# Managing Clinical Lab Data With SAS/AF®
## - an Online Demonstration

by Eric Brinsfield and Brad Klenz
Meridian Software, Inc.

**Introduction**

Increasingly, pharmaceutical research organizations find the growing volume of incoming lab test results challenging to manage.  In this presentation, we identify the characteristics of laboratory test data that complicate lab data processing and propose steps to simplify the situation.  In an on-line demonstration, we present a SAS® software-based solution that can alleviate most problems associated with handling large volumes of laboratory data.  The demonstration will focus on the SAS/AF portion of the system.

**Current Situation**

Although laboratory test data have been an integral part of clinical research for many years, recent developments have made managing these data increasingly difficult.  The data are changing and the demands for those data are changing.

Changes in the data

As clinical research grows more complex and pharmaceutical companies and clinical research organizations expand, the volume of lab test data that must be managed by any organization also grows.  The data are swelling, not only due to increases in the number of simultaneous studies underway, but also because more types of lab tests are being provided. Computer storage space becomes a growing problem from a hardware standpoint, while data redundancy becomes a costly problem from the standpoint of staff efficiency and data quality.

To add to these developments, pharmaceutical companies can utilize a growing number of different central laboratories to provide the laboratory test results.  Although standardized file formats for lab data exist, each central lab may still provide its data in different formats.  Companies trying different central labs to find the best, or using multiple labs to increase throughput, must deal with multiple data input formats.

In addition, these diverse record formats may change over time.  Each new study may have unique requirements for the data, requiring a new format.

Changes in the requirements for data

Increasingly, clinical researchers want to see the lab data at an earlier stage in the data collection process.  Even though the data may not be completely validated, early review of the data may be critical during studies with life-threatening implications.  This also means that less sophisticated computer users will be trying to use the data without programmer assistance.

At the same time, the FDA is justifiably demanding greater quality and reliability of lab test data, while pharmaceutical companies try to obtain "clean" data faster and faster. In addition to validity, the FDA is requiring accountability. Changes to the data must be tracked and controlled.

The personnel responsible for managing these data hope to spend less time dealing with everyday data management and more time improving the process, but doing so is difficult as the volume of data increases. They need to automate repetitive tasks, so they can spend more time striving for quality.

**A SAS Software-based Solution**

Our SAS/AF application addresses many of these problems using Version 6 of the SAS System. The system was developed using SAS version 6.06 under OS/2. It has now been tested under OS/2 with SAS 6.06 and SAS 6.08, as well as under VMS with SAS 6.07. The code is designed to be portable.

The system guides the users through each task with a series of SAS/AF windows. Users have the choice of using the mouse, functions keys, and pmenus within the limits of their hardware.

The major subsystems of MERLIN are shown below:

Using DATA steps, procedures, Structured Query Language, and Screen Control Language, we were able to create a relational database of SAS data sets. As a result, we have reduced data redundancy and therefore space requirements and update management problems.

Lab results are stored vertically rather than horizontally. In other words, each lab test does not require its own variable. You can add or delete lab tests without restructuring the data sets. As a result, the application can handle an unlimited number of lab test codes, which can vary from study to study within the same database.

The system allows programmers to "plug in" additional import modules in order to handle new or added central lab file formats. The programmer simply programs the import module according to documented standards and instructions, then defines the module to the menu system using specific windows.

The application converts all import data into a standard format. As a result, changes in the format of import data files do not affect the master database and are transparent to researchers using the data. Each study can have a different set of lab tests and input file formats without affecting the main SAS/AF system.

Validation is performed as the data are imported. As a result, the quality of the data is improved at an earlier stage and data are available for preliminary analysis much sooner. Validation criteria can be established by study-specific personnel, eliminating backlogs for lab data management personnel.

SYSTEM DEFAULTS: DATA SET VARIABLES
Return=F3 Help=F1

<Value List or Range=F7>

| Row | Variable | Description | Validate? | Action |
|-----|----------|-------------|-----------|--------|
| 1 | ALERTFLG | Alert Flag | NO | ERROR |
| 2 | DELTAFLG | Significant Change Flag | NO | ERROR |
| 3 | EFFECTDT | Effective Date | NO | <RANGES> |
| 4 | ETERMFLG | Early Termination Flag | NO | ERROR |
| 5 | INVCD | Investigator Code | NO | WARNING |
| 6 | LABDT | Lab Date | YES | <RANGES> |
| 7 | LABTM | Lab Time | NO | <RANGES> |
| 8 | LABTYPE | Update Type from Lab | NO | ERROR |
| 9 | LABVISIT | Lab Visit Identifier | NO | ERROR |
| 10 | LWRC | Lower Range - Character | NO | <RANGES> |
| 11 | LWRN | Lower Range - Numeric | NO | <RANGES> |
| 12 | OCCUR | Occurrence | NO | <RANGES> |
| 13 | PATDOB | Patient Date of Birth | NO | <RANGES> |

At any point during a study, users can access the data in several ways. If authorized, users can browse the data online through a variety of paths. They can define personalized subsets of the data for generating reports or exporting for use on other systems.

Select report(s).                                    <Run Interactive>
Then choose a run option<Run Batch>
(interactive or batch).<View Selections>

| Report | Description | # of Each |
|---|---|---|
| ATSUM | Audit Trail Activity Summary | 1 |
| ATDETAIL | Audit Trail Activity Detail | |
| LABRPT | Lab Report - By Central Lab, Protocol and Patient2 | |
| MASTSUM | Master Data Set Summary | |
| MASTSTDY | Master Data Set Summary by Study | |
| CONTENT | Merlin Data Set Contents Listing and Cross-Reference | |
| VSUMALL | Validation/Import Summary - All Labs and Studies | |
| VSUMSTDY | Validation/Import Summary - By Study | |

In addition, the system administrator can define SAS views to the application, so users can read data directly from the master database, without extensive knowledge of its structure.

Although it is assumed that most updates to the lab data arrive from the central labs, the system administrators can determine which variables in the database may be updated by authorized users. Some variables, such as visit numbers and lab test codes, need two sets of values. One set is typically used by the central lab, while the other is used by the pharmaceutical company. The application automatically updates the pharmaceutical company's codes as long as administrators provide the initial mapping.

Protocol code: JUP565                    <Apply=CTL A>
                                         <Get Unmapped=CTL U>

| Row | Lab Visit # | Unscheduled Visit Flag | Early Term. Flag | Lab Period ID |
|---|---|---|---|---|
| 1 | VISIT01 | 0 | __ | V1 |
| 2 | VISIT01 | 1 | __ | V1.1 |
| 3 | VISIT04 | 0 | __ | V4 |
| 4 | VISIT04 | 1 | __ | V4.1 |
| 5 | VISIT07 | 0 | __ | V7 |
| 6 | VISIT10 | 0 | __ | V10 |
| 7 | VISIT13 | 0 | __ | V13 |

The application also provides a complete audit trail facility, for both batch and interactive updates. Additions and deletions are tracked during the background import process. If interactive updates are permitted, each modification is logged. As a result, menu system also provides a "rollback" facility that enables administrators to undo changes that occurred after a certain point in time.

Audit information:     User ID:  BWK        Observation number:  1
                       Date:     17MAY93    Type of update:      C
                       Time:     18:19      Type of processing:  I

Observation changed:

Source ID:        LABRES
Study code:       JUP565
Specimen ID:      L09008

Variable changed:    LABPERID
Type of variable:    CHARACTER
Previous value:      _____
Current value:       V1-0

The application provides security at the menu level and the variable level. The functions available to given users are controlled by the main menu they can access. This access is determined by the administrator.

In addition, the administrator can define an unlimited set of security groups. Each of these groups are associated with a fixed list of variables that the members of that group can update. Any variable that can be modified is automatically tracked by the audit trail system.

To reduce the workload of the lab data management staff, the system can nightly and automatically import any new lab data that have arrived. The import process can be initiated on demand, if necessary. Ideally each central lab logs into the pharmaceutical company's system and uploads new data without requiring assistance from lab data management personnel. Selected reports execute automatically as the result of a nightly import and e-mail transmissions notify selected individuals of the status of each import.

Highlight report(s) to run automatically by selecting the row number:

| Row | Report Information |
| --- | --- |
| 1 | Title:      Validation/Import Summary - All Labs and Studies |
|   | Output to: _____ |
| 2 | Title:      Validation/Import Summary Report - By Study |
|   | Output to:  c:\merlin\merout\valsum.lis |
| 3 | Title:      Validation/Import Summary - By Lab |
|   | Output to: _____ |
| 4 | Title:      Validation Detail Report |
|   | Output to:  e:\merlin\merout\vdetail.lis |
| 5 | Title:      Validation Detail: Rejections/Warnings - by Study |
|   | Output to:  e:\merlin\merout\vrejwrn.lis |
| 6 | Title:      System Validation Criteria Summary Report: Variables |
|   | Output to:  e:\merlin\merout\vcrit.lis |

Select a topic:

Accessing online documentation
Accessing other user menus
Adding a new study
Adding help index topics
Adding reports and other programs
Administering audit trails
Administering master files and audit trails
Archiving audit trail data
Assigning users to security groups
Assigning validation criteria (administrators)
Assigning validation criteria (non-administrators)
Browsing master files (administrators)
Browsing master files (non-administrators)
Creating audit trails
Creating subset definitions
Defining and browsing data views (administrators)
Defining automatic validation reports

To complement the menu-driven approach to data management, the application provides an extensive help system. You can access help at the point of need for the task or the variable. You can browse the online documentation or use a help index to branch to any topic from any point in the system. To speed movement through the system, you can jump from any branch in the menu system to the bottom of another branch.

The following figure illustrates a portion of the application's help index.

**Summary**

We identified several trends relating to laboratory test data and proposed data processing practices that can be utilized to respond to those trends. We demonstrated a lab data management system that was developed using SAS software. Although the system is designed to handle lab test data, its data-driven approach can be applied to any data management application.

Version 6 of the SAS system enables SAS programmers to construct a flexible menu-driven system. Through Screen Control Language, windows can access multiple SAS data sets simultaneously, permitting the use of relational data structures.

SAS is a registered trademark of SAS Institute Inc., Cary, NC.