PhUSE 2006

Paper TS10

A Integrating Specialized Graphics and Analysis into Drug Development Production Reporting Systems

Joel Hoffman, Insightful Corporation, New York, US
Michael O'Connell, Insightful Corporation, Research Triangle, US

ABSTRACT

The flow of data for a preclinical or clinical study has many steps from specification to acquisition to storage to conformance to computation to reporting. For TK/PK studies this is especially complex because of the study design, sample collection at multiple time points within a day/visit, multiple sources of data, and multiple computational and graphics systems. In many current situations, this leads to inefficient use of resources as high level colleagues conduct the manual work and QC for assuring the quality of these tasks.

In contrast, a controlled and automated flow of data across the information lifecycle can be flexible and adaptive and produce significant saving of resources and reduce the time from database lock to study report.

This paper describes current procedures, processes and systems used by biopharmaceutical companies to perform these tasks. An alternative model will be proposed which reflects some current work in standards organizations. Second, an interface that can be used for implementing this alternative using S-Plus will be described. Third, a method for integrating S-Plus software into a current SAS based production reporting environment will be described.

INTRODUCTION

The flow of data for a preclinical or clinical study has many steps from specification to acquisition to storage to conformance to computation to reporting. For TK/PK studies this is especially complex because of the study design, sample collection at multiple time points within a day/visit, multiple sources of data, and multiple computational and graphics systems. In many current situations, this leads to inefficient use of resources as high level colleagues conduct the manual work and QC for assuring the quality of these tasks.

In contrast, a controlled and automated flow of data across the information lifecycle can be flexible and adaptive and can produce significant saving of resources and reduce the time from database lock to study report.

This paper will describe current procedures, processes and systems used by biopharmaceutical companies to perform these tasks. An alternative model will be proposed which reflects some current work in the standards organizations. Second, an interface that can be used for implementing this alternative using S-Plus will be described. Third, a method for integrating S-Plus software into a current SAS based production reporting environment will be described.

Pipeline pressures forced by new discovery technologies, heightened competition and regulatory demands for increasingly detailed studies mean the pace and number of clinical trials continue to accelerate. At the same time, budgets are shrinking, so pharmaceutical companies need to assess the value of each study, particularly in the early stages. As Kenneth Kaitin, director of the Tufts Center for the Study of Drug Development, has stated, companies need to learn to 'Kill early. Kill often.' Pharmacokinetists often play a significant role in such go/no-go decisions. By advising management if a drug is poorly metabolized, minimally efficacious or showing safety concerns, and recommending an early kill, these scientists can save companies the millions of dollars it would cost to move a dubious product into the next stage of development.

Pharmacokinetic studies are detailed and complex. The sheer scale and complexity of the data gathered, as well as the specific requirements of regulatory bodies with regard to how data are presented, adds many complications. This is compounded by the need to combine these data with other related clinical information, a process which is made all the more challenging with separate operational and informational silos. Too often, PK scientists are forced to devote

much of their time to data management and specification of routine displays rather than actual science. This results in poor productivity, delays in go/no-go decision, and affects how rapidly a company can bring its product to market.

Opportunities to simplify the process can be enabled by technology, but this presents challenges with regard to compliance, compatibility with in place production systems, and smoothly transitioning new technologies into the production lifecycle. Efficient methods will require companies to reconsider how each step is integrated into the decision making and reporting process.

FUNCTIONAL PROCESSES AND SYSTEM IN TK/PK AND PD ANALYSES

An integrated approach to the analysis and reporting of TK/PK and PD studies can be seen (Figure 1) as the piecing together of several data systems such as pre-clinical data stores, clinical database management systems, and access to assay results (typically in the form of Excel spreadsheets.

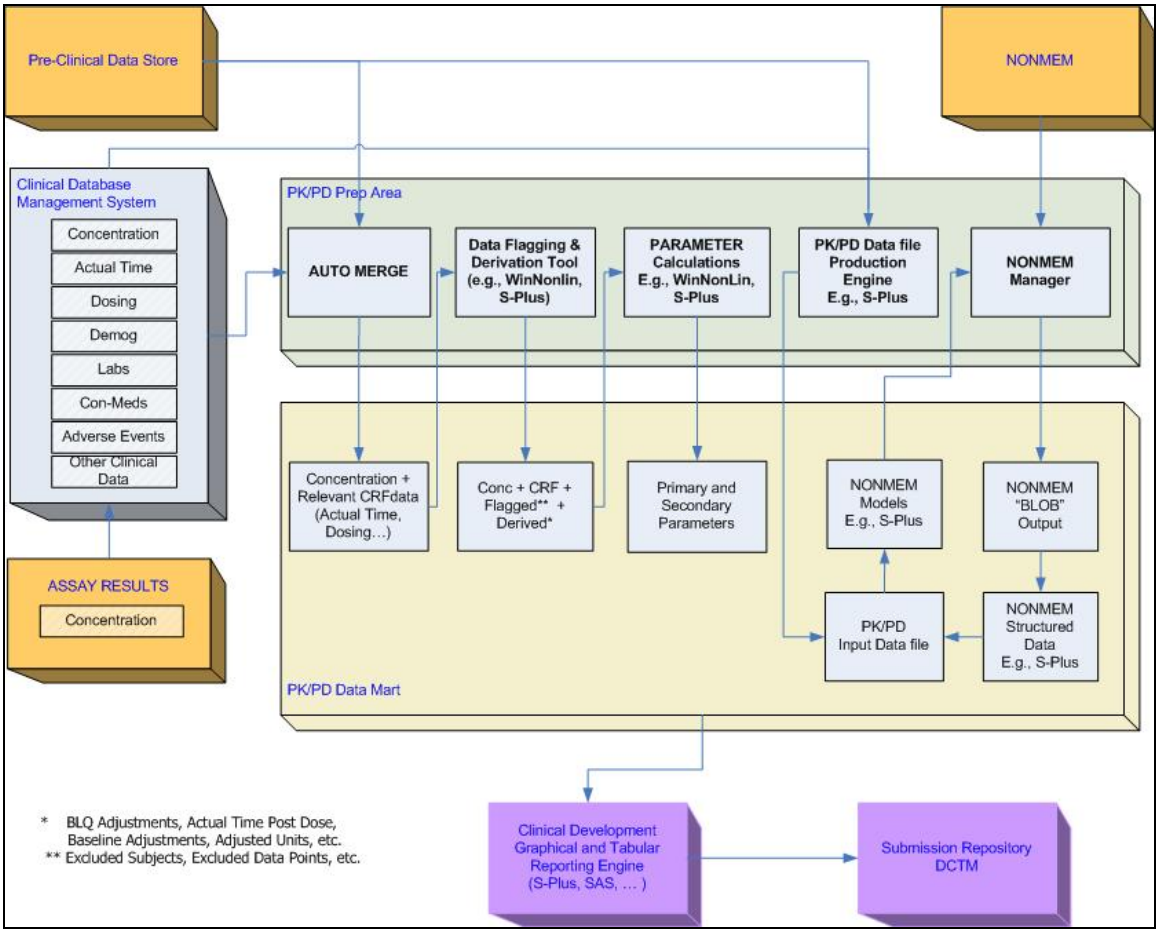

Figure 1: Data and Analysis Systems for TK/PK and PD Studies

These data sources are typically knitted to analysis and reporting systems with a set of manual steps.

PROCESSES FOR ACQUISITION, ACCESS, ANALYSIS AND REPORTING

For these types of studies this is especially complex because of the study design, sample collection at multiple time points within a day/visit, multiple sources of data, and multiple computational and graphics systems.

When scientists and pharmacologists perform analyses of toxicology and pharmacokinetic data, they need to bring data together from multiple sources, conduct analyses with a variety of specialized tools, and then assemble their results into reports to share with colleagues and regulatory agencies. In many current situations, scientists, pharmacologists and statisticians use manual processes for each of these steps as well as QC. As shown in Figure 2, there are multiple hand-offs between functional groups. Once a study is “finalized,” an Analytic Group may

coordinate the efforts of external CROs and Analytic Labs for the acquisition and assay of samples. Pre-Clinical PK groups may also be called up to perform assays, especially for First in Human (FIH) studies where the company may be unwilling to invest in the transfer of the assay technology until it has an early indication of a compounds viability as a medicine.

Once the assay data is available and approved by the PK scientist, there is the task of merging it with other clinical data collected from the Case Report Forms such as actual date/time the post-dose sample was taken, any adverse events which might impact the results (e.g., nausea or vomiting), urine volume, and demographic information such as patient weight. There are then multiple steps for the PK or PD analysis of the data that are of compliance risk (see red boxes in Figure 2)

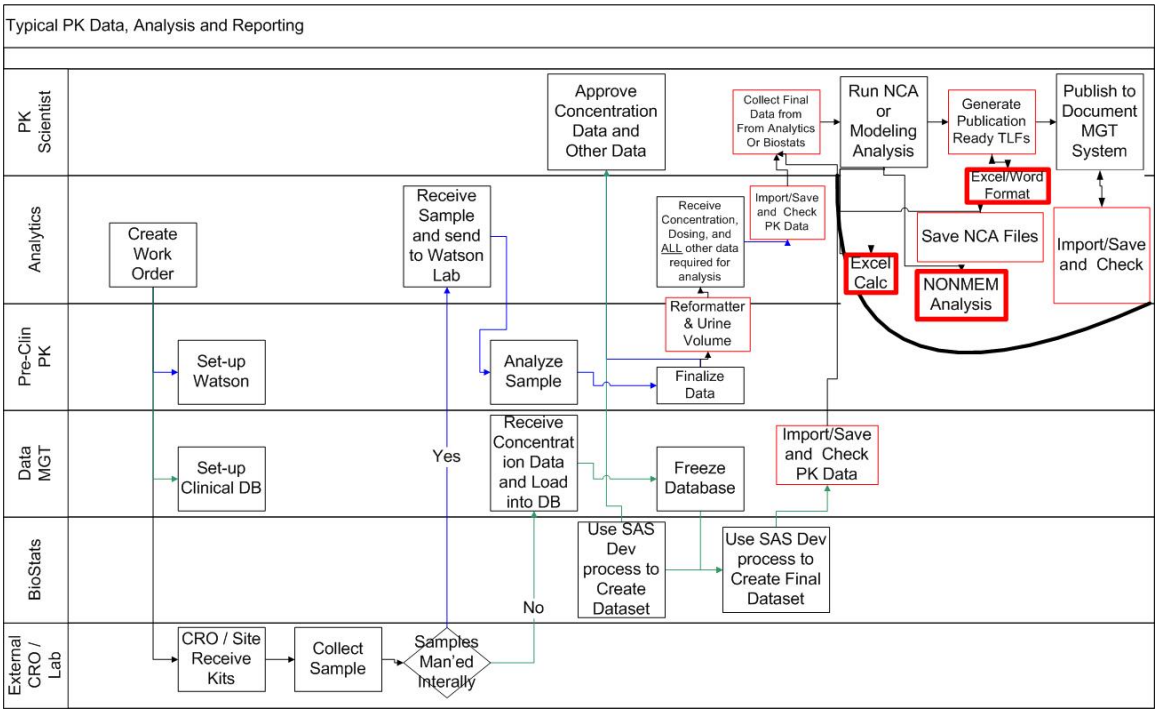

Figure 2: Roles and flow of concentration data

Closer inspection of the analysis process (Figure 3) reveals multiple manual steps through the noncompartmental analysis (NCA) process. In these steps, senior scientists and analyst manually enter and check formulas and transformations into Excel and WinNonlin in order to complete the NCA. And, of course, each one of these manual processes is associated with the manual checks required to ensure quality and lower compliance risk.

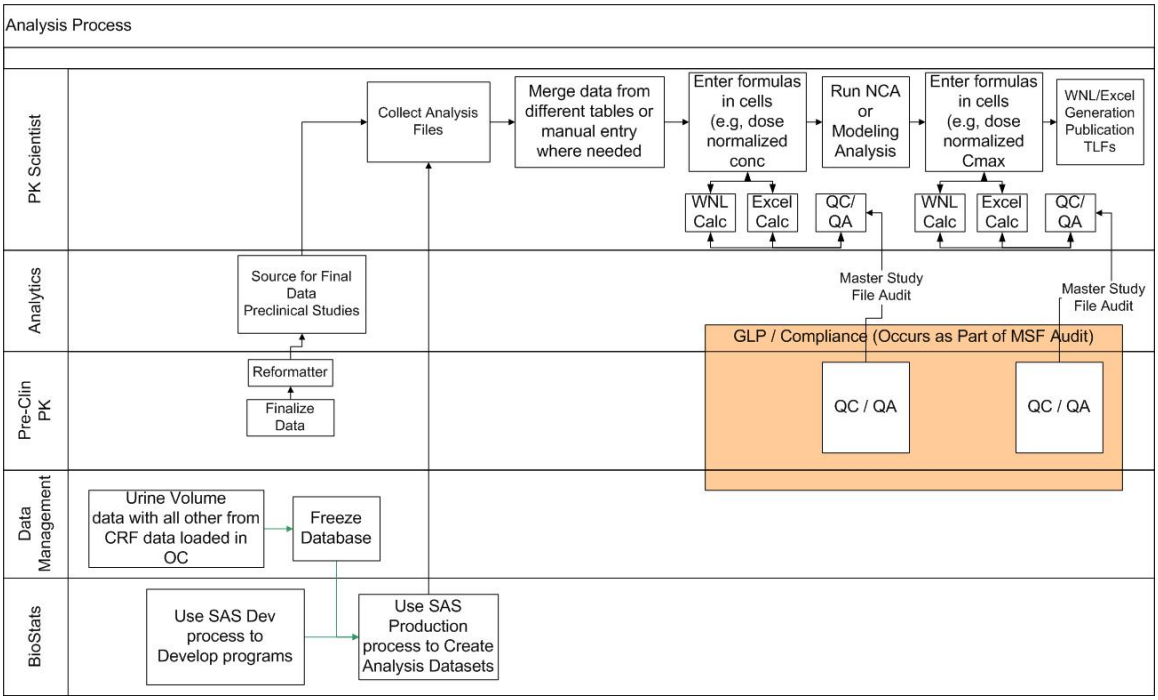

Figure 3: TK/PK Detailed Noncompartmental Analysis Process

SYSTEMS FOR ACQUISITION, ACCESS, ANALYSIS AND REPORTING

A closer look at the systems (Figure 4) that are knitted together in this way also highlights compliance risks where the manual processes are used. Of particular concern can be the use of desktop tools like MS-Excel and MS-Word where it is very difficult to systematically ensure quality. Even tools like WinNonlin and S-Plus when used as desktop tools will not, in and of themselves, provide adequate protection from regulatory risk.

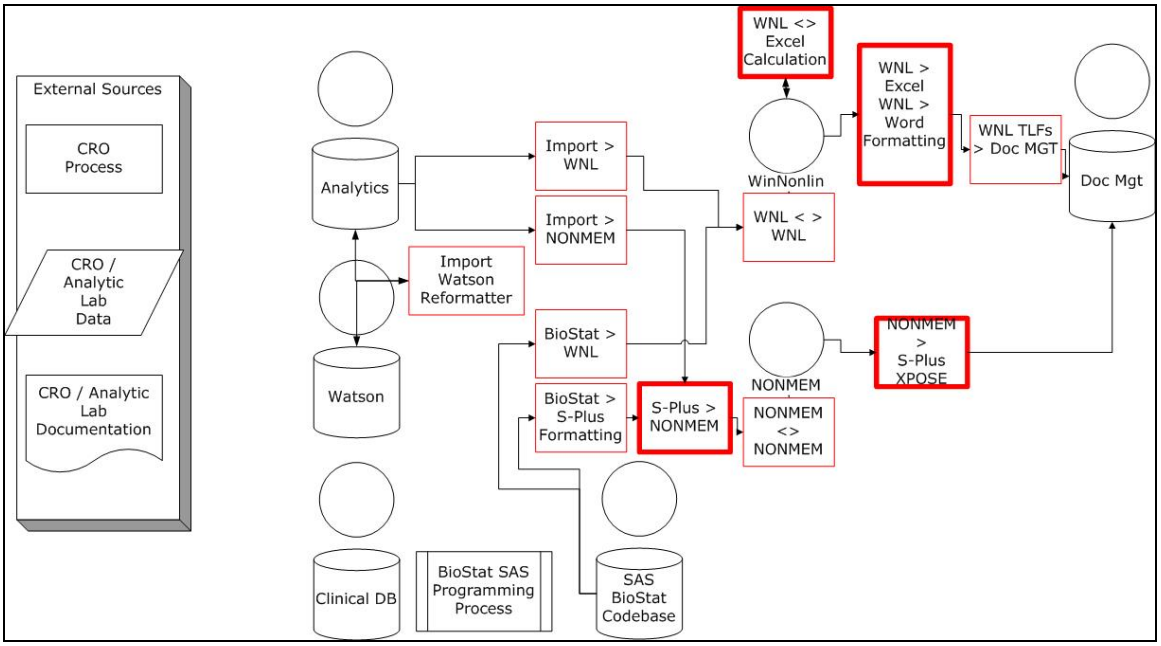

Figure 4: Systems used for PK/PD and potential compliance risk

REPORTING PROCESS ALTERNATIVES

In contrast, a controlled and automated flow of information across the information lifecycle (Figure 5) can produce significant saving of resources and reduce the time from database lock to study report.

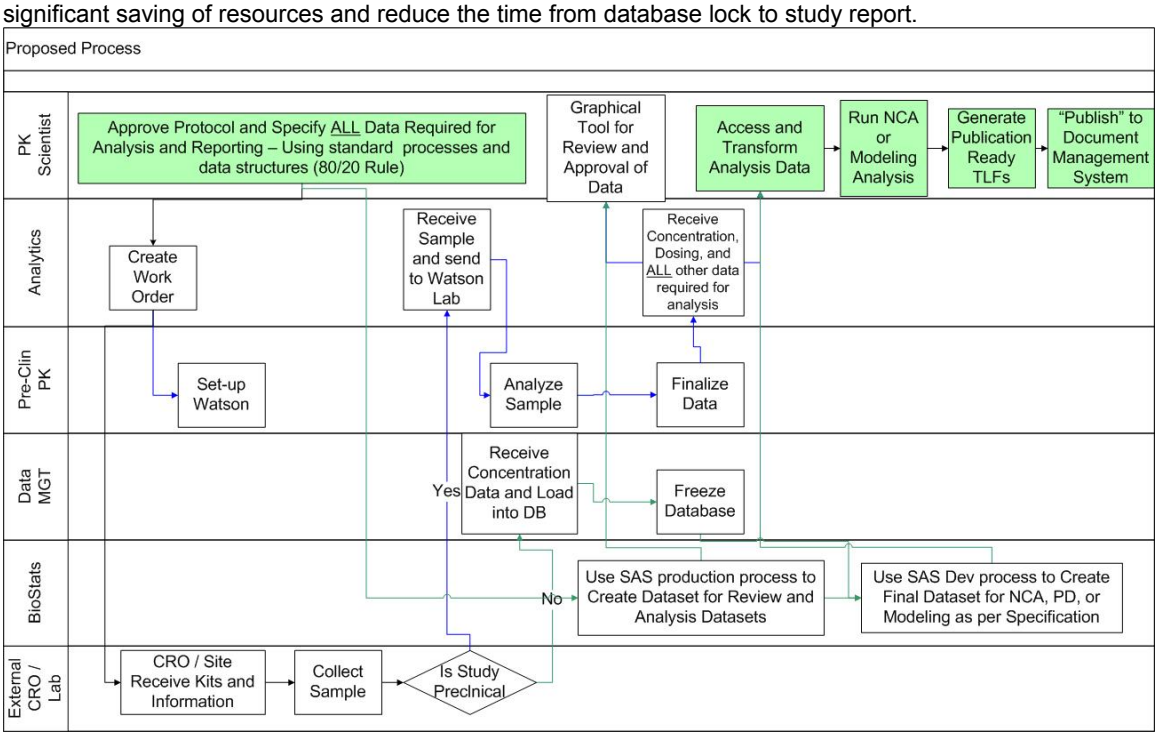

Figure 5: Alternative roles and flow of concentration data

Figure 5, when compared to Figure 2, illustrates the significant reduction in manual hand-offs and steps when systems and processes are tightly integrated. These include: (1) the integration of review and NCA analysis tools with the specification of these analyses in the Protocol and Statistical Analysis Plan (SAP), (2) integration with the production data environment, and (3) the integration of standard and non-standard report into the production reporting environment.

INTEGRATION OF ANALYSIS AND FIGURE SPECIFICATION WITH CODE PRODUCTION

Many steps are saved when the specification of review and submission figures are integrated with the production cycle. PK Scientists and Analysts no longer need to repetitively format and program code and cell formulas and macros to produce figures required output. In the following section, we will describe a system that enables the generation of production code from such a specification tool.

INTEGRATION OF REVIEW TOOLS

When available over the production data environment, review tools will quickly show discrepancies between the data that are collected compared with what is expected. For example, if PK scientists are expecting 2400 sample – data from 20 patients taken once an hour over 24 hours for five days – through visual analytics they can quickly ascertain if data are missing, and whether what is missing are data from one patient an hour's observation from all patients. In the example shown in Figure 6, an interactive data review tool can be used to highlight and identify a patient with extreme values across a range of laboratory test.

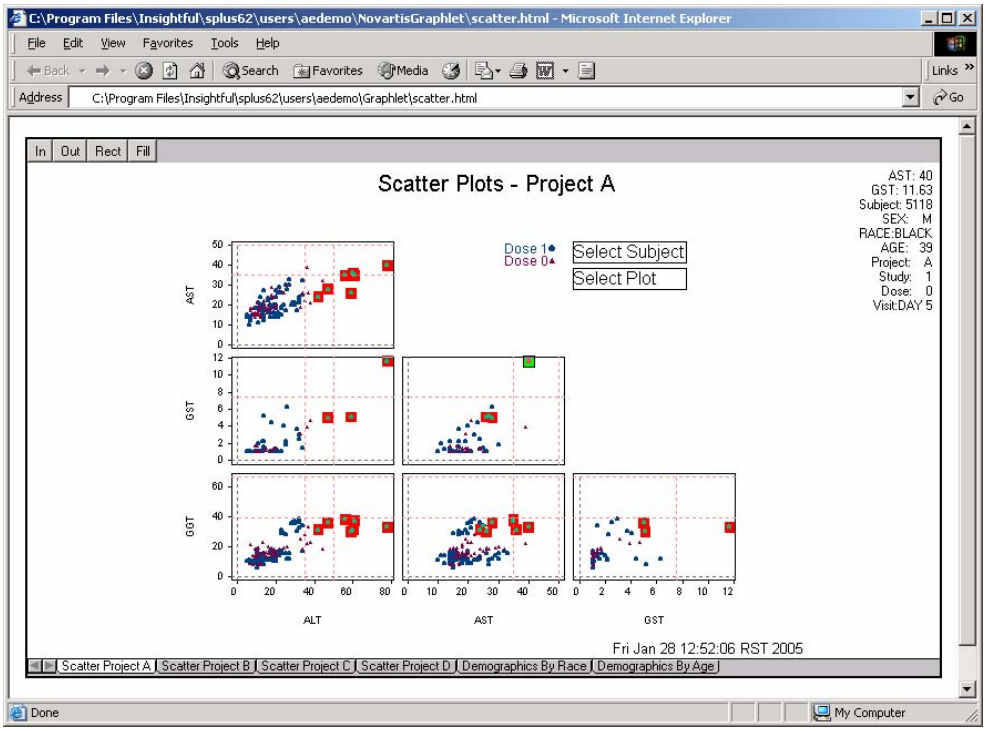

Figure 6: Interactive data review tool

INTEGRATION OF REPORTING AND COMPUTATIONAL TOOLS

For the reasons noted above, SAS is often not used for complex analyses and fine grained formatting of figures required to meet company standards for regulatory submissions. Rather specialized tools like S-Plus, WinNonlin, or NONMEM are used. These tools have two key roles: the ability to perform complex analyses and the creation of complex tables and graphs.

Tightly integrating these systems into the production data and reporting environment can have two significant impacts. First, it can drastically reduce the amount of manual work required for access, merging, and reformatting data. Second, and perhaps of greater impact, is the virtual elimination of manual QC steps.

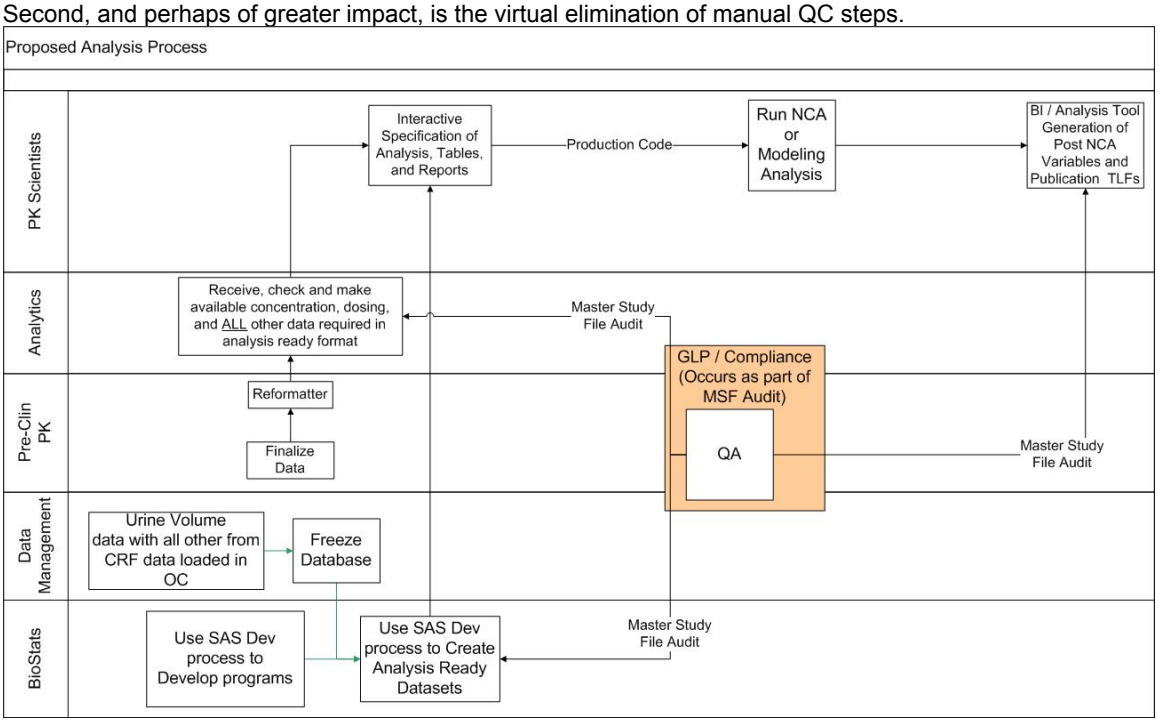

Figure 7: Alternative TK/PK Detailed Noncompartmental Analysis Process

When it comes to creating tables, graphs and listings, the regulatory demands as well as the company's particular requirements, such as font, point size, table length, are pervasive. Because the data flow goes outside of the standard reporting environment, companies will either create the tables and graphs manually or develop parallel systems. If done manually, there might be 100 tables and graphs for a single study and a manual process would require someone to type in a title for 100 different tables, resetting the font size and type for each listing. If a problem then arose, every listing would have to be rerun and reformatted. Automation means users can integrate the tables and graphs with other kinds of analysis and reporting elsewhere in the clinical process, thereby ensuring all relevant data and updates are appropriately reported (Figure 7). In addition, using the same code base to produce interactive data review and for submission graphics means that:

- Scientists will be understanding the data and interpreting it just as a reviewer would. Thus, the knowledge gained though out the review process can be used when writing the PK report. Compare Figure 8 below with Figure 6 above.

- The project team can be sharing looks at the data (in protected ways that do not impact blindedness). Thus, there are more people looking at the data as it becomes available in order to better interpret and understand it.

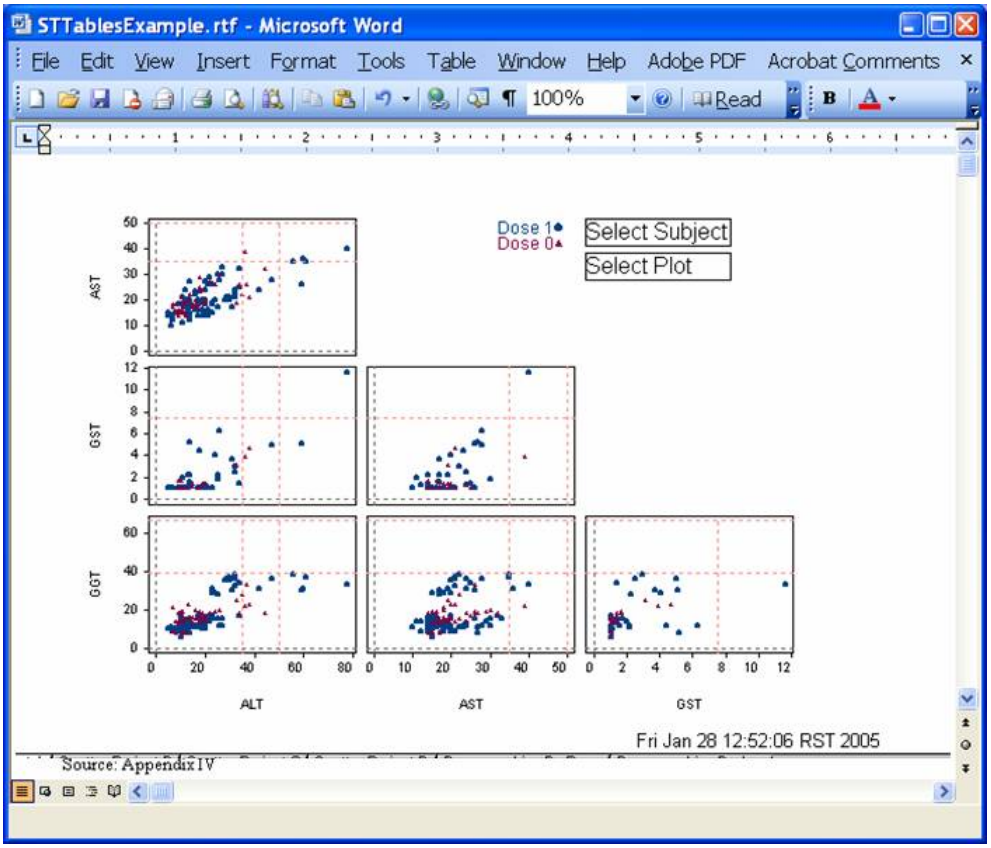

Figure 8: Production figure generated from S-Plus graphics specification

Integrating S-Plus as a statistical and reporting tool can afford significant benefits in both enabling interactive access to TK/PK data as well as providing production quality graphics that are tightly integrated with an SAS® based reporting system. The statistical graphics are based on S-PLUS Graphlets™ which enable drill-down, brushing, and display of meta-data and are well-suited to exploratory analysis and clinical data review. The reports are produced using the S-PLUS Clinical Pack for SAS Users, which enables the incorporation of S-PLUS Graphlets and specialized S-Plus analytics into SAS ODS output.

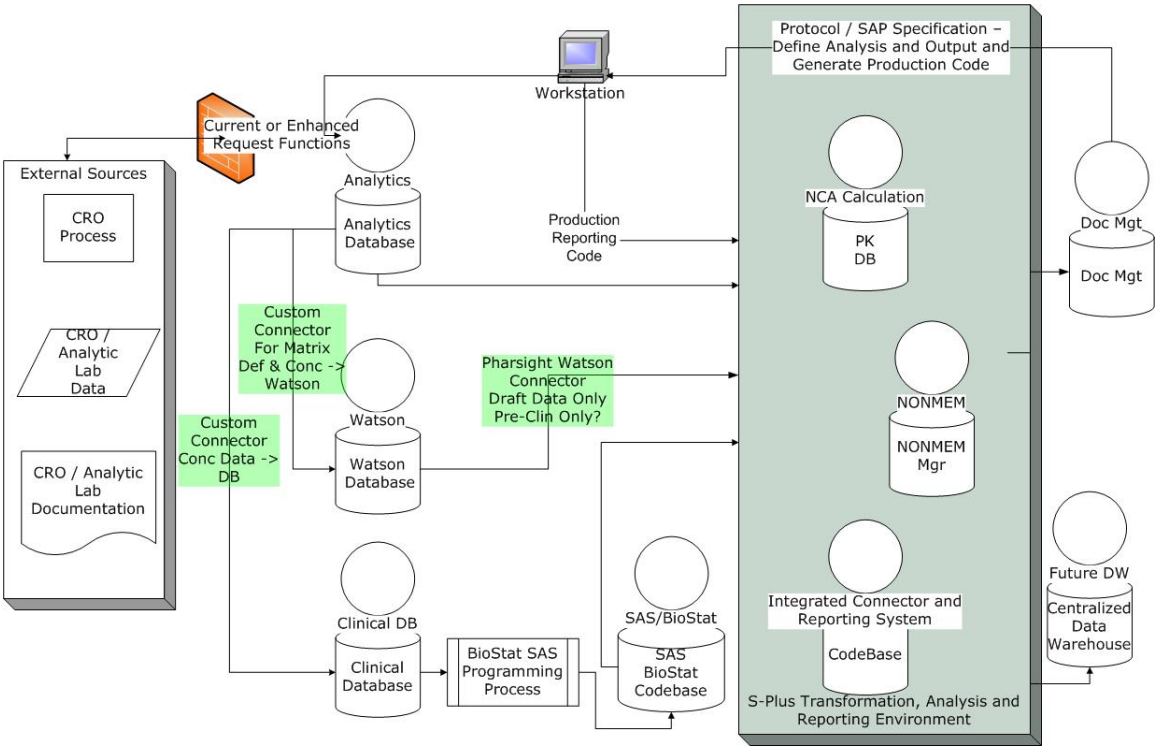

Figure 9: Alternative systems used for PK/PD and potential compliance risk

GRAPHICS OUTPUT SPECIFICATION APPLICATION

The S-PLUS environment facilitates reporting of key results, in the form of publication-quality tables and graphs, into RTF documents where one can add narrative text and automatically adhere to standard styles and formats. Automation of reporting using S-PLUS to produce reports (e.g. in MS Word) including S-PLUS tables and graphs includes the following benefits to the organization:

- Creates considerable productivity gains,
- Nails down workflows so as to avoid errors in cut-and-paste software such as Excel,
- Complies with 21 CFR 11/GCP guidelines,
- Enforces standard report styles and formats

By using the S-PLUS environment, many of the tasks in the workflow can be controlled by a menuing system that provides a central location for users to control the workflow. Providing central workflow control standardizes the process of creating the tables and figures and thus deploys best practices throughout the organization.
Some key benefits expected from the standardization and automation are as follows:

- Transparent tables and figures (unified formats and standard parameter names) across studies/projects improve the consistency of information shared amongst teams.
- Publication-quality tables and figures are generated in a validated and reproducible manner
- A major time saving will be achieved. Pharmacologists can spend more time on data analysis and other exploratory work, instead of word-processing focused activities.
- Clinical study reports will be delivered faster to meet demanding timelines.

GENERAL OPTIONS

The General Options tab is where the user selects the data on which to report and selects what tables and figures to include in the report. All the tables and figures are selected by default. The user also selects the name and location for the output file on this tab. Other options on this tab enable the user to store items for use later. The Table Design group allows the user to save and restore the selections on the dialog tabs. This allows a user to design a report that they may want to use again in the future.

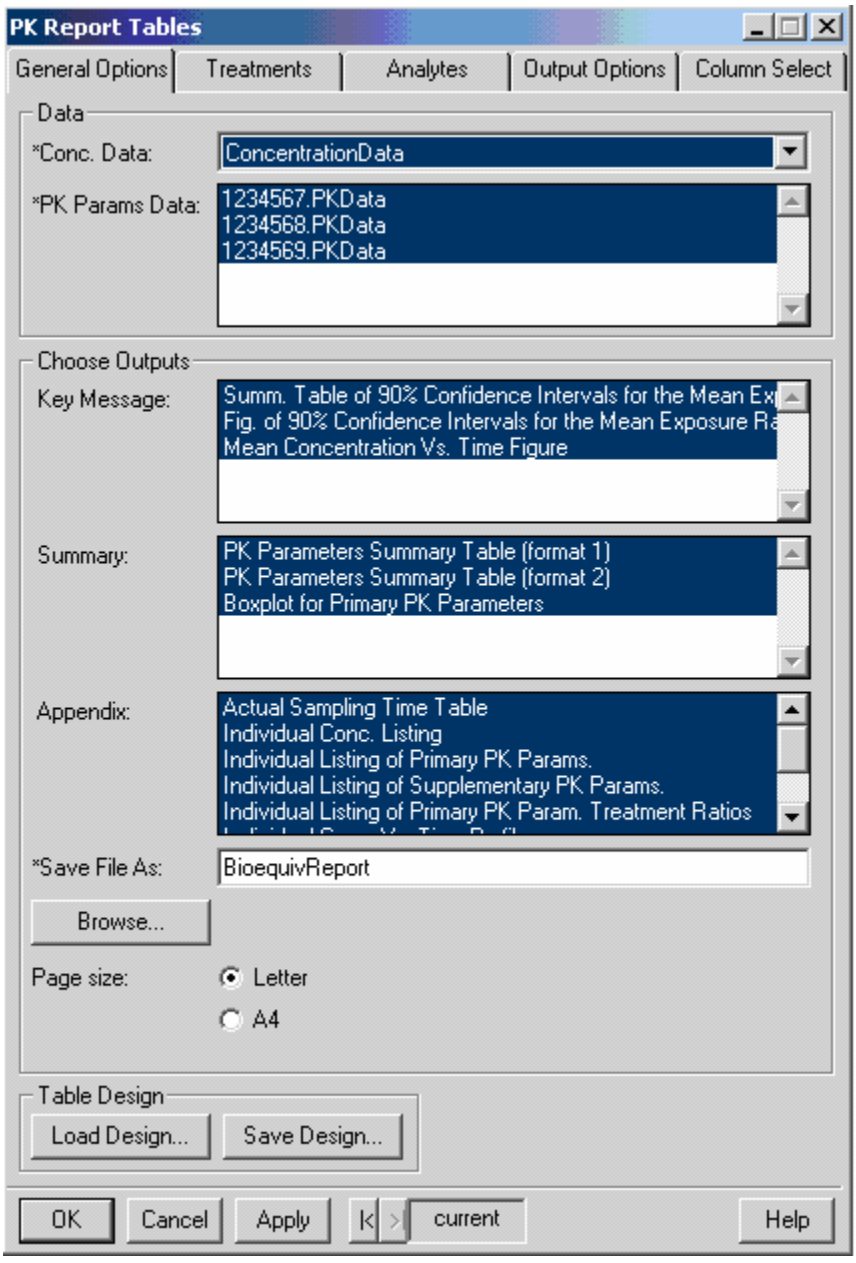

Figure 10: PK Report Dialog - General Options Tab

TREATMENTS

On the Treatments tab, the user specifies which treatments are reference and which are test. The user also can specify aliases for the treatments so that the tables and figures will not just display “Treat A” but can also have some descriptive information for Treat A, e.g. Drug A Capsule. In the Treatment Relations field, the reference treatment will be preceded by REF: and the test treatments will be preceded by TESTS.

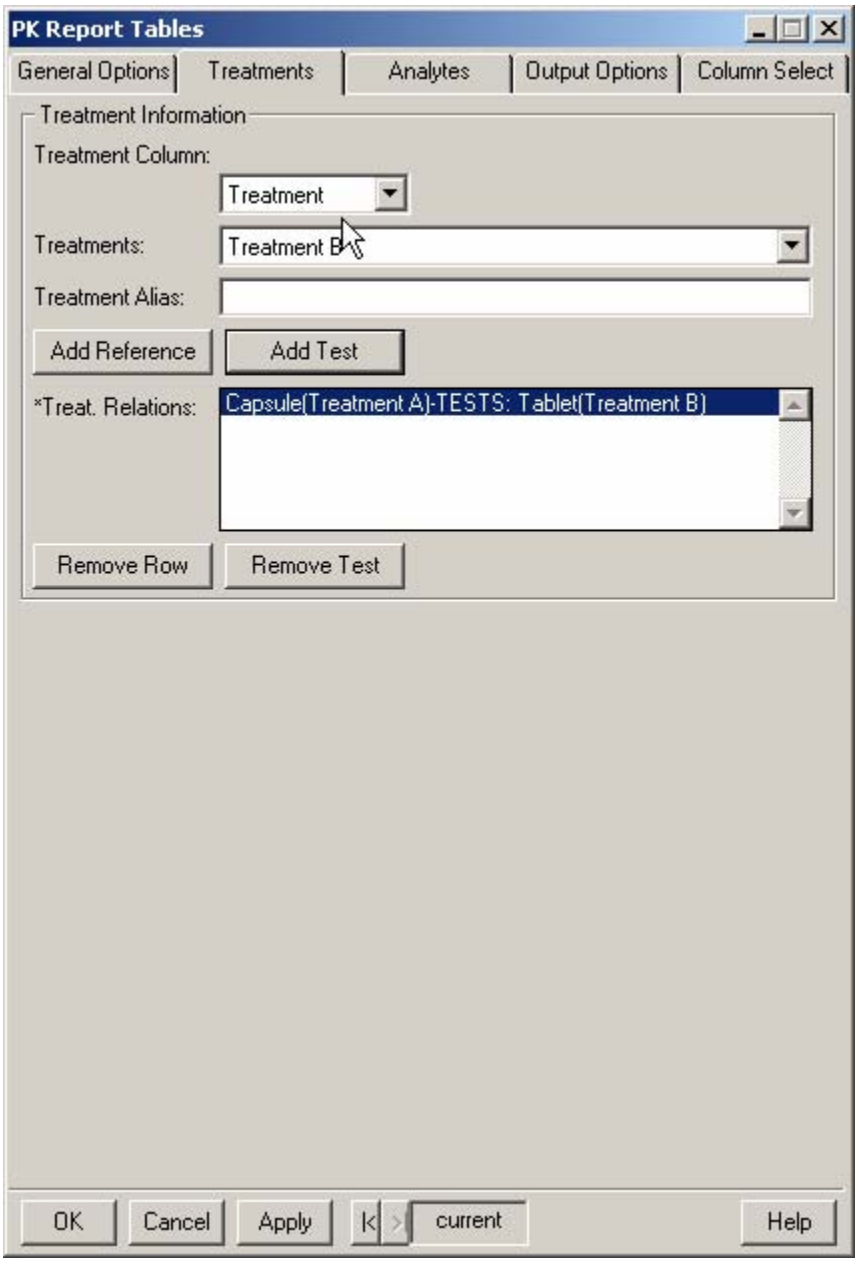

Figure 11 PK Report Dialog - Treatments Tab

ANALYTES/TITLE

On the Analytes/Title tab, the user specifies which columns have analyte concentration information and the aliases and treatment information for those analytes.

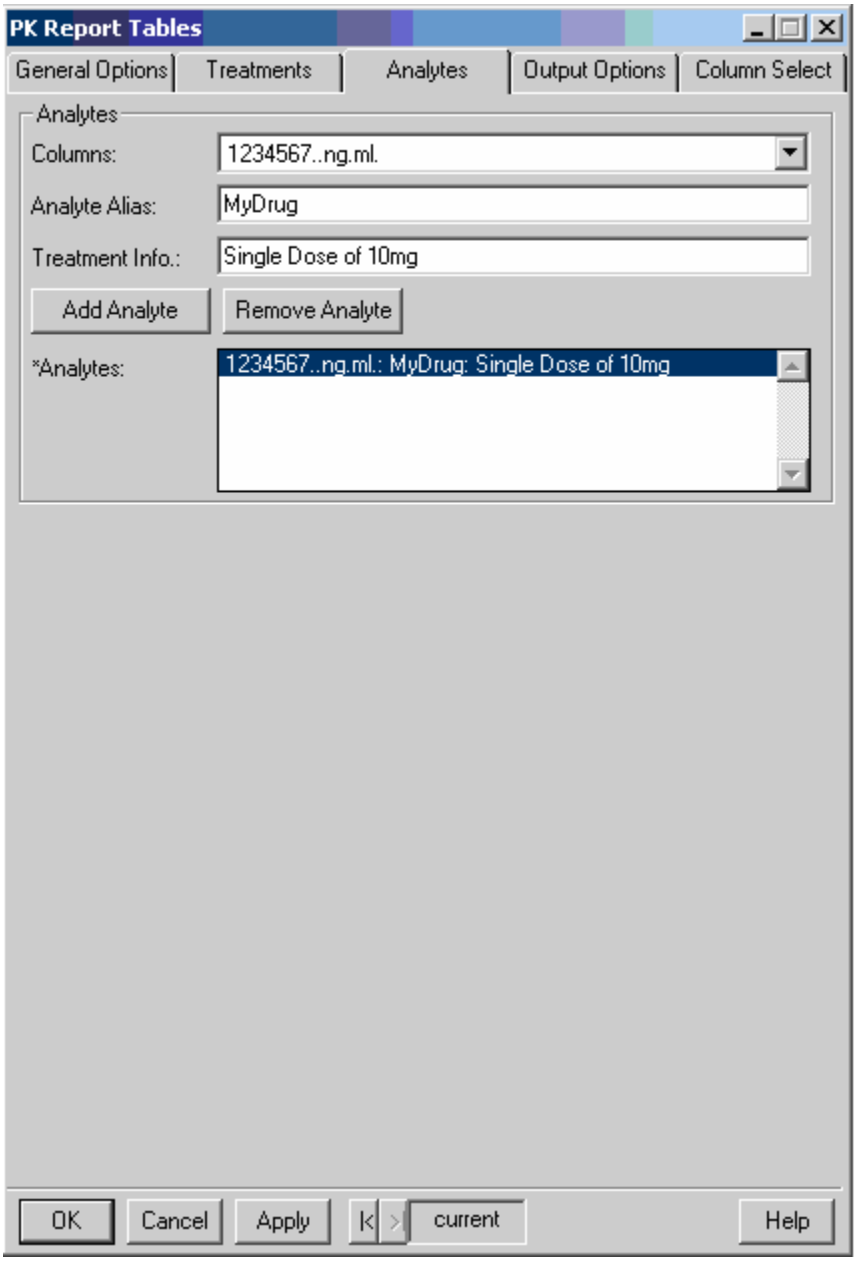

Figure 12: PK Report Dialog - Analytes/Title Tab

OUTPUT OPTIONS

The Output Options tab allows the user to modify the outputs for the selected tables and figures. In particular, the user can modify the PK parameters and summary statistics for the various tables and figures. The user also specifies whether the participants were subjects or patients; this determines the heading on tables, e.g. Patient versus Subject. The ability to Store Script allows a user to store an S-PLUS script that created the report. This script can be later run from the S-PLUS command line interface. The script will be stored to the selected file and can be run in any platform as long as the correct datasets are available.

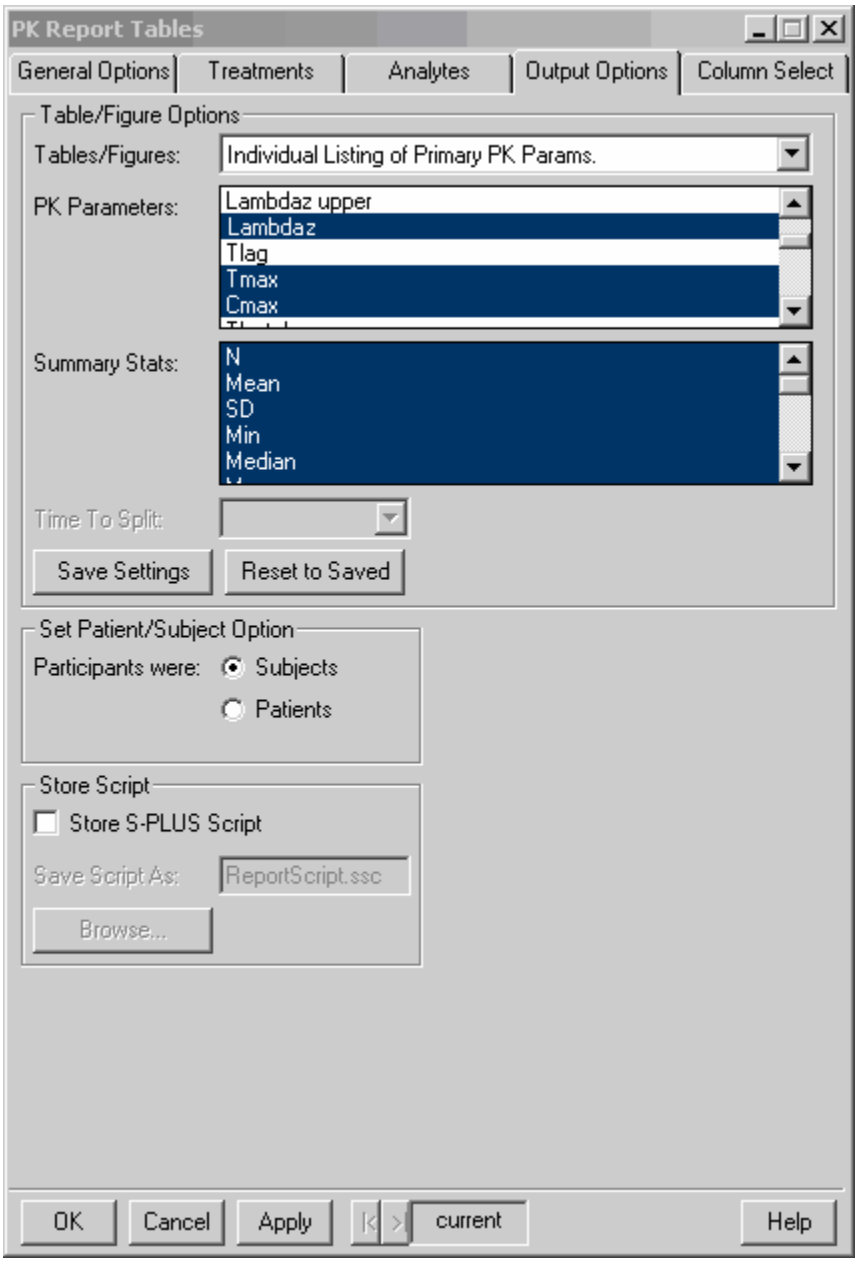

Figure 13: PK Report Dialog - Output Options Tab

COLUMN SELECT

If the data for the report do not have the standard columns for Patient, Treatment and Time values, then the user selects columns that should be used for those values. The defaults are:

- Concentration Data
  - USUBJID column for patient
  - Treatment column for treatment data
  - Relative.Actual.Time for relative actual time
  - Relative.Nominal.Time for relative nominal time
- PK Parameters Data
  - USUBJID column for patient
  - Treatment column for treatment data

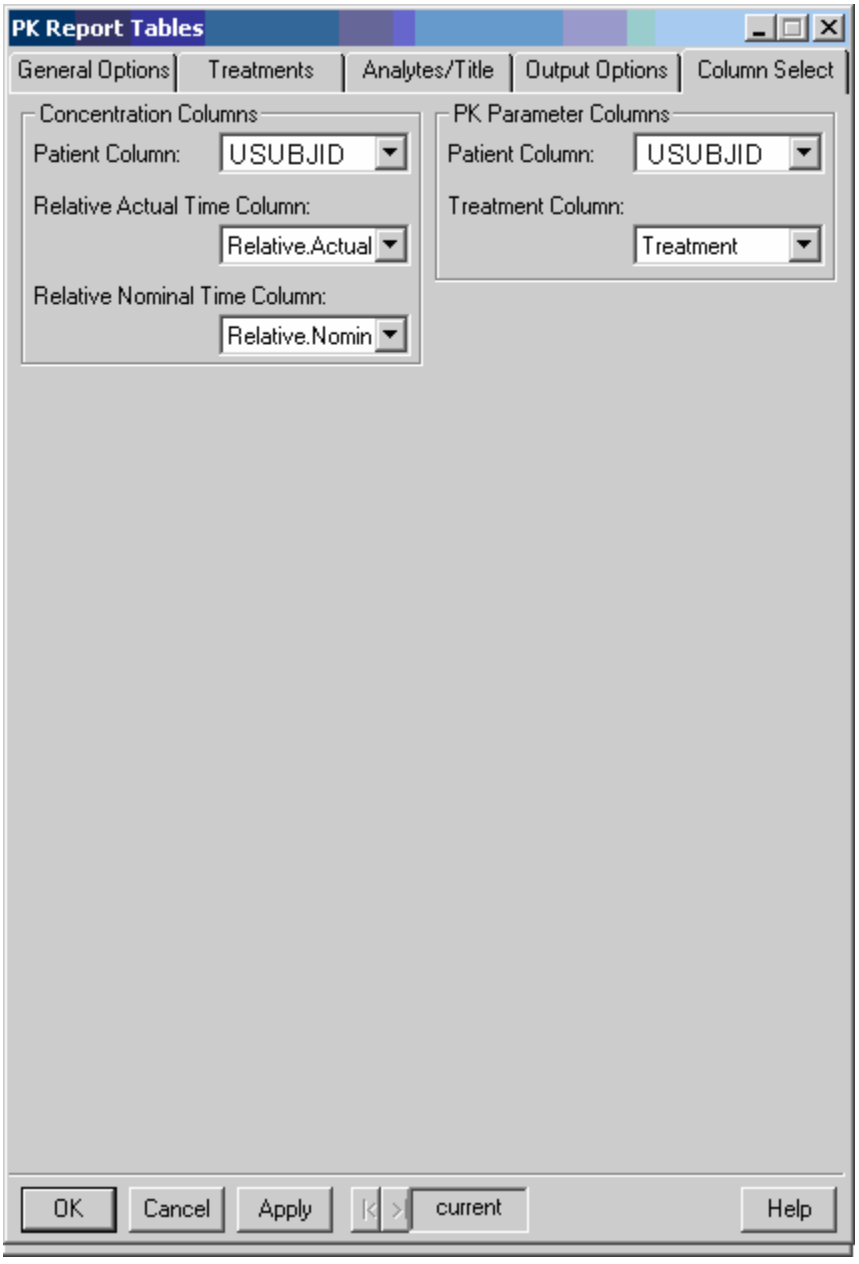

Figure 14 PK Report Dialog – Column Select Tab

CODE INTEGRATION

SAS users often would like to be able to access S-PLUS to take advantage of its sophisticated plotting capabilities or to implement a statistical method that is not available within SAS. At the same time, many pharma companies have a large user base familiar with SAS but fewer users with knowledge of S-PLUS. These SAS users may want to access S-PLUS functionality without having to learn the language. The S-PLUS® Clinical Pack for SAS® Users provides a method for SAS users to easily access the functionality of S-PLUS while still preserving a controlled computing environment and audit trail of activities.

The S-PLUS Clinical Pack for SAS Users provides a base level of functionality while demonstrating how to extend the concepts to any graph or statistical method desired. The provided family of macros demonstrates this capability by allowing a SAS user to generate nearly any type of S-PLUS TRELLIS graphic as well as invoking scripts created via the S-PLUS graphical workbench for Windows.

The S-PLUS Clinical Pack for SAS Users is implemented as a set of complementary SAS macros and S-PLUS functions. The S-PLUS Clinical Pack for SAS Users contains the following:

• SAS Application Macros

• SAS Helper Macros

• S-PLUS Library containing S-PLUS Functions

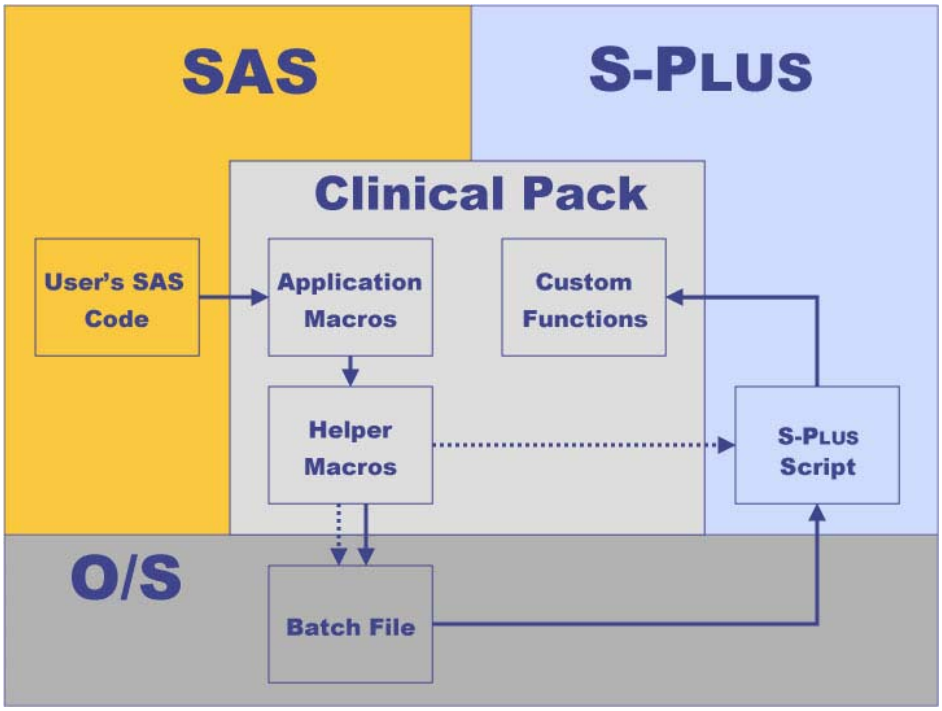

Figure 15: Overview of S-Plus Clinical Pack for SAS

It is simple for the user – you just have to execute a SAS macro within your familiar SAS environment. The macro can be executed in an interactive session, or in batch mode. You can even integrate the macro within another macro.

This macro (XYPLOTTRELLIS) makes a trellis graph in S-PLUS.

Submit a SAS macro call:

%XYPLOTRELLIS(

dataLibrary=mylib,

dataSet=data,

graphFormula = 'y ~ x | z')

Helper SAS macros write and execute batch S-PLUS job

S-PLUS creates plot

The arguments to the macro are the SAS data, the S-Plus session name, the plot formula as in S-PLUS, Title lines and the ODS html reportfile.

Note that there is no difference between a and “plot” and a “and datasets” (in S-Plus its just objects) – hence we could equally have sent a new data set to SAS. This option is very useful in cases where you want to do complex transformations (such as running means etc), but also in cases where you want to do matrix manipulations and S-Plus is an easier alternative than SAS or SAS/IML.

CONFIGURE THE ENVIRONMENT

There are a number of steps required in order for the S-PLUS Clinical Pack for SAS Users to work properly. SAS and S-PLUS must be accessible to the user along with the S-PLUS Clinical Pack for SAS Users SAS macro library and S-PLUS function library.

SAS

The SAS macros in the S-PLUS Clinical Pack for SAS Users must be accessible by the SAS executable. The file SPLUSMacros.sas is provided to include the source of each macro into a SAS session, although it is anticipated that most users will use an autocall macro library with the code of the macros under some form of revision control as part of a regulatory compliance process. Once the macros are available to SAS, a number of global macro variables needed by the S-PLUS Clinical Pack for SAS Users are set by a call to SPLUSENV.sas.

CONTROLLING OUTPUT

The original use of the S-PLUS Clinical Pack for SAS Users was to incorporate S-PLUS results into the SAS ODS. Supported targets include HTML and RTF. Note, however, that multiple, simultaneous ODS targets are not supported. Since that original release, the Clinical Pack has been expanded so that it can be used independently of SAS ODS. Although S-PLUS supports many different graphic file formats, the S-PLUS Clinical Pack for SAS Users provides support for JPG, PNG, WMF, PDF and PostScript. The format of the created graphic is controlled by the global macro variable graphFileFormat that is set in the SPLUSENV.sas macro.

The user can control the size of plot and margins, generation and location of log and output files (SAS and S-PLUS files, etc.) and other features via the macro parameters.

The S-PLUS Clinical Pack for SAS Users is designed to “insulate” the SAS user from S-PLUS; however, a basic understanding of the S-PLUS Clinical Pack for SAS Users interacts with S-PLUS is useful, especially for troubleshooting any problems, which may arise.

Figure 16, below, illustrates how the S-PLUS Clinical Pack for SAS Users calls S-PLUS in batch mode. While the S-PLUS script is running, it depends on a custom function library.

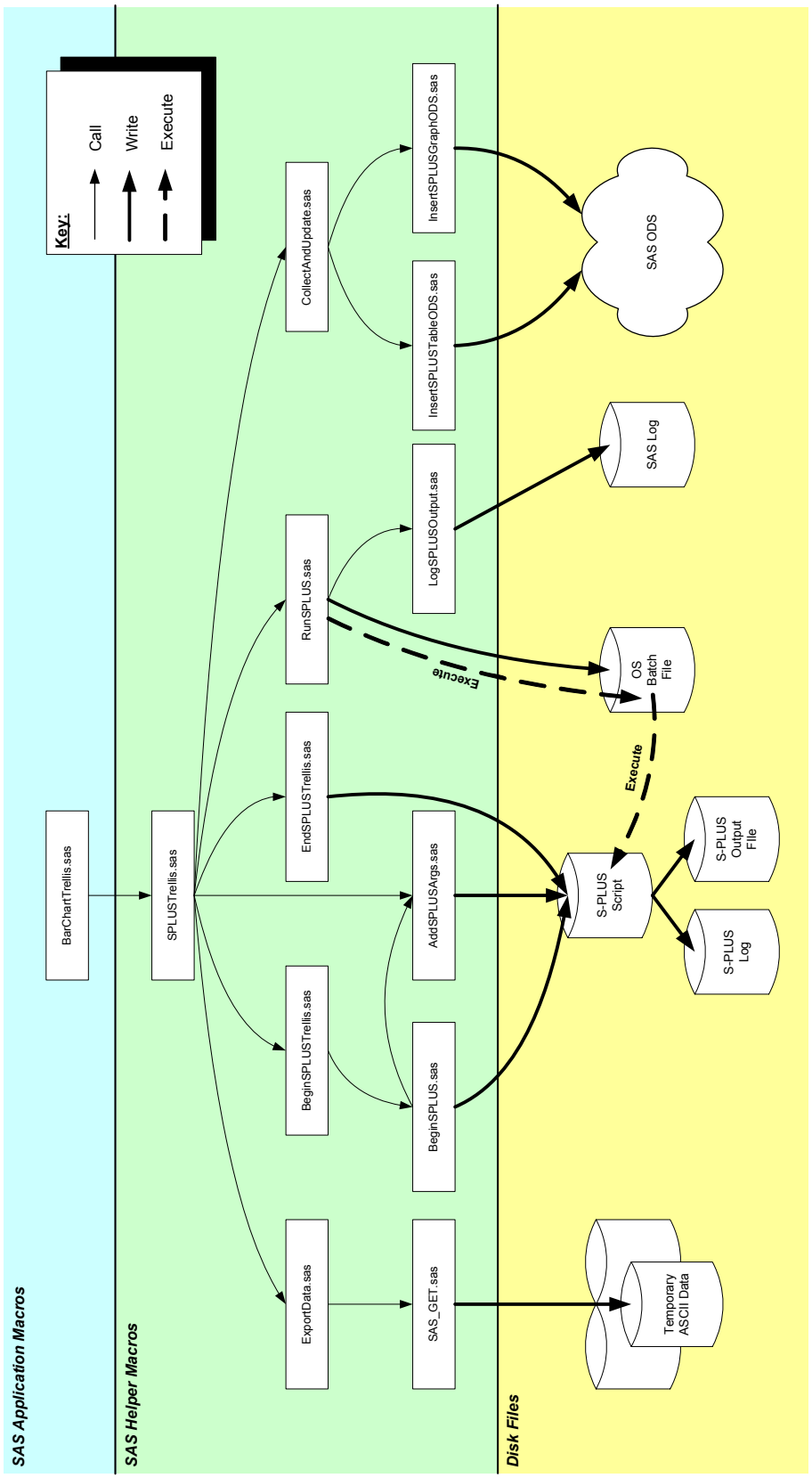

FIGURE 16: Trellis Macro Call Diagram

SAS APPLICATION MACROS

The SAS Application Macros can be thought of as a programming interface to the Clinical Pack functionality. SAS users who only wish to employ S-PLUS Clinical Pack for SAS Users functionality need only call these macros. These macros call other SAS macros that provide the link between SAS and S-PLUS.

SAS Application Macro calls, supplied with the appropriate parameters, are entered by a programmer within a SAS job. These macros automatically invoke SAS Helper Macros and S-PLUS Helper Functions to invoke S-PLUS functionality from within the SAS environment. At run time, SAS Application Macros invoke SAS Helper Macros to create an operating system (O/S) batch file and S-PLUS script file(s). The batch file is then started via a call to the O/S to execute the S-PLUS script, which produces a log file and any requested output. Additionally, the SAS Helper Macros will include the results of the S-PLUS script in the ODS target, if desired. (NB: During the presentation, example code will be shown.)

SAS HELPER MACROS

The SAS Helper Macros are not called directly by the user but are called by the SAS Application Macros. These Helper macros handle the details of creating and executing the O/S batch files and S-PLUS script files.

S-PLUS HELPER FUNCTIONS

The S-PLUS Helper Functions provide a library of complementary S-PLUS for the SAS Application Macros.

AUDIT TRAIL

The macros and functions of the Clinical Pack produce an audit trail consisting of S-PLUS log files and statements inserted into the SAS log. In addition, the location of all O/S batch files and scripts can be specified by the user to ensure conformance with applicable procedures and controls.

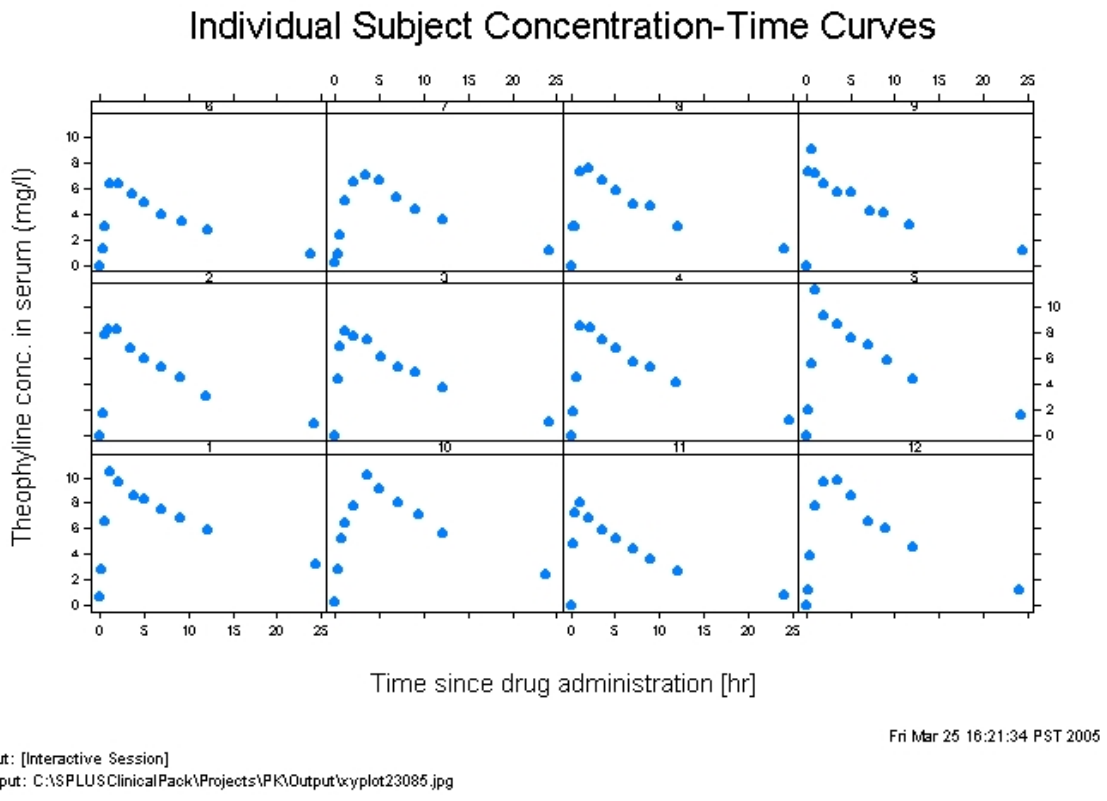

Figure 16: Example Output trellis figures of concentration data

SUMMARY

PK studies are a critical part of the drug development process, and improved efficiencies will enable the scientists to concentrate on their core job, which is understanding the actions of a drug, rather than being bogged down by the administrative aspects of PK analysis. In the long run, that can only be beneficial to an industry battling regulatory demands, competition, time constraints and shrinking budgets.

In addition, the saving associated with the direct generation of graphing code by the PK Scientist frees both the scientist and the programming support form the mundane and repetitive tasks of producing the production graphics.

Finally, the simple path for integrating this process into the overall clinical analysis and reporting environment enables the use of this technology without significant changes to other processes and systems.