Brief/Technical Note

Using DTA and DTAARRAY Variables and Programming in WinNonlin ASCII Models to Streamline User-Defined Calculation and Data Analysis

Jun Shen,^{1,4,6} Shuanglian Li,² Ronald R. Bowsher,³ and Andrew Vick^{1,5}

Received 24 March 2014; accepted 23 December 2014; published online 14 January 2015

Abstract. In pharmacokinetic (PK) analysis, there are many occasions where user-defined calculations need to be performed before or after the primary PK modeling/analysis. Conventionally, these calculations are often executed outside of the primary PK analysis by pre- or post-processing data from multiple sources, manually entering formulas and multiple additional set-ups. Such analysis approaches increase the risk of generating data defects and can employ software that is not fully compliant. We propose a method of leveraging DTA and DTAARRAY variables plus simple programming techniques in an ASCII model to automate these user-defined calculations in WinNonlin and eliminate the need for manual handling of data outside of the primary analysis. We demonstrated the application of this strategy through three case study examples. In case 1 (post-processing data), DTA variables were used to calculate three user-defined parameters in the primary PK model. In case 2 (pre-processing data), a baseline correction decision tree was programmed into the PK model to account for both the endogenous baseline level as well as the presence of residual drug. In case 3, DTAARRAY variables were used to perform a looping operation to calculate the difference factor (F_1) and the similarity factor (F_2) in support of *in vitro* bioequivalence evaluations.

KEY WORDS: ASCII model; decision tree; DTA; DTAARRAY; programming; user-defined calculation; WinNonlin.

INTRODUCTION

In conventional pharmacokinetic (PK) analysis, there are many occasions where user-defined calculations are performed following primary PK modeling to specifically augment the data analysis. For example, PK study of coagulation products differ from that of most conventional drugs in that plasma levels are quantified by bioassays of coagulation activity rather than plasma concentrations. As such, there are some coagulation-specific PK parameters which include incremental recovery: $K = C_{\max}/\text{dose}$, *in vivo* recovery (%) = $100 \times K \times \text{plasma volume}$ (where plasma volume can be defined in a simple function of weight and height of a patient) and the time post dose when coagulation activity has declined to 1 IU/dL (1). Conventionally, these calculations have to be performed outside of the primary PK analysis by multiple additional computations in either WinNonlin or Excel.

Electronic supplementary material The online version of this article (doi:10.1208/s12248-014-9711-7) contains supplementary material, which is available to authorized users.

¹ Seventh Wave Laboratories, Chesterfield, Missouri, USA.

² Biogen Idec Hemophilia, Cambridge, Massachusetts, USA.

³ B2S Consulting, Indianapolis, Indiana, USA.

⁴ Present Address: Vertex Pharmaceuticals Incorporated, 50 Northern Avenue, Boston, Massachusetts 02210, USA.

⁵ Present Address: WIL Research, Ashland, Ohio, USA.

⁶ To whom correspondence should be addressed. (e-mail: emailshenjun@gmail.com)

Another example, in generic drug development, is to calculate the similarity factor (F_2) for two dissolution profiles to evaluate the *in vitro* bioequivalence (2). The calculation of F_2 does not involve any PK analysis, but it is based on defining multiple intermediate entities and assembling them to obtain the final result. Conventionally, this has to be done by setting up a series of formulas in a worksheet. Although these types of calculation are usually not difficult, repeated “copy and paste” data from multiple sources, manually entering formulas, and other additional analyses out of the primary analysis are not only inefficient but error-prone, which adds unnecessary burden on regulatory compliant studies.

We propose a method of taking advantage of DTA and DTAARRAY variables plus some simple programming in an ASCII model to allow automation of these user-defined calculations in WinNonlin and eliminate the need of manual handling of data outside of the primary analysis. In the following sections, we discuss three typical scenarios to illustrate the application of this strategy. In case 1, DTA variables were used to calculate three user-defined parameters in the primary PK model. In case 2, a baseline correction decision tree was programmed into the PK model to account for both the presence of endogenous baseline level as well as the presence of residual drug. In case 3, DTAARRAY variables were used to perform a looping operation to calculate the difference factor (F_1) and the similarity factor (F_2) for *in vitro* bioequivalence evaluation.

METHODS

Case 1: Using DTA Variable to Calculate User-Defined Parameters

DTA(*N*), where *N* is the column number, is a WinNonlin variable that provides access to input data. Usually this variable is used to extract an indicator value to specify to which set of data the observations belong (WinNonlin Users Guide, version 5.2). Due to this property, DTA(*N*) can also be used to extract any numerical value from the input dataset such as weight, height, and dose as long as these values are properly set up in the dataset. A simple assignment statement will suffice. The only problem in using the DTA feature is that DTA(*N*) does not tolerate any missing value or non-numerical value, which will terminate the whole program. In some situations, this can cause a problem. Our solution is to add an additional column in the dataset to indicate if a value is missing or non-numerical (e.g., 0 for numerical values, 1 for missing or non-numerical values) so with an IF-THEN conditional statement, DTA(*N*) can skip reading that value.

The potential application of DTA variable in the calculation of user-defined PK parameters was illustrated with the FVIII activity PK data collected from phase1/2a clinical trial of rFVIIIFc, which was designed for the treatment of hemophilia A (3). The user-defined PK parameters for rFVIIIFc include incremental recovery (*K* value) and *in vivo* recovery (%), IVR), where $K = C_{\max}/\text{dose}$ and $\text{IVR} = 100 \times K \times \text{plasma volume}$, with $\text{plasma volume} = (23.7 \times \text{height} + 9.0 \times \text{body weight} - 1709)/100$. When height, body weight, and dose information can be directly read in through three DTA variables in the START block, the formulas can be set up in the SECONDARY block, an example ASCII code is provided in part 1 of the supplemental material. Another DTA variable can be used to find the observed $C_{\max}$ as follows. When the model goes through the entire dataset, each observation value can be read in by DTA and compared so the maximum value can be easily retained at the end of the primary analysis for further calculation. Model-predicted $C_{\max}$ can be either analytically calculated (model-dependent method) or obtained by comparing all the model-predicted values through the entire time course when the differential equation system proceeds (model-independent method). Therefore, *K* value and IVR can be calculated on both observed and model-predicted values for comparison purpose within the same frame of primary analysis.

The third user-defined PK parameter, the post dose time when model-predicted factor VIII activity has declined to 1 IU/dL (time to 1%, TBLP1), is conventionally obtained by performing an additional simulation after the primary PK modeling. We propose a general model-independent method that is to take advantage of the differential equation system and scan through the entire time course in the process of the primary PK modeling in a similar way to obtaining the model-predicted $C_{\max}$ by setting a simple conditional statement (e.g., IF Conc ≥ 1 IU/dL THEN TBLP1 = *T*, where *T* is the differential equation system time) in the DIFFERENTIAL block. However, since the model prediction stops at the last observation time, it is possible that the model-predicted factor VIII activity has not declined to 1 IU/dL yet when the model prediction stops. Therefore, a checkpoint needs to

be in place to check if the TBLP1 estimate is very close or equal to the last observation time. If so, a value of zero will return to indicate model-predicted factor VIII activity may not decline to 1 IU/dL up to the last observation time. The analyst has to manually inspect the output. To get the model prediction to proceed beyond the last observation time, a row of “dummy” observation record needs to be added at the end of the profile. This dummy record contains a time point that is large enough to allow factor VIII activity to decline to or below 1 IU/dL. Re-running of the model until no TBLP1=0 is present in the output will result in TBLP1 being estimated correctly.

Case 2: Program a Baseline Correction Decision Tree into a PK ASCII Model

Many biological drugs have endogenous counterparts. Furthermore, prior to enrollment on clinical trials, patients are often already receiving therapeutic treatment. Consequently, some level of residual drug from the prior-to-trial treatment may be observed in those patients. If this treatment is bioanalytically indistinguishable from the trial test article, then the PK profile will have three convoluted entities: the test drug, endogenous counterpart, and residual drug. In some cases, if the study design is robust enough (e.g., properly designed PK sampling points over a period that is long enough for baseline to be reliably estimated), the three entities may be uniquely identified through modeling. In other cases, when modeling alone is not able to distinguish the three entities and it is still desired to account for the presence of endogenous counterpart and residual drug before PK analysis, some pre-defined baseline correction methodology may have to be employed in order to adequately describe the decay of the test article.

A clinical development program of rFIXFc (recombinant coagulation factor IX Fc fusion protein) presented such an example (4). The PK profile, plasma concentration of rFIXFc, is determined by the factor IX coagulation activity to which all three entities (e.g., test article, endogenous baseline, and residual drug) contribute. It is not possible to uniquely distinguish the residual drug (BeneFix) that is still expected to decay or the baseline (endogenous factor IX) from the test article by simply modeling the PK profile. It is, however, strongly desired to account for the presence of the residual drug and the baseline in the PK profile which otherwise will produce biased estimate on the parameter of primary interest, elimination half life. One solution is to define a decision tree using scientific judgment and all available information to manually correct baseline and residual drug as shown in Fig. 1. We propose to program this decision tree into the PK model. The decision tree in nature is a set of conditions leading to different actions, which can be coded in a hierarchy of IF-THEN statements in the START block when differential equations are used to define the PK model.

Case 3: Using DTAARRAY Variable to Calculate Difference Factor (F1) and Similarity Factor (F2)

DTAARRAY(*i,j*) variable is not a well-documented feature in the WinNonlin Users Guide. Similar to the

Fig. 1. A decision tree to account for the baseline and residual drug level. Note: The decision tree presented here is only for demonstration purpose and may be different from the one that was actually used in the real study (4)

DTA(N) variable, DTAARRAY(i,j) provides access to input data. The difference is DTA(N) can only access the “current” row the program is reading while DTAARRAY(i,j) can access the whole dataset by specifying a row (i) and column (j) number. Therefore, DTAARRAY(i,j) can be called as an array as its name has suggested. The dataset-wise access to input data provides an opportunity to perform some wrap-up calculation within an ASCII model when modeling is not even the interest.

In generic drug development, it is often necessary to provide experimental evidence of *in vitro* bioequivalence for a test drug with respect to a reference drug. For example, to demonstrate that a generic formulation of calcium acetate (test) is equivalent to the standard care (reference), an *in vitro* phosphate binding study needs to be performed (5). The test or reference drug is incubated with phosphate buffer prepared in various concentrations. A percentage of phosphate binding can be measured for test and reference drug at each concentration of phosphate buffer. Then, a difference factor ($F1$) and a similarity factor ($F2$) can be calculated as follows:

$$F1 = \left\{ \left[\frac{\sum_{i=1}^n |R_i - T_i|}{\sum_{i=1}^n R_i} \right] \right\} \times 100 \quad (1)$$

$$F2 = 50 \times \text{LOG} \left\{ \left[1 + (1/n) \times \sum_{i=1}^n (R_i - T_i)^2 \right]^{-0.5} \times 100 \right\} \quad (2)$$

Where R_i and T_i represent the percentage of phosphate binding for reference and test drug, respectively, at each of the n phosphate buffer concentrations. Conventionally, given

the primary dataset with R_i , T_i , and phosphate buffer concentration, it requires manual entering multiple formulas for each dataset. We propose to code the formulas in an ASCII model by using the DTAARRAY(i,j) variable in a DO-loop to automate the whole calculation. Since modeling the data is not the purpose, any model relationship (e.g., $y = a \times x + b$) can be specified in such an ASCII model to complete the model specification.

RESULTS

Case 1: Using DTA Variable to Calculate User-Defined Parameters

The sample code for implementation of DTA variables to perform user-defined calculation was provided in Part 1 of the supplemental material. In this phase I/IIa trial for rFVIIIc (3), 16 patients with a crossover design produced 32 PK profiles. With the user-defined model code (part 1), all parameters of interest (e.g., HL, KV, IVR, TBLP1 etc.) from 32 profiles were calculated in one run, which saved ~60–70% of analysis time compared with manually calculating those secondary parameters. Moreover, the results of user-defined PK parameters generated from user-defined code were comparable to that obtained from manual calculation, suggesting the user-defined code was valid. As described in the “Methods” section, if the TBLP1 results returned a value of zero, it was a signal that the original sampling period was not long enough for factor VIII activity to drop to 1 IU/dL. In several PK profiles, we observed a value of zero for TBLP1. For these profiles, a dummy observation record was added with a large time point (e.g., 168 h) and the analysis was simply “refreshed” to obtain the correct values for TBLP1 (the whole analysis was saved in a workspace). The inclusion of dummy observation record was indicated by adding an

additional column in the dataset (e.g., INDX), if the value in the column of INDX is 1, the data record in the same row is the dummy observation record, otherwise, the true observation data record. Alternatively, the dummy observation records with large time point can be inserted into the dataset for each PK profile at the very beginning to avoid the “check and refresh” procedure. Figure 2 demonstrated an example where a dummy observation record (at 168 h) was inserted to help estimate TBLP1. Furthermore, KV and IVR were calculated on both model-predicted (KV_M, IVR_M) and the observed (KV_OB, IVR_OB) $C_{\max}$ that was obtained by using a DTA variable and formulas presented in the SECONDARY block.

Case 2: Program a Decision Tree into a PK Model

The sample code for programming a decision tree in an ASCII model was provided in Part 2 of the supplemental material. In this phase I/IIa trial for rFIXFc (4), endogenous baseline and residual drug (both are convoluted with test drug in the measurement of factor IX activity) cannot be uniquely determined by modeling the PK data. Utilizing the information from screening and pre-dose samples, the decision tree dictated a baseline level (BL) and a first-order elimination rate (K_R) for the residual drug so that the amount of pre-existing residual drug (Resid) and other PK parameters can be estimated through a two-compartment model. When a decay of residual drug (BeneFix) needed to be accounted for, a reported half life (FixHL, 18 h) for BeneFix was used to calculate a first-order elimination rate constant, FixK_R (see part 2). The dataset followed a similar format to what is shown in case 1 (see the sample dataset in part 1) with an additional column (Col_SCN) to define screening sample. To validate the code for the decision tree, the BL and K_R values were reported as secondary

parameters and verified against the values they were meant to be.

Case 3: Using DTAARRAY Variable to Calculate Difference Factor ($F1$) and Similarity Factor ($F2$)

The sample code for implementation of DTAARRAY variable to calculate difference factor ($F1$) and similarity factor ($F2$) was provided in Part 3 of the supplemental material. The number of data points (e.g., time points in the sample code) was provided through CON(1) in the dosing regimen tab. Since modeling was not the primary interest, the specified model (a linear model in the sample code) only served to complete the model specification. The calculation of intermediate variables (e.g., SUM1, SUM2, and SUMR) was performed in the TEMPORARY block by using several DTAARRAY variables within a Do-NEXT loop. The final calculation of $F1$ and $F2$ was performed in SECONDARY block. $F1$ and $F2$ were then reported as secondary parameters. The results obtained from the ASCII code were verified against the results by manual calculation.

DISCUSSION

Since WinNonlin is a validated software system, it is used widely for supporting pharmacokinetic assessments in regulatory compliant studies. However, some of the features in ASCII models are not as widely realized by the PK/PD community. In this article, we presented three typical cases to advocate taking advantage of these under-used features together with simple programming techniques to streamline data analysis.

DTA and DTAARRAY variables can be called in an ASCII model to perform user-defined calculation, which can either be part of the primary PK modeling as shown in Case 1 or be the sole purpose of the analysis as exemplified in Case

Fig. 2. TBLP1 calculation. The time of last observation (circle) around 72 h post dose is not long enough for the predicted drug concentration (solid line) to drop to or below the threshold (dotted line). When a “dummy” observation record was inserted with a time point at 168 h, the predicted drug concentration extends to 168 h post dose and drops below threshold so that the TBLP1 parameter can be correctly calculated

3. It is then possible to define the calculation method separately from the data. Once the ASCII model is validated, it can be repeatedly called to perform the same calculation for new data. This is particularly useful for many clinical development programs that require such calculation for each patient to support decision making (e.g., dose adjustment). In addition to the obvious benefit of time saving, it relieves the concern from quality assurance about the accuracy of those formulas every time the same type of calculation needs to be performed. The efficiency of the method will only become more significant when more datasets are processed.

We presented Case 2 with a special interest to demonstrate handling baseline subtraction. As discussed in the “Methods” section, the endogenous baseline convoluted with pre-existing residual drug may not always be identified through modeling. Thus, when baseline correction is still desired, it is necessary to define some *a priori* criteria to do so such as a decision tree. By programming the decision tree into the model, the manual baseline correction does not have to be “manual” anymore. In our experience, it certainly improved the consistency of the process of data analysis and reduced the subjective tendency of the analyst. This method can be easily adapted to different pre-defined criteria for pre-processing data.

To conclude, we found DTA and DTAARRAY variables were very useful features to automate user-defined calculations

through ASCII models. We encourage members of the pharmacokinetic community to consider applying these features more broadly in their routine data analyses to support drug development.

REFERENCES

1. Kessler CM, Gill JC, White GC, Shapiro A, Arkin S, Roth DA, *et al.* B-domain deleted recombinant factor VIII preparations are bioequivalent to a monoclonal antibody purified plasma-derived factor VIII concentrate: a randomized, three-way crossover study. *Haemophilia*. 2005;11:84–91.
2. Shah VP, Tsong Y, Sathe P, Liu JP. In vitro dissolution profile comparison—statistics and analysis of the similarity factor, *F2*. *Pharm Res*. 1998;15:889–96.
3. Powell JS, Josephson NC, Quon D, Ragni MV, Cheng G, Li E, Jiang H, Li L, Dumont JA, Goyal J, Zhang X, Sommer J, McCue J, Barbetti M, Luk A, Pierce GF. Safety and prolonged activity of recombinant factor VIII Fc fusion protein in hemophilia A patients. *Blood*. 2012.
4. Shapiro AD, Ragni MV, Valentino LA, Key NS, Josephson NC, Powell JS, *et al.* Recombinant factor IX-Fc fusion protein (rFIXFc) demonstrates safety and prolonged activity in a phase 1/2a study in hemophilia B patients. *Blood*. 2012;119:666–72.
5. FDA. Draft guidance on calcium acetate. 2011.