

KinetDS: An Open Source Software for Dissolution Test Data Analysis

Aleksander Mendyk^{1,*}, Renata Jachowicz¹, Kamil Fijorek²,
Przemysław Dorożyński¹, Piotr Kulinowski³, and Sebastian Polak⁴

e-mail: mfmendyk@cyf-kr.edu.pl

¹Department of Pharmaceutical Technology and Biopharmaceutics, Faculty of Pharmacy, Jagiellonian University – Medical College, Medyczna 9 St., 30-688 Kraków, Poland

²Department of Statistics, Faculty of Management, Cracow University of Economics, Rakowicka 27 Str., 31-510 Kraków, Poland

³Institute of Technology, Pedagogical University, Podchorążych 2 St., 30-084 Kraków, Poland

⁴Unit of Pharmacoepidemiology and Pharmacoeconomics, Faculty of Pharmacy, Jagiellonian University – Medical College, Medyczna 9 St., 30-688 Kraków, Poland

ABSTRACT

Because drug quality is the focus for pharmaceutical industry and regulatory agencies, the in vitro dissolution test becomes a standard tool for characterization of manufactured products. However, results of the dissolution test must be expressed in mathematical terms; this is realized by fitting various models to the cumulative dissolution curves. The models might be either mechanistic or empirical. The fitting process requires software (e.g., *KinetDS*) for automation and determination of possible release mechanisms of drug substances from the dosage forms. The software is FOSS (Free Open Source Software) and is available at <http://sourceforge.net/projects/kinetds/>.

INTRODUCTION

Bioequivalence (BE) is the basic criterion of similarity between two pharmaceutical products, namely generic and innovator drugs. Due to the probabilistic character of BE definition, pharmaceutical equivalence (PE) is mandatory for the pharmaceutical industry. In comparison with bioassays, it is based on the more reproducible in vitro dissolution tests. The latter produce time profiles of the drug amount released from the dosage forms, which are expressed in their cumulative form. An expression of dissolution profile similarity is based on two main approaches: (1) the direct comparison of the curves (1) (i.e., by similarity factor f_2), and (2) use of mathematical modeling to parameterize curves and compare resulting parameters of the models.

The latter method is also recognized as an element of a new development strategy of pharmaceutical products (2, 3), process analytical technology (PAT), and implicates possible mechanisms of drug release from the dosage form, which influences bioavailability (4). Curve fitting of drug release data is also commonly used as a step for in vitro–in vivo correlation (IVIVC) development. The fitting process of the model parameters requires software, which was the aim of this work and resulted in the *KinetDS* project. Its main objectives were intuitiveness of the graphical user interface (GUI), a high level of software automation, and free availability for commercial and noncommercial use.

FEATURES AND THEORETICAL BACKGROUND

KinetDS is software for curve fitting, designed particularly to describe the cumulative dissolution curve

by a simple equation or set of equations. However, other curves, derived from different data sources, might be also analyzed if their dependent-variable range is between 0 and 100. The equations were chosen from the most popular mechanistic and empirical models (Table 1) applied to the drug dissolution curve description (5) such as zero- to third-order kinetic models, Higuchi, Korsmeyer–Peppas, Weibull (2- and 3-parameter), Hixson–Crowell, Michaelis–Menten, and Hill equation.

Model-independent descriptions of the dissolution curve are also provided (Table 2) as dissolution efficiency, DE (6), and mean dissolution time, MDT (7). DE and MDT values are automatically computed for every available dataset and included at the end of the report file.

The best model choice is based on model performance on the available data. Several measures of goodness of fit are implemented (Table 3): (1) coefficient of determination, R^2 , and empirical coefficient of determination, R^2_{emp} ; (2) Akaike Information Criterion, AIC (8); (3) Bayesian Information Criterion or Schwarz criterion, BIC (9); and (4) root-mean-squared error, RMSE.

Additional diagnostics of the linear regression function is the standard error of the regression coefficients (10), also expressed in percent of the relevant parameter value (relative standard error of the regression coefficient).

Model parameters are established by linear and nonlinear regression (NLR). The latter is represented here by the simplex method. So far, all the models included in *KinetDS 3.0* are susceptible to linearization. Therefore, they might be fit by both linear regression and NLR. It is planned in Version 4.0 to introduce models suitable only for nonlinear regression. Regarding mathematical simplicity of the kinetic models from zero- to third-order, they usually do not require NLR. Specifically, because of its

*Corresponding author.

Table 1. Chosen Models Available in KinetDS 3.0

Model	Equation
Zero-order	$Q = k \cdot t + Q_0$
First-order	$\frac{1}{Q} = k \cdot t + \frac{1}{Q_0}$
Second-order	$\frac{1}{Q^2} = k \cdot t + \frac{1}{Q_0^2}$
Higuchi	$Q = k \cdot \sqrt{t}$
Korsmeyer–Peppas	$Q = k \cdot t^n$
Weibull with the lag time	$Q = 100 \cdot \left(1 - \exp \left[\frac{-(t - T_{LAG})^b}{a} \right] \right)$
Hickson–Crowell with lag time	$Q^{\frac{1}{3}} = k \cdot (t - T_{LAG}) + Q_0^{\frac{1}{3}}$
Michaelis–Menten	$Q = \frac{Q_{max} \cdot t}{k + t}$
Hill	$Q = \frac{Q_{max} \cdot t^n}{k^n + t^n}$

Q: amount (%) of drug substance released at the time t
 Q₀: start value of Q
 Q_{max}: maximum value of Q (100%)
 T: time
 k, a, b: constants
 T_{LAG}: lag time
 For a full list of the models, please refer to the manual available at <http://sourceforge.net/projects/kinetds/>.

linear nature, the zero-order model is excluded from the NLR procedure. Thus, the user checks the box “use nonlinear regression (NLR)” without concern for the models but only for the expected improvement of model performance. If such improvement is noticed by the software, it will be introduced to the final report. For the reference, the final report always includes both sets of coefficients derived from the linear regression and NLR. Regarding this, it is usually advisable to check the box “use nonlinear regression (NLR)” as a possible way to improve the model. However, there is a price for this in terms of computational resources required. Because NLR is based on the iterative

Table 2. Model-Independent Measures of Dissolution Process

Measure	Equation
DE (Dissolution Efficiency)	$DE = \frac{\int_0^t Q dt}{Q_{max} \times t} \times 100$
MDT (Mean Dissolution Time)	$MDT = \frac{\sum_{j=1}^n t_j^{AV} \times \Delta Q_j}{\sum_{j=1}^n \Delta Q_j}$ $\Delta Q = Q_{(t)} - Q_{(t-1)}$ $t_j^{AV} = (t_j + t_{j-1})/2$

Q_{max}: maximum amount of drug released (= 100%)
 n: number of timepoints

Table 3. Best Model Criteria

Measure	Equation
Empirical R ² (R ² _{emp})	$R_{emp}^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - y_{AV})^2}$
Akaike information criterion (AIC)	$AIC = 2k + N \cdot \left[\ln \left(\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N} \right) \right]$
Bayesian information criterion or Schwarz criterion (BIC)	$BIC = N \cdot \ln \left(\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N} \right) + k \cdot \ln(N)$
Root-mean-squared error (RMSE)	$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}}$

y_i: observed value
 ŷ_i: model-predicted value
 y_{AV}: average output value

procedure (simplex method), its convergence might be time-consuming in certain cases.

Linear regression provides starting points for nonlinear regression; therefore, this information is not required from the user, as opposed to other software such as SPSS (11). An original, linear search algorithm for a lag time was written and implemented for the Weibull, Korsmeyer–Peppas, and Hixson–Crowell models (5). This means that three-parameter models are also manageable by linear regression, which with the standard procedure is able to optimize only two-parameter linear equations. The basic idea behind this technique is that the system searches for the optimal two parameters of the linear transformation (if available) of the models relative to the established third parameter in the specific domain. The latter is very simple to set when assuming that it is T_{LAG}, which must be between 0 and the first dissolution time point. Although from the physical point of view, such assumption does not have to be valid in all cases; from a mathematical point of view, it allows implementation of an iterative search for the optimal T_{LAG} by simple repetition of the linear regression procedure with T_{LAG} adjusted gradually from 0 to the first dissolution time. The algorithm might be summarized in the following steps:

1. Set T_{LAG} value for iteration k (T_{LAG}^(k) = T_{LAG}^(k-1) + step).
2. Perform linear regression with the current value of T_{LAG}^(k) to establish values of remaining two parameters.
3. Compute performance of the obtained model (i.e., RMSE).
4. Set k = k + 1 and if k < k_{MAX}, go back to point 1.

The optimal value of the T_{LAG} is found for the model with the lowest prediction error obtained for the data. This iterative algorithm makes available models more flexible and at the same time is very fast, especially compared with nonlinear regression. For the latter, special procedures were implemented to enhance its predictive ability.

Nonlinear regression is well known for its dependence on the adjusted parameter starting values—usually their random choice is not the best in terms of the final model performance. Therefore, when nonlinear regression is applied, the starting values of the model are those adjusted by the linear regression in the previous run (Figure 1). This is regarded as the use of nonlinear regression for fine-tuning of the model parameters. However, as shown below (Figure 2), linear regression might be misleading in the adjustment of the model parameters, thus additional start-runs were projected for nonlinear regression. This time the user has control over additional runs by using option “NLR random start-point,” which adds two iterations with random estimates of the equation parameters for the simplex method. The procedure chooses the best solution obtained along with the standard nonlinear regression procedure and presents it in the “NLR results” section for each equation report.

To make algorithm-fitting even more flexible, a tool for setting a priori values of linear regression coefficients was introduced as a box in the GUI labeled “set equation params.” A major use of such a tool is usually in setting $b = 0$, thereby forcing the linear function to include the $(0, 0)$ point. This is mainly applicable to linear regression because NLR always adjusts all the coefficients as needed.

KinetDS works in two modes: (1) simple, where a single curve is described by a single model and a ranking of the models is created based on the goodness-of-fit criterion (Table 3), and (2) piecewise, where a single curve is described by a set of models chosen from the internal library.

The latter is especially useful for the detection of the breakpoints between the dominant mechanisms of drug release across the whole time of the dissolution test. Such a situation is quite common with biphasic or multiphase drug-release profiles.

All operations in the piecewise mode are automatic and parameterized by approximation accuracy expressed in the 0–1 range. The higher the accuracy, the more sensitive is the procedure to changes in the goodness-of-fit criterion—this results in a higher number of switch-points. The goodness-of-fit criterion is another parameter of this procedure, thus different fit measures result in different sets of models.

Batch processing is also implemented allowing for processing of an unlimited number of data files. It generates additional summary reports comparing model parameters for all the data files included in the analysis, which might be easily imported to the spreadsheet for further processing. Batch processing is available for both simple and piecewise models, thus the reports available from these modes will be different at the end of the program work. For the simple mode, the report will include parameters of all the models taken into consideration with their diagnostics (goodness of fit). For the piecewise mode, the report will introduce only the models

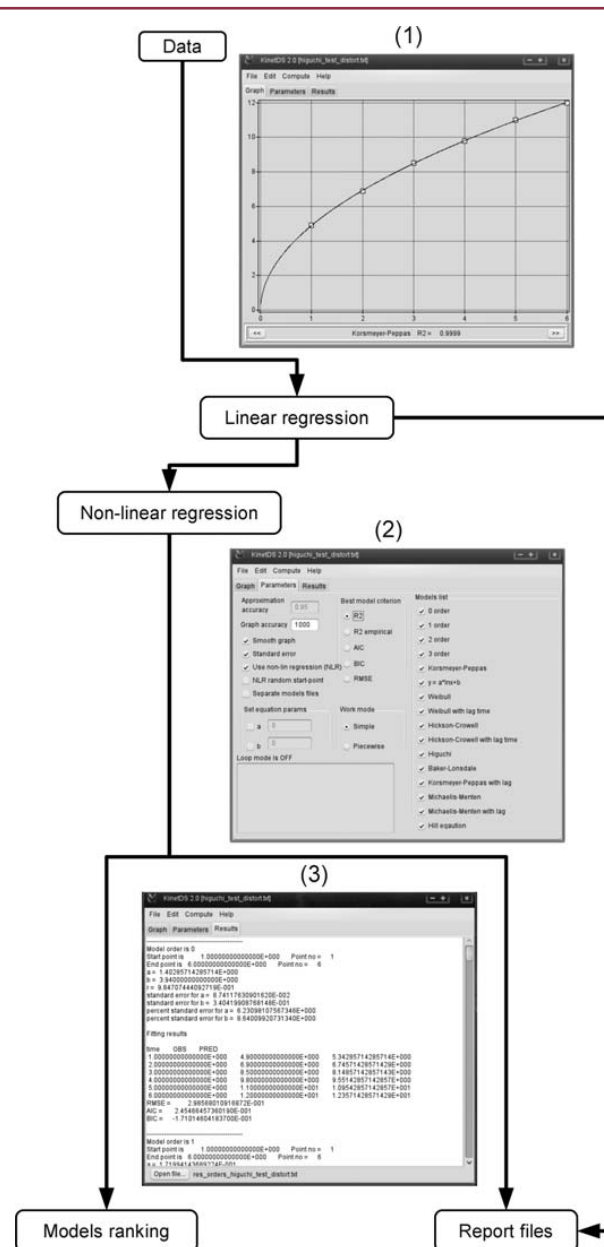


Figure 1. Working scheme of *KinetDS* 3.0. (1) Graph tab with data fitted to the Korsmeyer–Peppas model, (2) adjustable parameters, and (3) report tab.

chosen by the system for particular curve representation together with their domains, namely time points where they are applicable.

Apart from the algorithm-dependent tools, other tools for data handling and presentation are also included. An option of graph-smoothing is available under the checkbox “smooth graph” together with the text field “graph accuracy” denoting the number of steps between the original data points used for graph smoothing. “Graph accuracy” also has meaning for the search algorithm of the T_{LAG} value, where it is also a base of the step used for the time-axis search procedure. Another tool is a checkbox “separate models files” allowing the creation of output files including approximation of the data performed for

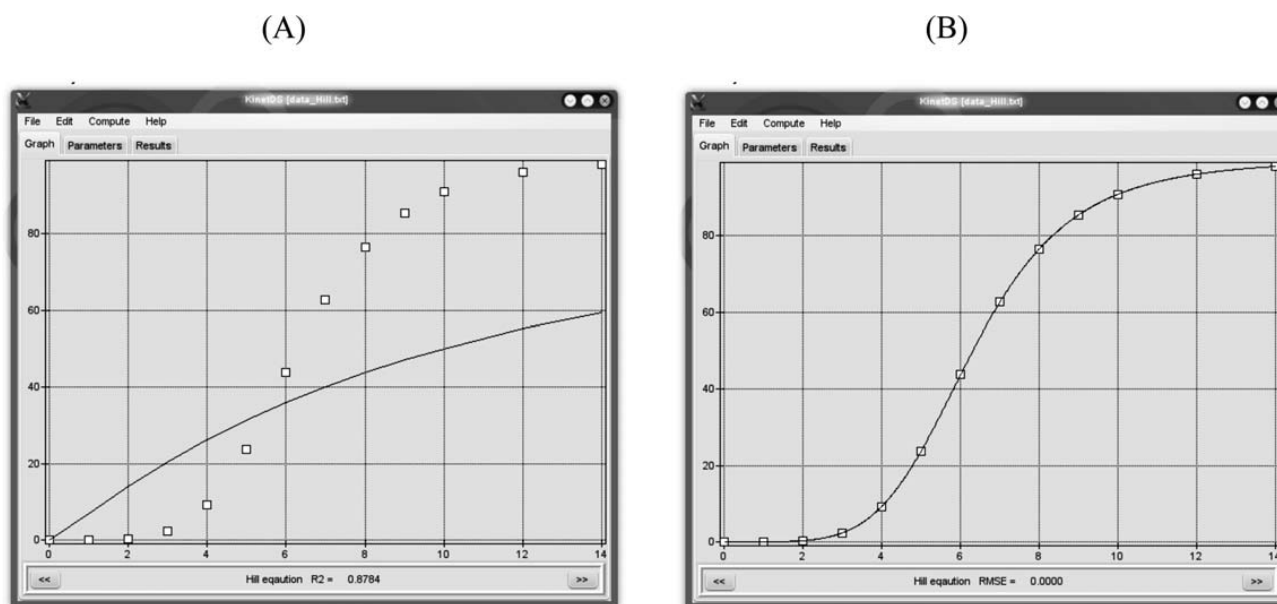


Figure 2. Results of the simple mode tests: (A) linear regression only and (B) nonlinear regression preceded by linear regression.

the particular model. Each file is dedicated to a single model only and allows for the approximation within (interpolation) and beyond (extrapolation) the original data range with specified accuracy. Once the “separate models files” is checked, a popup window is invoked for the range and approximation accuracy settings. After that, the whole procedure of file generation is performed alongside the main run of the software.

Case Study 1: Simple Mode

The simple mode of *KinetDS* was studied using the simulated data from the Hill equation (Table 4). Two main procedures were applied: linear and nonlinear regression.

Table 4. Data for Simple Mode Test of *KinetDS* 3.0

Time	Q (%)
0	0
1	0.01
2	0.32
3	2.37
4	9.29
5	23.81
6	43.74
7	62.70
8	76.62
9	85.52
10	90.91
12	96.14
14	98.17

For linear regression, R^2 was used as the fit criterion, whereas for a nonlinear regression, RMSE was chosen. Additionally, “NLR random start-point” option was enabled.

It is noticeable that for a sigmoid-shape dissolution curve, the linear regression (Figure 2A), although justified by linearization of the Hill model, was ineffective. The only valid results were obtained by nonlinear regression (Figure 2B). *KinetDS* 3.0 was able to reproduce perfectly the original parameters of the Hill equation used for the simulation of the dissolution profile; this might be illustrated by the unrealistic value of RMSE = 0 (Figure 2B). The data were based on the equation, thus the software was able to reproduce it with absolute accuracy; this was the cause of the zero value of RMSE.

Case Study 2: Piecewise Mode

The piecewise mode was also tested on the simulated data to validate the software ability to reproduce assigned models. The dataset was a simple merge between the Korsmeyer–Peppas model and the zero-order model of drug release (Table 5). It was a simulation of biphasic drug release with the first rapid release phase and the second following a linear pattern. The breakpoint between the release mechanisms was established at $t = 11$. As shown in Figure 3, *KinetDS* was able to reproduce the hidden pattern of dissolution with perfect accuracy, however, only for a specific set of nonstandard parameters (i.e., accuracy set to 0.999 and goodness-of-fit criterion set to empirical R^2).

The algorithm of piecewise regression is, in fact, an algorithm of curve morphology detection, and as such, it is very complicated and based on many assumptions. In *KinetDS*, its major version was developed with the simplest approach of a linear regression method; therefore, it is

Table 5. Data for Piecewise Mode Test of KinetDS 3.0

Time	Q (%)
0	0.00
0.2	24.68
0.4	30.39
0.6	34.32
0.8	37.41
1	40.00
2	49.25
3	55.62
4	60.63
5	64.83
6	68.47
7	71.71
8	74.64
9	77.33
10	78.00
11	79.00
12	80.00
13	81.00
14	82.00
15	83.00
16	84.00
17	85.00
18	86.00
19	87.00
20	88.00

justified that R^2 model performance measure was the optimal condition. Based on the above, it is advised that the piecewise regression algorithm in *KinetDS 3.0* should be used mostly with linear regression and its dedicated goodness-of-fit measures, whereas nonlinear regression should be used with care.

Case Study 3: Real-Life Example

The data for this study were derived from a *Dissolution Technologies* paper by Patel et al. (12) concerning dissolution tests of various oxcarbazepine formulations. Since there was no table of raw data available in the paper, graphical tool *g3data* was applied to scan the original graph (Figure 1 in the publication) and extract numbers. A test run with *KinetDS 3.0* was performed with the data series labeled in the original paper as "DCT." A few models were chosen: zero- and first-order, Weibull, and Hixson–Crowell. *KinetDS* was started in the simple mode with NLR enabled. To assess the Hixson–Crowell model for the amount of drug remaining in the formulation, the original dataset was transformed by simple subtraction of original

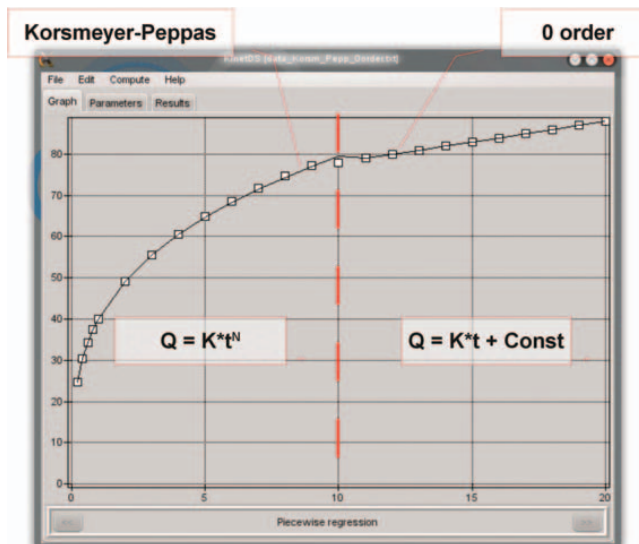


Figure 3. Results of the piecewise mode tests.

values of Q from 100%. The transformed dataset was processed separately by *KinetDS 3.0* with the same parameters as the original one.

In Table 6, the equation parameters found in the original paper (12) are compared with those generated by *KinetDS*. To allow for a direct comparison of the results, only the linear regression results are reported. A good accordance of the fitted values between *KinetDS* and Patel et al. was observed for all the models. Some small discrepancies were found. This might be caused by the graphical scanning procedure and its lack of precision. NLR always allowed performance improvement of such models (i.e., Weibull). The best model found by *KinetDS* was Michaelis–Menten, which was not included in the original paper; therefore, it was impossible to compare results for this particular model.

RELEASE AND FUTURE PERSPECTIVES

KinetDS is written in Object Pascal and developed in the Lazarus RAD (13) environment. There are versions compiled for 32-bit Windows environments (Windows 98, 2000, XP, Vista, and 7), Linux 32- and 64-bit, and Mac OS–X 10.6 64-bit version. The software is available as Free Open Source Software (FOSS) from the SourceForge server (14). It is free for personal and professional use and is distributed under GPL v3 (GNU Public License) together with the source code (15). According to the GPL, *KinetDS* is distributed "as is" without any warranty. Such a model of development allows the developer to focus on software building only. Any particular use and its outcomes are the users' responsibility. However, this task is much easier once the user has the source code available, which might be exploited (i.e., in the software validation procedure).

The source address is <http://sourceforge.net/projects/kinetds/>. So far, for about a year and a half of the project's existence, there were around 1600 downloads worldwide. It is still an ongoing project, and the next release

Table 6. Comparison of Patel et al. (12) Modeling Results with the Values Provided by KinetDS 3.0

Coefficient	Patel et al.	KinetDS 3.0
Zero-order		
K	1.9209	1.9467
R ²	0.8415	0.8490
First-order		
K	0.0234	0.0562
R ²	0.7595	0.7590
Hixson-Crowell		
K	-0.043	-0.0435
R ²	0.8711	0.8820
Weibull		
β	1.0638	1.0891
α	31.22	34.00
R ²	0.9496	0.9541

scheduled in June 2012 will be a major release, where all known bugs will be corrected and extensive validation of the software provided, together with possible extension of piecewise mode.

ACKNOWLEDGMENTS

The Simplex method is based on code by Dr. Jean Debord from his Open Source Pascal library *tpmath* (16). Special thanks to Angus McLean (BioPharm Global, USA) for very valuable remarks and challenging beta tests.

This software was partially supported by a Polish State project carried out at the Pharmaceutical Faculty, Jagiellonian University Medical College, project no K/ZBW/000219.

REFERENCES

- Moore, J. W.; Flanner, H. H. Mathematical Comparison of Dissolution Profiles. *Pharm. Tech.* **1996**, *20*, 64–74.
- Hinz, D. C. Process analytical technologies in the pharmaceutical industry: the FDA's PAT initiative. *Anal. Bioanal. Chem.* **2006**, *384* (5), 1036–1042.
- PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance; Guidance for Industry; U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), U.S. Government Printing Office: Washington, DC, 2004. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070305.pdf> (accessed Jan 20, 2012).
- Cao, Q.-R.; Choi Y.-W.; Cui, J.-C.; Lee, B.-J. Formulation, release characteristics and bioavailability of novel monolithic hydroxypropylmethylcellulose matrix tablets containing acetaminophen. *J. Control. Release* **2005**, *108* (2–3), 351–361.
- Costa, P.; Sousa Lobo, J. M. Modeling and comparison of dissolution profiles. *Eur. J. Pharm. Sci.* **2001**, *13* (2), 123–133.
- Khan, K. A.; Rhodes, C. T. Effect of compaction pressure on the dissolution efficiency of some direct compression systems. *Pharm. Acta Helv.* **1972**, *47* (10), 594–607.
- Podczec, F. Comparison of in vitro dissolution profiles by calculating mean dissolution time (MDT) or mean residence time (MRT). *Int. J. Pharm.* **1993**, *97* (1–3), 93–100.
- Akaike, H. A new look at the statistical model identification. *IEEE T. Automat. Contr.* **1974**, *19* (6), 716–723.
- Schwarz, G. E. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6* (2), 461–464.
- Weisberg S. *Applied Linear Regression*, 3rd ed.; Wiley Series in Probability and Statistics; Wiley-Interscience: Hoboken, NJ, 2005.
- IBM SPSS Web site. <http://www.spss.com> (accessed Jan 20, 2012).
- Patel, N.; Chotai, N.; Patel, J.; Soni, T.; Desai, J.; Patel, R. Comparison of In Vitro Dissolution Profiles of Oxcarbazepine-HP β-CD Tablet Formulations with Marketed Oxcarbazepine Tablets. *Dissolution Technol.* **2008**, *15* (4), 28–34.
- Lazarus Home Page. <http://www.lazarus.freepascal.org> (accessed Jan 20, 2012).
- Sourceforge Home Page. <http://sourceforge.net/projects/kinetds> (accessed Jan 20, 2012).
- GNU Operating System Web site. <http://www.gnu.org/licenses/gpl.html> (accessed Jan 20, 2012).
- Debord, J. Informatique & Chimie. http://www.unilim.fr/pages_perso/jean.debord/index.htm (accessed Jan 20, 2012).