

Metadata-Driven Data Analysis

Peter Schaefer, CEO

Validated Cloud Applications, Inc.

PhUSE Single Day Event, June 2nd, 2016

Gaithersburg, MD



Just Two Topics



- Impact of using metadata on user interfaces
- What does that mean for analysis tools?

The Prototype of “Complex” User Interfaces



Why was that?



Along came some Metadata ...



Los Angeles Times

The Time Machine : Using Listed Codes VCR Plus Will Simplify Recording Off TV

November 25, 1990 | DANIEL CERONE | Times Staff Writer

For those confused VCR owners who have pitched a permanent tent in front of their television sets so as not to miss their favorite TV shows, it's time to break camp.

This week marks the local introduction of a user-friendly gadget that its developers claim can program the most complicated home VCR with the ease of using a touch-tone telephone.

It's called VCR Plus--a remote-control device that automatically records programs by entering a code number printed in TV listings.

Starting this week in Southern California, the code numbers (PlusCodes) will be printed exclusively in the Los Angeles Times Sunday TV Times magazine and daily Calendar TV listings. The PlusCodes in The Times will appear beside broadcast and cable TV programs in five categories: prime-time, movies, sports, soap operas and children's programming.

How the use of Metadata changed the User Interface ...



Traditional

- Applies forward thinking for problem solving
- Buy more flexibility and pay with higher complexity
- High risk of user errors
- Rather low risk of data errors

Metadata-driven

- Rethinking the problem from the end
- Simpler user interface but with limitations
- Fewer mistakes by user
- Higher potential for errors from data
- Additional effort to create metadata

Back to Life Sciences:

Let's Create Concentration-Time Plots

- Think about the user input for this kind of plot

Select Profile Dataset:
Choose File: ADNCA-like...osing.csv

Variable	Column
X_Axis	ARELRTM
X_Unit	RELRTMU
Y_Axis	AVAL
Y_Unit	AVALU

Profile_Keys

☐ STUDYID	☐ TRTRSDTM	☐ SPEC
☐ USUBJID	☐ NRELRTM	☐ ROUTE
☐ SITEID	☐ ARELRTM	☐ COUNTRY
☐ TRTP	☐ RELRTM	☐ TRTDESC
☐ TRTA	☐ AVAL	☐ PKFL
☐ SEX	☐ AVALU	☐ PKRFL
☐ DOSE	☐ PARAM	☐ PKEXMFL
☐ DOSEU	☐ PARAMCD	

Group_By

☐ STUDYID	☐ TRTRSDTM	☐ SPEC
☐ USUBJID	☐ NRELRTM	☐ ROUTE
☐ SITEID	☐ ARELRTM	☐ COUNTRY
☐ TRTP	☐ RELRTMU	☐ TRTDESC
☐ TRTA	☐ AVAL	☐ PKFL
☐ SEX	☐ AVALU	☐ PKRFL
☐ DOSE	☐ PARAM	☐ PKEXMFL
☐ DOSEU	☐ PARAMCD	

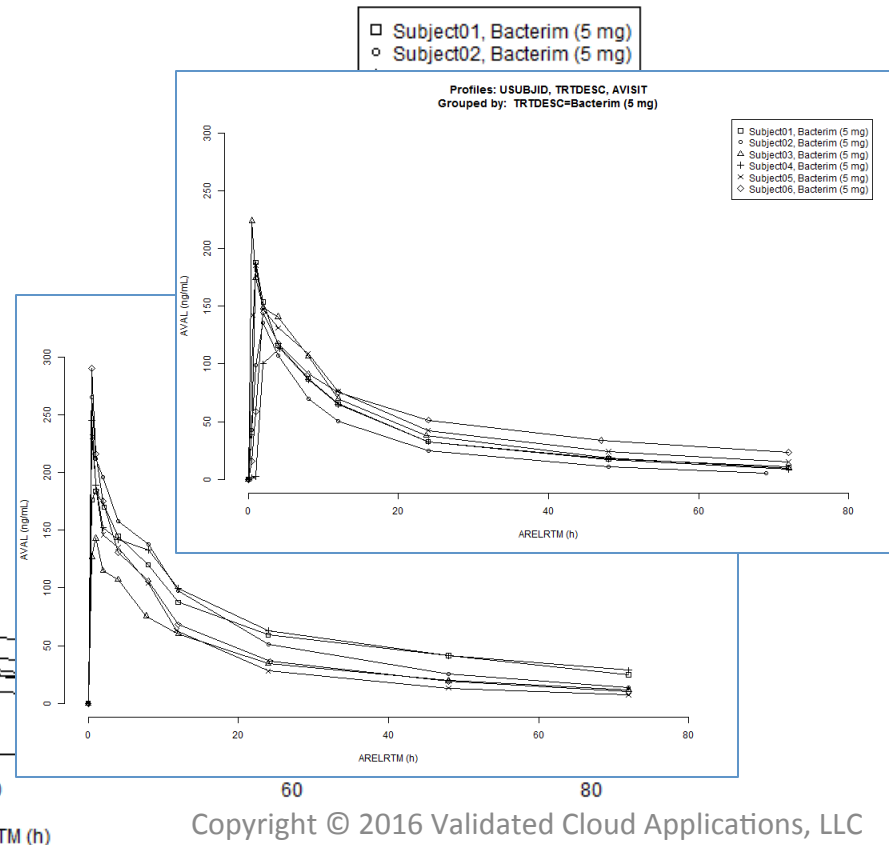
Filter Include

Add Filter

Select File Format for Plots:
☒ Portable Network Graphics

Profiles: USUBJID, TRTDESC, AVISIT
Grouped by: TRTDESC=Bacterim (5 mg)

- Variables for x- and y-axes
- Variables to uniquely identify a profile
- Variables to group profiles into one plot
- Additional settings (file format, plot size, etc.)



And what if you want even more?

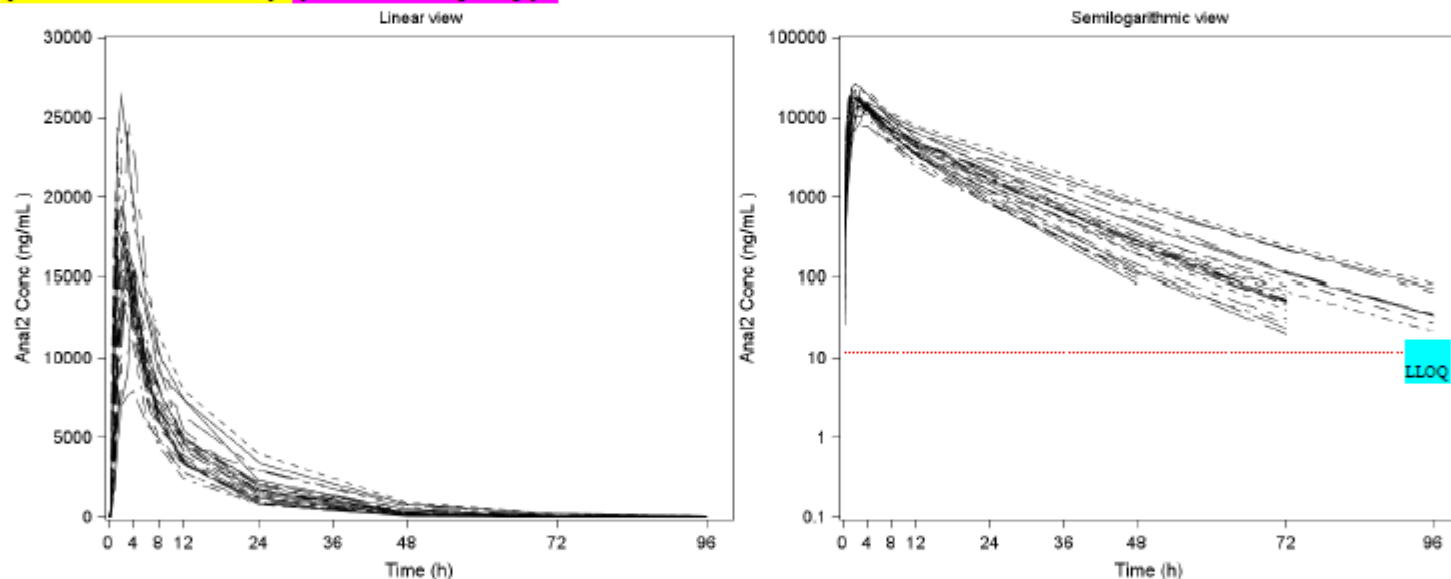
PROTOCOL/PRODUCT INFO

(page x of x)

Figure 16.2.5-x.x Overlaying individual [dose-normalized] concentration-time profiles
by compound, matrix, analyte and [actual/randomized] [treatments/group]
Analysis Set: PK analysis set

Compound: XXX, Matrix: YYY, Analyte: ZZZ

[Actual/Randomized] [treatment/group]: AAAAAA



Values < LLOQ (xxx uom) were considered as [zero;missing;LLOQ; LLOQ/2;...].

PATH DATA/PROGRAM/OUTPUT

PRODUCTION STATUS/RUN DDMMYYYY: HHMM

Figure 7-5. Shell for

From the PhUSE White Paper
“Analyses and Displays Associated to Non- Compartmental Pharmacokinetics ”

Using Metadata, you'll get more with less

The screenshot shows a software interface for PK analysis. At the top, there is a dropdown menu labeled 'Selected Solution: Figure_7_5_7_6' and an 'Execute' button. Below this is a section with 'Solution Name: Figure_7_f5_7_6.R' and 'Description: PK Concentration Profiles by Treatment and by Subject.' The next section has 'Local Source: ☒' and 'Select ADaM input dataset:' with a 'Choose File' button and the text 'ADNCA-like...cterim.csv'. Below this is a table with 'Filter' and 'Include' columns, and an 'Add Filter' button. The bottom section is titled 'Select the type of figures to create' and has two checked options: 'Plots with all subjects separated by treatment and analyte' and 'Plots with all treatments separated by subject and analyte'. There is another 'Execute' button at the bottom. Red arrows point from the text 'Standardized datasets allow to replace user input with defined and well-understood metadata' to the 'ADaM input dataset' field, and from 'User only selects desired output from predefined analyses' to the figure type options. A red box at the bottom left contains the text 'Predefined analyses based on PhUSE White Paper on PK'.

Selected Solution: *Figure_7_5_7_6* ▾

Execute

Solution Name: *Figure_7_f5_7_6.R*
Description: *PK Concentration Profiles by Treatment and by Subject.*

Local Source: ☒

Select ADaM input dataset:

Choose File ADNCA-like...cterim.csv

Filter	Include
Add Filter	

Select the type of figures to create

☒ Plots with all subjects separated by treatment and analyte

☒ Plots with all treatments separated by subject and analyte

Execute

Predefined analyses based on PhUSE White Paper on PK

Cloud Applications, Inc.

- Standardized datasets allow to replace user input with defined and well-understood metadata
- User only selects desired output from predefined analyses

So, what can be done?



- Using metadata for simplifying the UI can be applied in many situations.
- For example,
 - Non-Compartmental PK Analysis (NCA)
 - Standard method applied to concentration-time profiles to understand what happens in the body after dose administration
 - Typical user input includes
 - Mapping of variables (time, concentration, dosing, ...)
 - Some profile and sampling characteristics
 - Some analysis parameters

Bringing it all together ...

Metadata-driven UI for NCA

Replace these ...

	None	Sort	Carry	Time	Concentration
Subject	☐	☐	☐	☐	☐
Study_ID	☐	☐	☐	☐	☐
Age	☐	☐	☐	☐	☐
BMI	☐	☐	☐	☐	☐
BSA	☐	☐	☐	☐	☐
Smoke	☐	☐	☐	☐	☐
Gender	☐	☐	☐	☐	☐
Height	☐	☐	☐	☐	☐

After loading the dataset:

- Mapping of variables (sort/key variables, time, concentration, dosing, ...)
- Profile and sampling characteristics
- Analysis / model parameters

Select
Adam-NCA
Input Dataset

Local Source: ☒

Select an ADaM-NCA compliant dataset as input

Choose File | ADNCA-like...cterim.csv

Select the dose type: ☐ ADNCA-like_Bacterim.csv

☒ Non-steady-state, typically after single dose administration

☐ Steady-state, typically after multiple dose administration

CoreNCA

Solution Name: NCA_Core.R

Description: Non-Compartmental Analysis

AUC Calculation Method

☒ Linear

☐ Logarithmic

☐ Linear-logarithmic

Some NCA
settings

Summary

- 
-  Tools should be easier to use when user interfaces are using metadata
 -  But, tools will be less flexible – think “canned analyses”
 -  There should be little to no overhead for providing metadata to drive tools if data standards are used
 -  However, it seems that current data standards (including CDISC) do not incorporate enough metadata
 - In other words, a lot of meta-information is implicit and typically not tool accessible
 - Specifically well suited for some well established analyses
 - Examples are NCA, dataset conversion, initial dataset exploration