values also depend on the time units (hours, minutes) used.

2.3.2. Logistic

The logistic function is also one of the better choices to fit dissolution curves:

$$X_{(t)} = \frac{\alpha}{1 + \exp[(\beta - t)/\gamma]} \tag{4}$$

with $X_{(t)}$ the percentage of drug dissolved at time t; α the value of the horizontal asymptote as t approaches infinity ($t \to \infty$); β the time at which $X_{(t)} = \alpha/2$; γ the scale parameter on the time axis which represents the distance on the time axis between β and the point where the response is $\alpha/(1 + e^{-1}) \approx 0.73\alpha$. The meaningful graphical interpretation of the model parameters facilitates the choice of the starting estimates: for α a value of 100 (%) is reasonable, for β the (average) time at which 50% of the drug is dissolved ($\sim$ dissolution rate) and for γ the difference between the (average) times at which 50 and 73% of the drug are dissolved, respectively.

The non-linear mixed effect version of model Eq. (4) is:

$$y_{ij} = \frac{\alpha_0 + \alpha_1 B_i + a_i}{1 + \exp[((\beta_0 + \beta_1 B_i + b_i) - t)/(\gamma_0 + \gamma_1 B_i + c_i)]} + \varepsilon_{ij} \tag{5}$$

The interpretation is similar as for model Eq. (3): y_{ij} the percentage dissolved at time j for tablet i; α_0, β_0 and γ_0 represent the values for the fixed effects parameters of the reference set; α_1, β_1 and γ_1 the differences between the fixed effects parameters of the reference and test set; B_i indicates whether tablet i belongs to the reference ($B_i = 0$) or the test batch ($B_i = 1$). Deviations of individual tablets are represented by the random effects a_i, b_i and c_i, which are assumed to be independent and normally distributed with mean $\boldsymbol{0}$ and covariance matrix $\boldsymbol{D}$. The residuals ε_{ij} are assumed to be normally distributed with mean 0 and covariance matrix $\boldsymbol{\Sigma}_i$. They are also supposed to be independent from each other and from the random effects.

2.3.3. First-order

Analogous to the previous models, the mixed effects version of the first-order model to fit dissolution curves can be represented as:

$$y_{ij} = (\alpha_0 + \alpha_1 B_i + a_i) \times (1 - e^{-(\beta_0 + \beta_1 B_i + b_i)t}) + \varepsilon_{ij} \tag{6}$$

The interpretation is similar as above. The assumptions concerning the random effects and residuals are also made here.

Since $(\alpha_0 + \alpha_1 B_i + a_i)$ corresponds to the value of the horizontal asymptote as in the logistic model, a reasonable starting value for α_0 is 100 (α_1 can be given 0 as starting value, supposing there is no difference between the two sets). Starting estimates for β_0, which reflects the release rate of the reference set, are more difficult to derive. Values between 0.03 and 0.15 are mentioned (Polli et al., 1997; Yuksel et al., 2000). As for the Weibull function, the β-values are also dependent on the time units used.

2.3.4. Gompertz

The Gompertz function can be written in its mixed effects form as:

$$y_{ij} = (\alpha_0 + \alpha_1 B_i + a_i) \times \exp[-\exp(-(\beta_0 + \beta_1 B_i + b_i))(t - (\gamma_0 + \gamma_1 B_i + c_i)))] + \varepsilon_{ij} \tag{7}$$

The interpretation and assumptions are the same as in the previous sections. As in the logistic and first-order function, $(\alpha_0 + \alpha_1 B_i + a_i)$ corresponds

to the horizontal asymptote when $t \to \infty$, $(\beta_0 + \beta_1 B_i + b_i)$ depends on the curvature of the profile and $(\gamma_0 + \gamma_1 B_i + c_i)$ describes the rising part ($\sim$ dissolution rate) of the curve. Also here raises the problem of the initial starting estimates. As before, for α_0 '100' is reasonable. Possible starting values for β_0 and γ_0 were determined experimentally.

2.4. Estimation of the model parameters and comparison of models

For the estimation of the model parameters an alternating algorithm is used (Pinheiro and Bates, 2000). It alternates between a penalised non-linear least squares step (PNLS) and a linear mixed effects (LME) step until a convergence criterion is met. In the PNLS step, the Gauss–Newton algorithm is used to estimate the fixed and random effects. The LME step returns the maximum likelihood (ML) or restricted maximum likelihood estimators (REML) of the model parameters using a hybrid approach of Expectation–Maximisation and Newton–Raphson iterations. Several structures can be attributed to the within group covariance matrix. By default, the unstructured type is defined.

The goodness of fit of a model can be evaluated using the Aikaike information criterion (AIC). This criterion takes into account the number of model parameters. In S-plus, the lower the AIC value, the better the model fits. The likelihood ratio test (LRT) is a statistical test to decide if the difference between two nested models is significant or not (a model is called 'nested' with another if it is a special case of the other model, e.g. the Weibull model without random effects is nested in the Weibull model with random effects).

2.5. Working procedure to fit NLME models

To avoid convergence problems due to overparametrisation, first a model with only fixed effects parameters is fitted. If no convergence is obtained, one can try other starting values. So, essentially one starts to fit a 'classical' non-linear model with no random effects parameters. Together with the estimated values for the model parameters, their significance is given in the output. Next, the random effects parameters are added individually, as well as all possible combinations. For example, the Weibull can be fitted using three different combinations of random effects: a model with only a_i, one with only b_i and one with a_i and b_i. For some combinations however, it is possible that no convergence is reached using the standard settings of the program. Adapting these settings is possible, but it can be time consuming and it is not evident for occasional users of this software. Other starting values can sometimes help too.

For each model that can be fitted, the AIC is calculated and different nested models can be statistically compared using the LRT. During refinement of the model, i.e. the determination of random effects that improve the model, and for model comparison (Weibull, logistic,…), the ML estimates must be used. Although the differences between the ML and REML estimates are small, REML is preferred for the final parameter estimates since it gives more accurate results.

The deviations of the individual tablets from the fixed effects parameters can be represented in a random effects plot. The model fit can be further evaluated using diagnostic plots: the residuals versus fitted values and the autocorrelation of the residuals versus lag. Following the assumption $\varepsilon_{ij} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_i)$, the residuals should be homogeneously spread around the zero line. The autocorrelation or serial correlation of the residuals varies between 1 and -1: for the same time point ($\text{lag} = 0$) the correlation is always 1; for time points 'further away' ($\text{lag} > 0$) the values should be relatively small. It is important to notice that some 'artificial' correlation for $lags > 0$ can be observed when the model is somewhat deficient and not all information of the original data is modelled. This is normally also reflected in the residuals versus fit plot.

2.6. Software

The NLME models were studied using S-plus 2000 (Mathsoft, Seattle, WA) in the mode Statistics □ Mixed Effects □ Non-linear. Unless

otherwise mentioned, the standard settings of the program were used.

3. Data

The NLME models were applied to two types of dissolution profiles: one corresponding to an immediate (data A) and another to a slow (data B) dissolving formulation. Data A were obtained from the industry and data B from the literature (Tsong and Hammerstrom, 1994). The same data were already analysed by principal component analysis (Adams et al., 2001a) and linear mixed effects models (Adams et al., 2001b). Both data A and B consist of a reference and a test set. For each set, 12 tablets were measured at different time points.

For data A, measurements were performed at 15, 30, 45 and 60 min. The 12 dissolution profiles of the reference set, together with three profiles of the test set, are shown in Fig. 1. Notice that the percentages at the first time point (15 min) are already relatively high. The f_2 factor for these batches amounts to 83 so that they can be considered as pharmaceutically equivalent according to the present FDA guidelines ($f_2 > 50$).

For data B, the percentages dissolved were determined at 1, 2, 3, 4, 6, 8 and 10 h. So, these time points are unequally spaced. Fig. 2 shows that the curves of Br increase more gradually than the curves of data A. Since the f_2 factor for data B is 64, both batches are also considered as pharmaceutically equivalent.

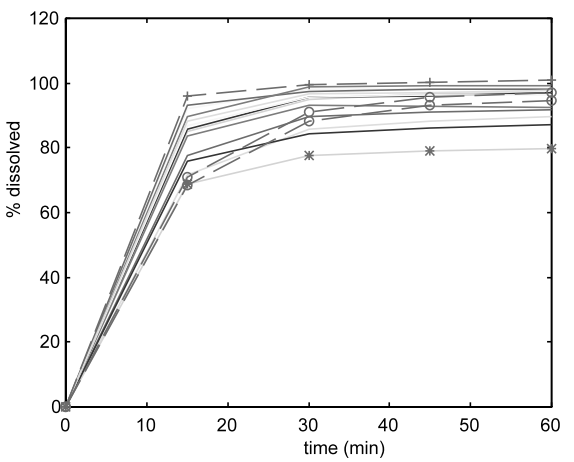

Fig. 1. Data A: the 12 dissolution profiles of the reference set with dissolution profile 8 indicated with '–∗–'. Dissolution profiles 2 and 7 of the test set are indicated by '- -○- -' and profile 11 by '- - + - -'.

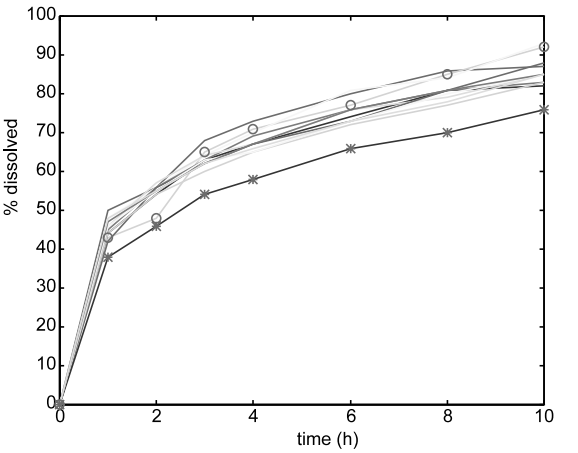

Fig. 2. Data B: the 12 dissolution profiles of the reference set with dissolution profile 2 indicated with '–○–' and profile 9 with '–∗–'.

Remark that neither for data A nor for data B measurements are performed at time point 0. Consequently, the point (0, 0) is not used in the model calculations.

4. Results and discussion

4.1. Data A

4.1.1. Weibull

As described in the working procedure, model Eq. (3) without random effects parameters is fit by ML. $X_{(\infty)}$ is set equal to 100 (%). The resulting parameter estimates for α_0 and β_0 yield a good fit for the mean profile of the reference batch while the combination with α_1 and β_1 does the same for the test batch. The AIC_{ML} amounts to 630.4. The differences indicated by parameters α_1 and β_1 are found not to be significant at the 5% level ($P =$ 0.1255 and 0.1162, respectively) so that both batches can be considered similar. Next the random effects are included to see whether a better fit

is obtained. Inclusion of a_i gives an AIC_{ML} of 516.4, of b_i 501.3 and of a_i and b_i 483.8. According to the LRT, the latter model is clearly the best. The final REML estimates are given in Table 1. This table also contains the P-values for α_1 (0.0077) and β_1 (0.0278). They now indicate that both batches are different at the 5% level. A rather strong negative correlation (-0.834 for both the reference and the test set) is found between the scale (α) and the shape factor (β) of the Weibull equation. This means that when α increases, β decreases and vice versa. The behaviour of the individual tablets is shown in the random effects plot of Fig. 3(I). Tablet 8, which has the lowest profile in Fig. 1, has the lowest b parameter of the reference set. For the test set, tablets 2 and 7 (14 and 19 in Fig. 3) show the most deviant values in a. It can be seen in Fig. 1 that their dissolution starts relatively low and ends relatively high. Fig. 4(I) shows the diagnostic plots. The spread of the residuals versus the fitted values is rather homogeneous although some $\cap$-shape can be recognised. This indicates that some information is still left in the residuals. This is also reflected in the autocorrelogram where the relatively high value at lag 2 rather indicates a somewhat deficient model than an important correlation.

The influence of taking other values than 100 for $X_{(\infty)}$ is also examined. When $X_{(\infty)}$ is considered as an additional parameter to be determined, no convergence is obtained. Also when $X_{(\infty)}$ is set to 95 instead of 100, no convergence is obtained. When 105 is used, convergence is reached, but the AIC increases to 526.2.

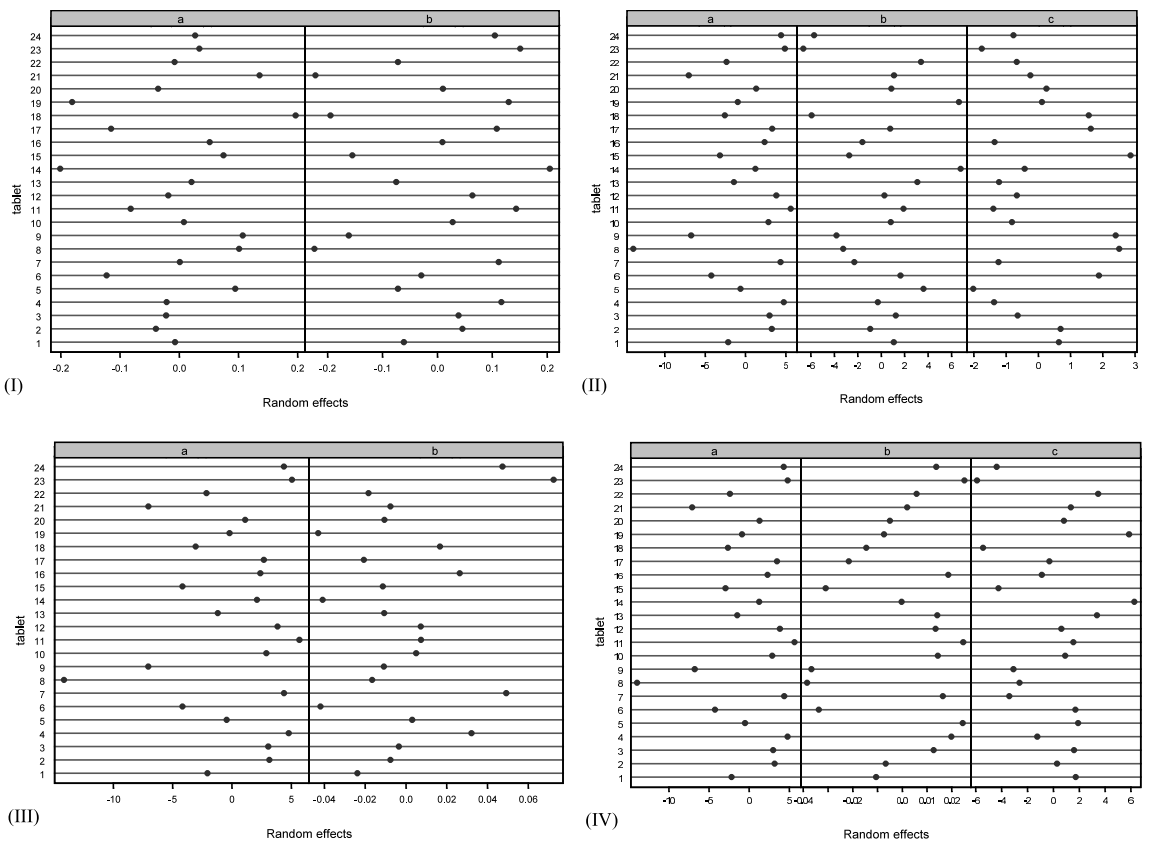

Fig. 3. Plot of the random effects a_i, b_i and eventually c_i for the tablets of data A (1–12: reference, 13–24: test set); (I) Weibull, (II) logistic, (III) first-order and (IV) Gompertz.

Table 1
REML estimates of the model parameters for data A

Parameter	Weibull		Logistic		First-order		Gompertz	
	Estimate (S.E.)	P-value (t-test)	Estimate (S.E.)	P-value (t-test)	Estimate (S.E.)	P-value (t-test)	Estimate (S.E.)	P-value (t-test)
α_0	0.546053	<0.0001	93.59588	<0.0001	93.60919	<0.0001	93.63329	<0.0001
	(0.0572915)		(1.400542)		(1.432800)		(1.404269)	
α_1	−0.204981	0.0077	2.14167	0.2836	1.87041	0.3593	2.19277	0.2737
	(0.0747159)		(1.981411)		(2.026751)		(1.986925)	
β_0	0.453748	<0.0001	1.88042	0.1999	0.14914	<0.0001	0.14861	<0.0001
	(0.0500534)		(1.452498)		(0.008857)		(0.010612)	
β_1	0.160192	0.0278	−2.43549	0.2187	−0.02076	0.1011	−0.04312	0.0019
	(0.0712604)		(1.961400)		(0.012496)		(0.013324)	
γ_0			6.38318	<0.0001			0.23360	0.8589
			(0.627369)				(1.308823)	
γ_1			2.74135	0.0020			−2.56427	0.1586
			(0.850520)				(1.798797)	
S.D. (a_i)	0.133819		4.842544		4.947703		4.855198	
S.D. (b_i)	0.148115		4.105160		0.030200		0.024957	
S.D. (c_i)			1.725224				3.739879	
S.D. (ε_{ij})	1.504193		0.371179		0.619411		0.373769	
AIC_{ML}[a]	483.8		400.3		413.3		398.0	

[a] AIC values of the ML estimates are given to allow model comparison.

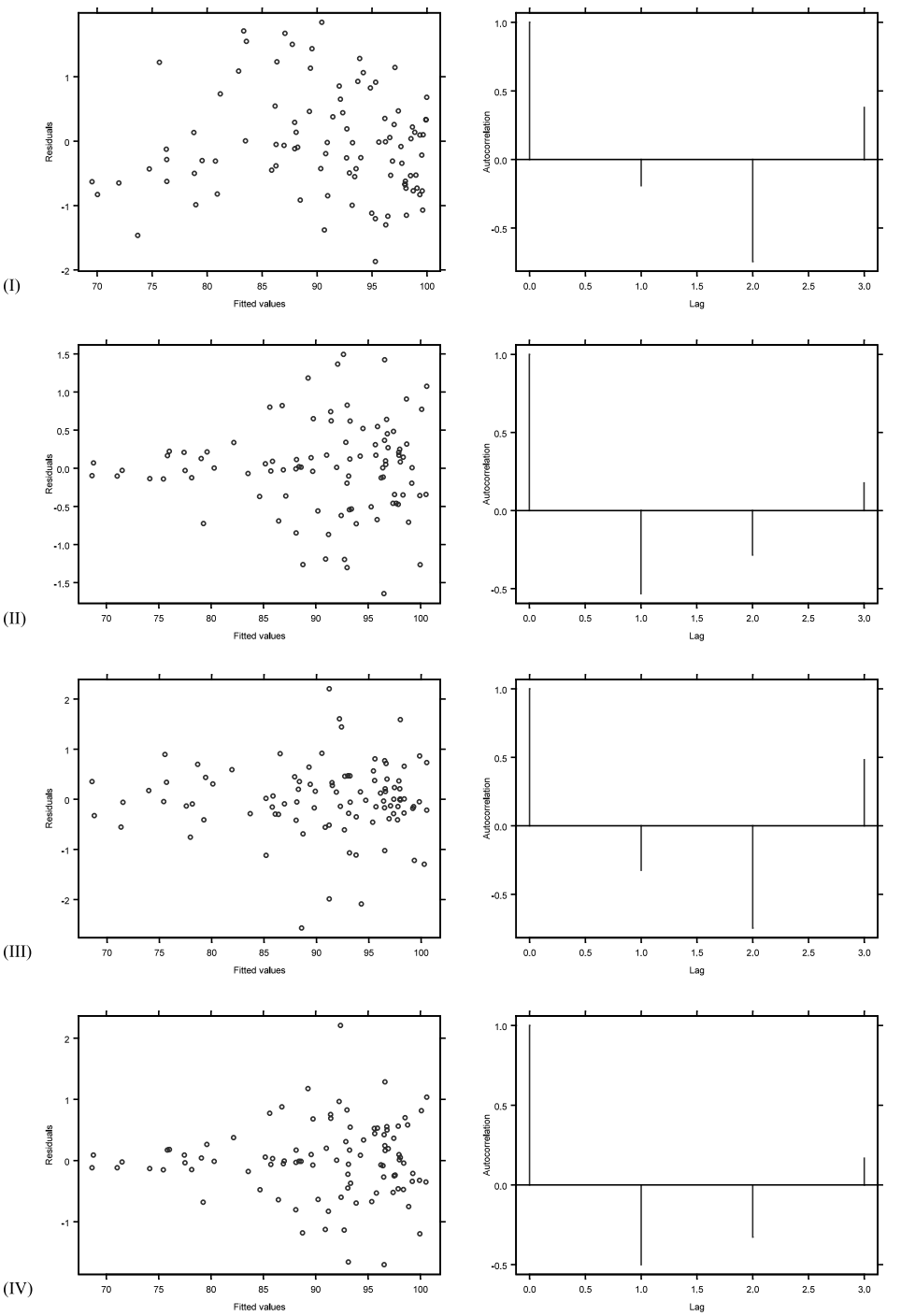

Fig. 4. Diagnostic plots for data A: residuals versus fitted values (left) and autocorrelation of the residuals (right); (I) Weibull, (II) logistic, (III) first-order and (IV) Gompertz.

4.1.2. Logistic

Fitting model (5) without random effects parameters yields an AIC_{ML} of 627.2. No significant differences are observed between both batches since the P-value for $\alpha_1 = 0.3161$, for $\beta_1 = 0.8763$ and for $\gamma_1 = 0.6659$. Next the random effects are added to the formula. Seven combinations are possible: (a_i), (b_i), (c_i), (a_i b_i), (a_i c_i), (b_i c_i) and (a_i b_i c_i). The model including all random effects is found to be the best (AIC_{ML}: 400.3). The final REML estimates for the parameters are also shown in Table 1. The P-values indicate that both batches do not differ in α or β (both batches approach the same asymptote for $t \to \infty$ and also their dissolution rate is similar). The interpretation of the P-value for γ_1 depends more on the the level of significance that is applied: γ_1 is not significant at a 0.1% level, but it is at a 1% level. However, since the γ parameter describes mainly the behaviour of the curves between 50 and 73% of their maximum dissolution, it is not the most relevant parameter of the logistic function to decide that two sets are different or not. Remark also that the measurements at the first time point are already above 73%. The random effects plot is illustrated in Fig. 3(II). The low profile of tablet 8 of the reference set (Fig. 1) can be clearly recognised in a. This is logical because the numerator in Eq. (5) corresponds to the horizontal asymptote when $t \to \infty$. Tablets 2 and 7 of the test set (14 and 19 in Fig. 3(II)) are clearly deviating in b. Since this parameter describes when the curve reaches half of its asymptotic height, it can be correlated to the shape of both profiles shown in Fig. 1. As already mentioned for γ above, also the plot for c (the random effect parameter of γ) contains no important information. The diagnostic plots can be found in Fig. 4(II). The residuals show a dispersing pattern (fan shape) indicating that the lower percentages (measured at the first time point) are better fitted than the higher ones (later measurement times). The autocorrelation of the residuals is good.

4.1.3. First-order

When model (6) is fitted without random effects an AIC_{ML} of 627.5 is obtained. The P-values of α_1 (0.2875) and β_1 (0.1310) indicate that both batches are equivalent. Adding the random effects gives the best results when both a_i and b_i are included (AIC_{ML}: 413.3). The parameter values, estimated by REML, are given in Table 1. Both batches can still be considered as similar (see P-values for α_1 and β_1). The behaviour of each tablet is illustrated in Fig. 3(III). As expected, tablet 8 of the reference set is recognised in a ($\sim$ horizontal asymptote). Tablets 2 and 7 of the test set (14 and 19 in Fig. 3(III)) are less striking than in the previous plots. In the test set, the b-value for tablet 11 (23 in the figure) is high and indicates a fast release rate. This is confirmed by the dissolution curve in Fig. 1. The diagnostic plots of Fig. 4(III) do not show something special: the residuals are rather homogeneously spread and only the autocorrelation at lag 2 is relatively large (similar to the autocorrelogram of the Weibull).

4.1.4. Gompertz

The AIC_{ML} for model Eq. (7) without the random effects parameters amounts to 627.2. According to the P-values for α_1 (0.3187), β_1 (0.8346) and γ_1 (0.7180), both sets can not be considered different. Adding the random effects, the model with random effects for all fixed effects is also here found to have the lowest AIC_{ML}: 398.0. For some combinations of random effects (only c_i or b_i c_i), the calculations do not converge. The REML estimates for the parameters of the best Gompertz model can be found in Table 1. At a 0.1% level of significance, none of the difference indicating parameters (α_1, β_1 and γ_1) is significant, but at a 1% level β_1 is. The random effects plot is given in Fig. 3(IV). Tablet 8 of the reference batch can be recognised in a ($\sim$ horizontal asymptote). Tablets 2 and 7 of the test set (14 and 19 in Fig. 3(IV)) have deviant values for c which can be explained by the different dissolution rate. The relatively small differences between the b_i-values indicate that all profiles have nearly the same curvature. This is also reflected in the relatively small standard deviation of b_i (see S.D. (b_i) in Table 1). So, the most relevant differences between the tablets can be derived from a_i and c_i. This also implies that the difference between both batches indicated by β_1, is not so important. The

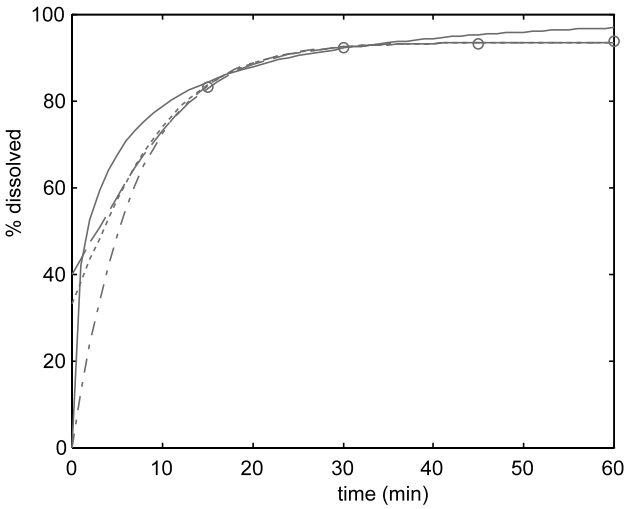

Fig. 5. Average percentages at each time point for the reference set of data A (○) together with the fitted profiles using the parameters of Table 1: '—' Weibull, '– – –' logistic, '– - – - –' first-order and '- - -' Gompertz.

spread of the residuals and the autocorrelation (Fig. 4(IV)) are comparable to those obtained with the logistic function.

4.1.5. Conclusions for data A

For data A, the model with the lowest AIC_{ML} or best fit is the Gompertz function, although the AIC_{ML} of the logistic is only slightly higher (Table 1). Since both models are not nested, they can not be statistically compared by the LRT. The Weibull however, which is usually assumed to give good fittings, shows here the highest AIC value. The four functions are also illustrated in Fig. 5, where the fixed effects parameters of the reference set are used to fit the average percentages. Although (0,0) is not used in the calculations, the Weibull and the first-order go through the origin and so simulate better a natural dissolution profile. The two other functions (logistic and Gompertz) transsect the Y-axis above the origin. A consequence is that the first measurement point is fitted better compared to the later ones, resulting in a fan shaped 'residuals versus fitted values' plot. Tablets with a deviant profile are not always clearly recognised in the random effects plot of a model. For example, tablet 8 of the reference set can be noticed more easily in the random effects plot of the logistic, first-order and Gompertz than in the Weibull plot. Also the final decision that two batches are statistically similar or not, depends on the model chosen to fit the data. According to the f_2 factor of 83, both sets can be considered as pharmaceutically equivalent.

Compared to the classical non-linear regression, which fits no random effects, the AIC_{ML} values for the NLME models are clearly better. For all four model functions, the best fit is obtained when random effects are included for all fixed effects parameters. However, one has to be careful with the interpretation of the results: the introduction of random effects in the formula implies an accurate fit for each tablet, but this does not necessarily mean that the average curve is fitted better. The NLME approach is also more discriminative since it indicates differences between two sets that are not recognised in the classical approach. The NLME models can also be compared with the best LME model, as applied by Adams et al. (2001b). The best LME model yields an AIC_{ML} of 387.2 what is marginally better than for the best NLME model (AIC_{ML}: 398.0).

An important drawback of the non-linear approach in general is that the 'chance' of obtaining convergence is dependent on the choice of the starting estimates for the parameters to be determined. In several cases, the initially chosen starting values had to be adapted.

4.2. Data B

4.2.1. Weibull

Similar to data A, the Weibull model is also examined for data B. Eq. (3) fitted with only fixed effects, yields an AIC_{ML} of 892.5. The P-values for α_1 and β_1 clearly indicate that both sets are statistically different (both < 0.0001). The best NLME model is that with both random effects a_i and b_i (AIC_{ML}: 814.3). This value can be further improved by adjusting $X_{(\infty)}$. As for data A, no convergence is obtained if $X_{(\infty)}$ is added as an additional variable to the formula. When $X_{(\infty)}$ is changed manually, a minimum AIC_{ML} value of 772.4 is found for $X_{(\infty)} = 120$. The parameter values, estimated by REML, are shown in Table 2.

Table 2
REML estimates of the model parameters for data B

Parameter	Weibull		Logistic		First-order		Gompertz	
	Estimate (S.E.)	P-value (t-test)	Estimate (S.E.)	P-value (t-test)	Estimate (S.E.)	P-value (t-test)	Estimate (S.E.)	P-value (t-test)
α_0	0.457536	<0.0001	86.59145	<0.0001	79.39921	<0.0001	88.24694	<0.0001
	(0.0086140)		(1.363773)		(2.112428)		(1.472546)	
α_1	−0.094145	<0.0001	5.31368	0.0041	10.40702	0.0010	5.76738	0.0038
	(0.0119289)		(1.819484)		(3.102215)		(1.958611)	
β_0	0.426894	<0.0001	0.66304	<0.0001	0.60409	<0.0001	0.29338	<0.0001
	(0.0107784)		(0.075985)		(0.044386)		(0.017211)	
β_1	0.178785	<0.0001	1.03951	<0.0001	−0.19986	0.0009	0.06428	0.0038
	(0.0154904)		(0.090355)		(0.059216)		(0.021857)	
γ_0			2.69979	<0.0001			−0.42125	<0.0001
			(0.150887)				(0.091614)	
γ_1			−0.58130	0.0008			1.17561	<0.0001
			(0.169475)				(0.101486)	
S.D. (a_i)	0.024553		3.261562				3.361828	
S.D. (b_i)	0.029178							
S.D. (c_i)	1.803323		2.385660		5.335579		2.169044	
S.D. (ε_{ij})	772.4		831.2		1051.3		803.7	
AIC_{ML}[a]								

[a] AIC values of the ML estimates are given to allow model comparison.

The P-values for α_1 and β_1 remain very small (< 0.0001) and indicate that both sets differ in both the scale (α) and shape (β) factor. Compared to data A, the correlation between α and β is not very important here (-0.477 for the reference and -0.471 for the test set). The random effects plot for each tablet is given in Fig. 6(I). The most remarkable tablet for parameter a is tablet 9 of the reference set. It can be seen in Fig. 2 that this tablet has the lowest profile. Concerning random effect b, tablet 2 of the reference set attracts attention. Fig. 2 shows that the profile of tablet 2 is deviant at the second time point. The diagnostic plots are illustrated in Fig. 7(I): the residuals are homogeneously spread and the autocorrelation for lag > 0 is small.

4.2.2. Logistic

Without taking care of the random effects, the AIC_{ML} of model (5) amounts to 902.4. The P-value for α_1 is 0.0051, for $\beta_1 < 0.0001$ and for γ_1 0.0159. Examining the random effects reveals that fitting all random effects (a_i, b_i and c_i) leads to convergence problems. The best fit is obtained when only a_i is included in the formula (AIC_{ML}: 831.2). So, the main difference between the individual profiles is the value of the horizontal asymptote to which each profile approaches when the time becomes infinite. The REML values are given in Table 2. Note that the addition of random effects for α has also an influence on the estimation of the fixed effects parameters (e.g. the P-value of γ_1 decreases from 0.0159 to 0.0008). Even at a low level of significance (e.g. 0.01%), both batches can not be considered as similar because of β_1 (both batches differ in their dissolution rate). The behaviour of the individual tablets is shown in Fig. 6(II). The lower profile of

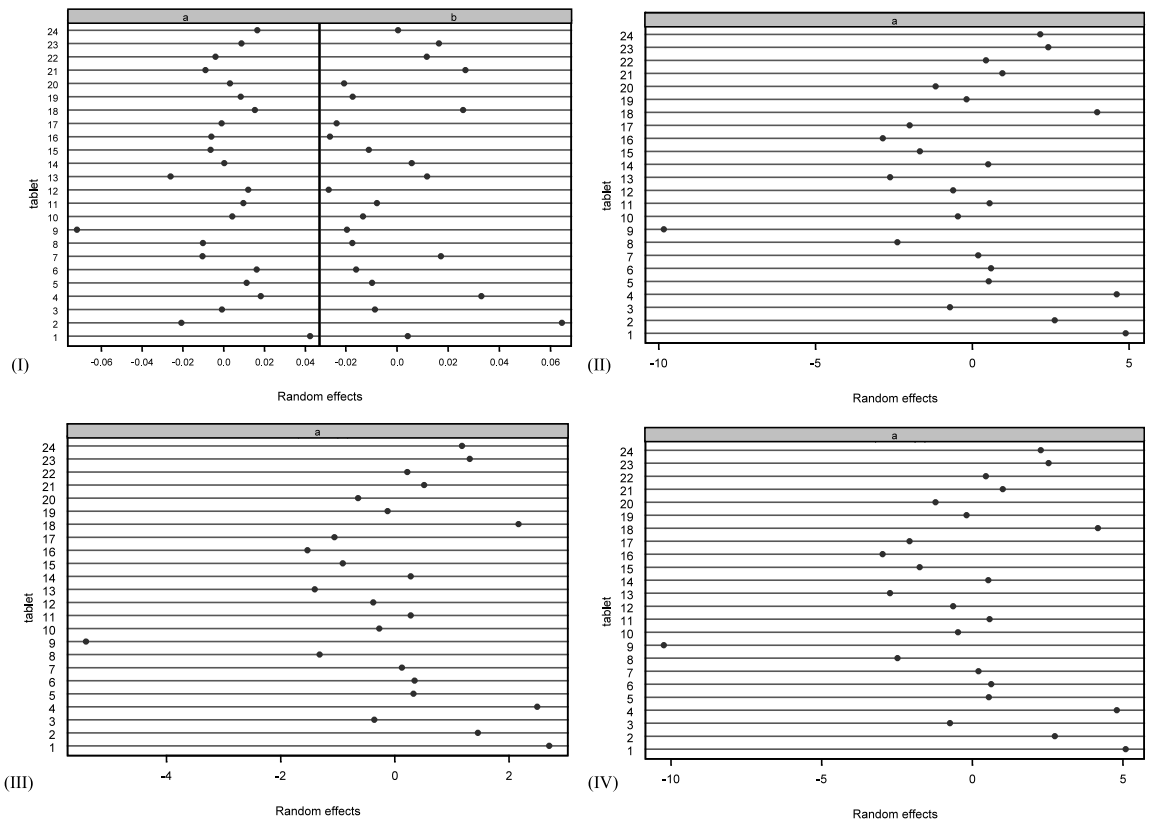

Fig. 6. Plot of the random effects a_i and eventually b_i for the tablets of data B (1–12: reference, 13–24: test set); (I) Weibull, (II) logistic, (III) first-order and (IV) Gompertz.

reference tablet 9 can be clearly recognised in a ($\sim$ horizontal asymptote). As expected, the dip in the profile of tablet 2 is not reflected in a. In Fig. 7(II), the residuals show a somewhat wavy pattern which is reflected in the relatively large autocorrelation at lag 6. As mentioned before this is not very relevant for the study of the correlation, but indicates rather that the model is somewhat deficient and does not fit all information included in the data.

4.2.3. First-order

Fitting model Eq. (6) with only fixed effects parameters resulted in an AIC_{ML} of 1051.3. All parameters are significant (P-values < 0.0001) so that both batches are statistically different. The study of the random effects leads to convergence problems when only b_i is included. The model with the lowest AIC_{ML} (1041.9) includes only a_i as random effect. However, at a 5% level of significance, the latter model is not significantly better than the first one without random effects ($P_{LRT} = 0.0655$). Since the model without random effects contains less variables to be estimated, it is considered the best. The REML estimates for the fixed effects parameters can be found in Table 2. The estimates for the fixed effects parameters of the first-order model including a_i are nearly the same, but the P-values for α_1 and β_1 are different: the most refined model (with a_i) indicates a difference between both batches (P-values < 0.0001) while the simpler model is less discriminative (P-values 0.0010 and 0.0009, respectively). An advantage of the more complex model is that deviations of individual tablets can be traced in the random effects plot (Fig. 6(III)). It is logical that tablet 9 has a clearly deviant value for a ($\sim$ horizontal asymptote), while the dip in the curve of tablet 2 has no pronounced influence on this parameter. Residual plots are shown for both versions of the first-order function (Fig. 7(III)). A similar U-shaped profile can be noticed in the two cases. A high autocorrelation is found at lag 6 if random effects are included, but as mentioned above, this is not relevant.

4.2.4. Gompertz

As in the previous examples, the model equation is first fitted using only the fixed effects parameters (AIC_{ML}: 889.8). The P-values for α_1, β_1 and γ_1 are 0.0111, < 0.0001 and 0.0465, respectively. The best Gompertz model for data B is obtained when only the random effects parameter a_i is included (AIC_{ML}: 803.7). As for the logistic function, the individual tablets differ mainly in the asymptotic value when $t \to \infty$. The REML estimates are given in Table 2. The most important difference between both batches is found in γ_1, which indicates a difference in the dissolution rates. The random effects plot can be found in Fig. 6(IV). As expected, tablet 9 attracts attention and the random effect of tablet 2 is not deviant from the other. The residuals are somewhat wavy and the autocorrelation is good (Fig. 7(IV)).

4.2.5. Conclusions for data B

The best fitting model for data B is the Weibull function with $X_{(\infty)} = 120$ ($AIC_{ML} = 772.4$). However, when $X_{(\infty)}$ is set equal to 100 as usual, the AIC_{ML} increases to 814.3. In this case the Gompertz function shows a better fit ($AIC_{ML} = 803.7$). In contrast with data A, it was not always necessary for data B to add all random effects in the model. The application of the first-order function even revealed that the addition of random effects did not improve the fit. The fixed effects parameters of the four models studied are used to fit the average percentages of the reference batch as illustrated in Fig. 8. The logistic and the Gompertz do not pass through the origin (they also show similar diagnostic plots as follows from Fig. 7). Differences between tablets are mainly observed in the level of the asymptotic value of the profiles for an infinite time as reflected in the a_i parameter of the logistic, first-order and Gompertz function. The somewhat deviant percentage at the second time point of profile 2 of the reference set can only be recognised in the b_i parameter of the Weibull. The more refined the model, the more discriminative it is. This is illustrated in Table 2 by the P-values of α_1 and β_1 for the Weibull (best model) and the first-order (worst model). In general, both sets differ mainly in their dissolution rate. Notice that the f_2 factor indicates that both sets are pharmaceutically similar ($f_2 = 64$). The AIC_{ML} value of the best fitting LME model as

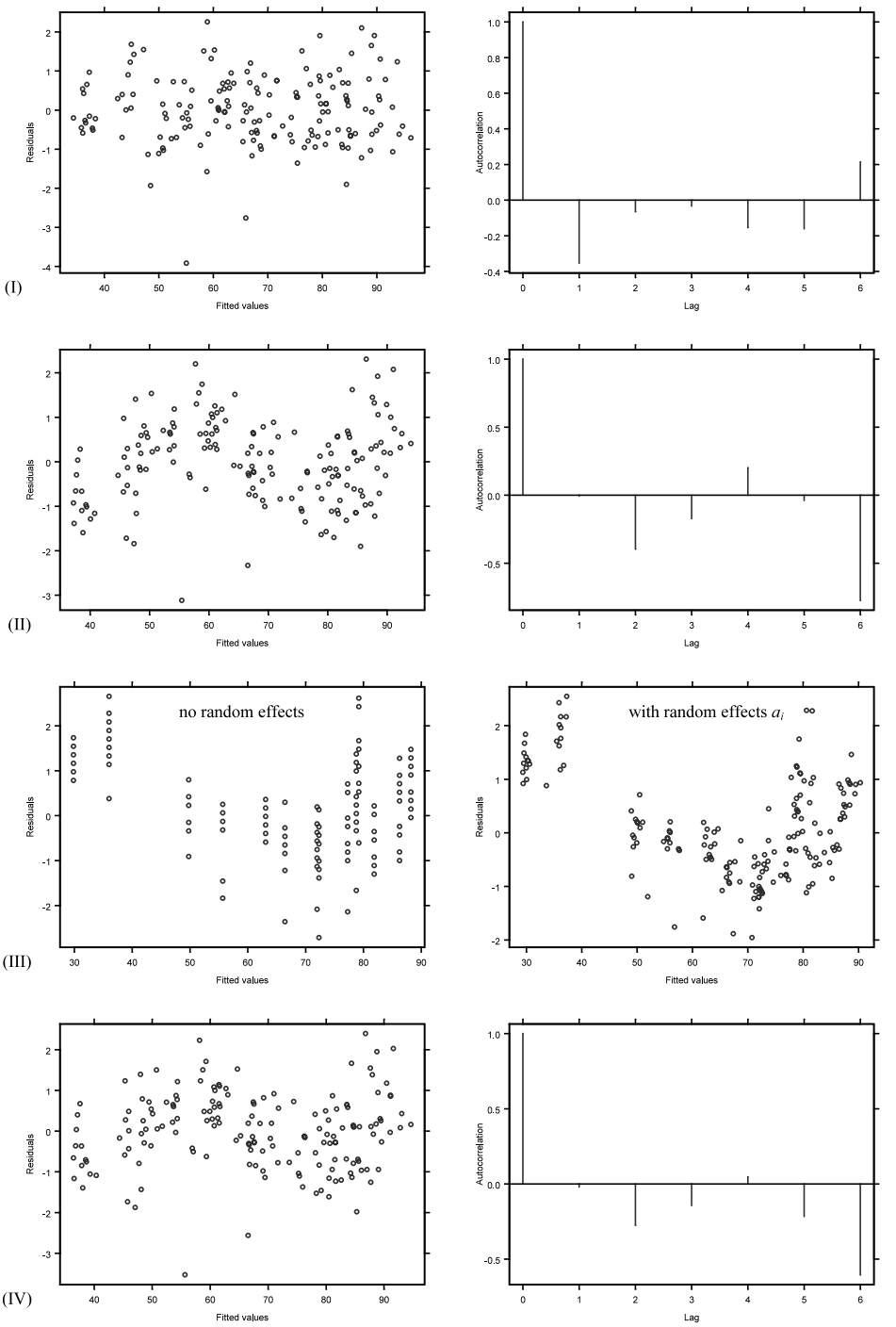

Fig. 7. Diagnostic plots for data B: residuals versus fitted values (left) and autocorrelation of the residuals (right, except for III); (I) Weibull, (II) logistic, (III) first-order and (IV) Gompertz.

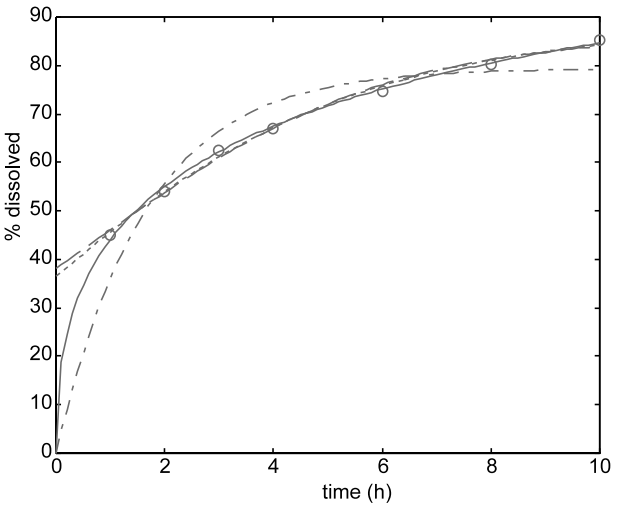

Fig. 8. Average percentages at each time point for the reference set of data B (○) together with the fitted profiles using the parameters of Table 2: '—' Weibull, '– – –' logistic, '– - – - –' first-order and '- - -' Gompertz.

described in Adams et al. (2001b) amounted to 773.4. This is as good as the best NLME model studied here.

As for data A, sometimes several attempts are necessary to obtain convergence due to 'badly chosen' starting values.

5. Conclusion

NLME models suffer from the same inherent disadvantages of modelling. None of the models is suitable to fit all kinds of dissolution curves with sufficient accuracy. Compared to LME models, the NLME models give worse (data A) to similar (data B) results. An attendant advantage of the LME models is that they are easier to apply because one does not have to choose between different types of models and no starting values have to be estimated. It is considered that NLME models have the benefit of a better prediction of responses outside the range of observed data. As a matter of fact, in the cases studied here, this is only valid for the Weibull and the first-order since they pass through (0, 0). For both LME and NLME models, convergence problems can occur and it is not always easy to fit all information available in the original dissolution data. The decision that 2 batches are similar depends not only on the model (LME, Weibull, logistic, first-order, Gompertz, etc.) chosen, but also on the level of significance used. At the moment, there are no agreements between the regulatory authorities and the pharmaceutical industry which level is pharmaceutically acceptable. Compared to the f_2 factor both LME and NLME models are more discriminative and informative.

Acknowledgements

This research was financed by a post-doctoral fellowship of the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT), Brussels.

References

Adams, E., De Maesschalck, R., De Spiegeleer, B., Vander Heyden, Y., Smeyers-Verbeke, J., Massart, D.L., 2001a. Evaluation of dissolution profiles using principal component analysis. Int. J. Pharm. 212, 41–53.

Adams, E., Coomans, D., Smeyers-Verbeke, J., Massart, D.L., 2001b. Application of linear mixed effects models to the evaluation of dissolution profiles. Int. J. Pharm. 226, 107–126.

Anderson, N.H., Bauer, M., Boussac, N., Khan-Malek, R., Munden, P., Sardaro, M., 1998. An evaluation of fit factors and dissolution efficiency for the comparison of in vitro dissolution profiles. J. Pharm. Biomed. Anal. 17, 811–822.

Costa, P., Sousa Lobo, J.M., 2001. Modelling and comparison of dissolution profiles. Eur. J. Pharm. Sci. 13, 123–133.

Crowder, M.J., 1996. Keep timing the tablets: statistical analysis of pill dissolution rates. Appl. Stat. 45, 323–334.

FDA Guidance for Industry, 1995. Immediate Release Solid Oral Dosage Forms—Scale-up and Postapproval Changes: Chemistry, Manufacturing and Controls; In Vitro Dissolution Testing and In Vivo Bioequivalence Documentation. Food and Drug Administration, Center for Drug Evaluation and Research, Rockville, MD, November, 1995.

FDA Guidance for Industry, 1997. Dissolution Testing of Immediate Release Solid Oral Dosage Forms. Food and Drug Administration, Center for Drug Evaluation and Research, Rockville, MD, August, 1997.

Langenbucher, F., 1972. Linearization of dissolution rate curves by the Weibull distribution. J. Pharm. Pharmacol. 24, 979–981.

Moore, J.W., Flanner, H.H., 1996. Mathematical comparison of dissolution profiles. Pharm. Technol. 20, 64–74.

Pinheiro, J.C., Bates, D.M., 2000. 'Mixed-Effects Models in S and S-plus, Statistics and Computing'. New York: Springer–Verlag, .

Polli, J.E., Rekhi, G.S., Shah, V.P., 1996. Methods to compare dissolution profiles. Drug Inf. J. 30, 1113–1120.

Polli, J.E., Rekhi, G.S., Augsburger, L.L., Shah, V.P., 1997. Methods to compare dissolution profiles and a rationale for wide dissolution specifications for metoprolol tartrate tablets. J. Pharm. Sci. 86, 690–700.

Sathe, P.M., Tsong, Y., Shah, V.P., 1996. In-vitro dissolution profile comparison: statistics and analysis, model dependent approach. Pharm. Res. 13, 1799–1803.

Sathe, P., Tsong, Y., Shah, V.P., 1997. In vitro dissolution profile comparison and IVIVR. Carbamazepine case. Adv. Exp. Med. Biol. 423, 31–42.

Tsong, Y., Hammerstrom, T., 1994. Statistical issues in drug quality control based on dissolution testing. Proceedings of the Biopharmaceutical Section of the American Statistical Association, pp. 295–300.

Tsong, Y., Hammerstrom, T.H., Sathe, P.S., Shah, V.P., 1996. Statistical assessment of mean differences between dissolution data sets. Drug Inf. J. 30, 1105–1112.

Yuksel, N., Kanik, A.E., Baykara, T., 2000. Comparison of in vitro dissolution profiles by ANOVA-based, model-dependent and -independent methods. Int. J. Pharm. 209, 57–67.